universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Calculation of the vertical net energy flux using an energy budget equation for the land surface

verfasst von / submitted by

Vanessa Nicole Seitner, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 614
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Meteorologie
Betreuer/ Supervisor:	Univ.-Prof. Mag. Dr. Leopold Haimberger

Contents

List of Tables

List of Figures

Abstract

In this thesis different calculation methods for the vertical net energy flux through the Earth's surface are discussed. Three different methods are used for this purpose, firstly the direct calculation from the net radiation and the turbulent fluxes as they can be downloaded from reanalyses like ERA5. Another calculation possibility is the calculation over an atmospheric column. The third method is the diagnostic evaluation of an equation for the description of the soil heat flux over the land surface.
At first the necessary theoretical foundations are explained. Afterwards the data and methods used are described. This provides the basis for comparing the different methods of calculating the vertical net energy flow.
A comparison of the calculation methods and a discussion of the differences is of great interest. For this purpose, the land surface from 40 °N to 90 °N is used as a study area. Since this is a very large area, individual sub-areas are also dealt with. Over land surfaces the consistency between the net vertical energy fluxes is reasonable on the seasonal time scale. Inconsistencies emerge in high surface regions and near steep orography. The land surface equation as it is presented here does not seem to work over ice sheets, likely because of different data representation there.

Zusammenfassung

In dieser Arbeit wird auf unterschiedliche Berechnungsmethoden für den vertikalen Nettoenergiefluss durch die Erdoberfläche eingegangen. Dafür werden drei unterschiedliche Methoden verwendet, zum einen die direkte Berechnung aus der Nettostrahlung und den turbulenten Flüssen wie man sie in Reanalysen wie ERA5 herunterladen kann. Eine weitere Berechnungsmöglichkeit stellt die Berechnung über eine atmosphärische Säule dar. Die dritte Methode ist eine diagnostische Auswertung einer Gleichung für die Beschreibung des Bodenwärmestromes über der Landoberfläche.
Zuerst werden die dafür benötigten theoretischen Grundlagen erläutert. Im Anschluss daran werden die verwendeten Daten und Methoden beschrieben. Dadurch ist die Basis für den Vergleich der unterschiedlichen Berechnungsmethoden für den vertikalen Nettoenergiefluss geschaffen.
Ein Vergleich der Berechnungsmethoden und eine Diskussion der Unterschiede ist von großem Interesse. Dafür wird als Untersuchungsgebiet die Landoberfläche von 40 °N bis 90 °N herangezogen. Da dies einen sehr großen Bereich darstellt wird auch auf einzelne Teilbereiche eingegangen. Über Landoberflächen ist die Konsistenz zwischen den vertikalen Nettoenergieflüssen auf der jahreszeitlichen Zeitskala vernünftig. Inkonsistenzen treten in höher gelegenen Gebieten nahe steiler Orographie auf. Die Gleichung für die Landoberfläche, wie sie hier vorgestellt wird, scheint über Eisschilden nicht zu funktionieren, wahrscheinlich wegen der unterschiedlichen Datenrepräsentation dort.

1. Introduction

1.1 Continental Heat Gain

According to Mayer et al. (2017) energy budgets are an important tool to denote the mean state, the variability and the trends in the climate system. To determine whether the climate system is in balance, the fluxes at the top of the atmosphere are considered.
Fig. 1 shows the global mean energy budget.

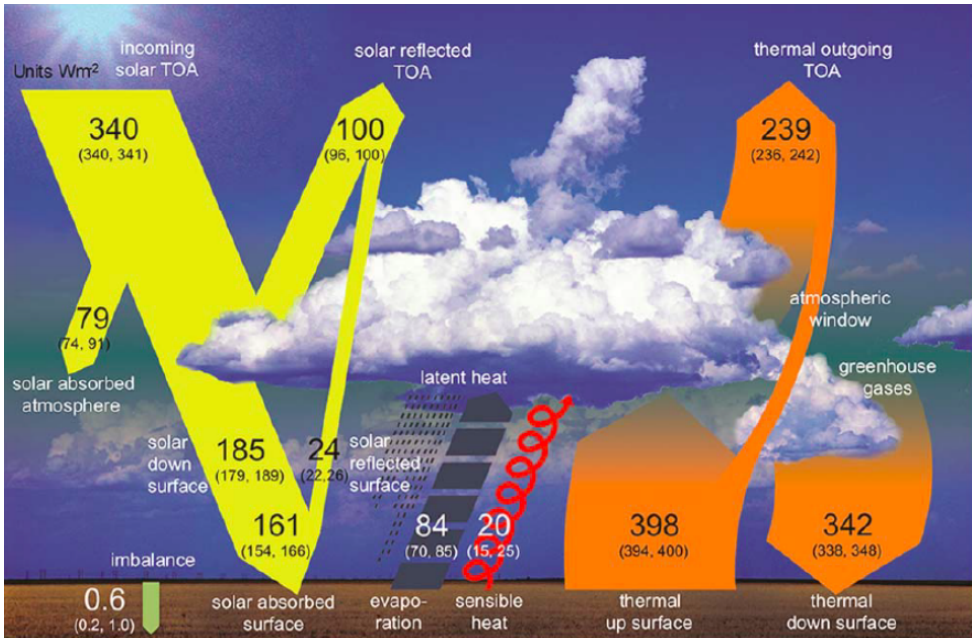


Figure 1: Global mean energy budget with magnitudes of energy fluxes in $W\,m^{-2}$ with uncertainty ranges in parentheses. (IPCC, 2013)

The incoming solar radiation at the top of the atmosphere (TOA) amounts to 340 Wm^{-2}. If the solar reflected radiation having a value of 100 Wm^{-2} is subtracted, 240 Wm^{-2} are resulting, which roughly corresponds to the value of the thermal radiation emitted. As known from IPCC (2013), the Earth's climate is warming and there must be a net flux of energy into the climate system, which goes into the surface, called imbalance in Fig. 1. If the global imbalance of 0,6 Wm^{-2} is considered, a closure of the budget follows. (IPCC, 2013)

Based on Beltrami et al. (2002), the imbalance leads to an increase of the heat contents within oceans, atmosphere and cryosphere. The continental lithosphere is also a major component of earth's climate system with regard to heat and moisture transports between surface and atmosphere. Table 1 shows the heat gained by ocean, atmosphere, components of cryosphere and lithosphere within the time range given.

The values in Table 1 show that the heat gained by the lithosphere is in the same order of magnitude as atmosphere and components cryosphere. Although it is an order of magnitude smaller than that of the ocean, it is not negligible and should be taken into account in the determination of the total heat gain of the climate system. (Beltrami et al., 2002)

Sphere	Time	Heat Gain in Joule
Ocean	1955-1996	$182,0 \cdot 10^{21}$
Atmosphere	1955-1996	$6,6 \cdot 10^{21}$
Continental glaciers	1955-1996	$8,1 \cdot 10^{21}$
Antarctic sea ice	1950-1970	$3,2 \cdot 10^{21}$
Lithosphere	1950-2000	$9,1 \cdot 10^{21}$

Table 1: Heat gained by ocean, atmosphere, components of cryosphere and lithosphere in J *in the indicated time period. (Beltrami et al., 2002)*

Similar values to Table 1 can be found in Church et al. (2011): In a period from 1972 to 2008, $193,0 \cdot 10^{21}$ J were stored by the ocean, $2,0 \cdot 10^{21}$ J by the atmosphere and $4,7 \cdot 10^{21}$ J were stored by continents. Also in von Schuckmann et al. (2016) can be found that the ocean is the major heat reservoir with a storage of 93 % of the imbalance, 4 % of the imbalance go into ice surfaces and the remaining 3 % lead to a warming of atmosphere and land surfaces.

1.2 Aim of the work

There are lots of papers which treat the heat gain in the ocean, but only a few that consider the heat gain in the lithosphere.
Due to the fact, that the heat gained by the lithosphere is not negligible, it is desirable to have a closer look at the energy balance over the land surface.
An accurate representation of the land surface budget is also required for medium range and seasonal prediction of land surface temperature and precipitation.
The aim of this thesis is to provide an energy budget equation for land surface for calculating the vertical net energy flux. This calculated energy flux is therefore compared to the net energy flux parameterized in the ERA5 reanalyses of ECMWF. In addition to that, a comparison is also made to the indirectly determined net energy flux from the atmosphere.
All methods contain one or the other inaccuracy, which leads to differences in the calculation of the vertical net energy flux. The differences have to be discussed to evaluate the budget equation of the land surface, if a good estimate of the vertical net energy flux over land is provided.

1.3 Period and area under consideration

The data used is considered for a time period from 2000 to 2018. In order not to lose information through the calculation of tendencies (explained in section 4.4), the data from December 1999 to January 2019 are downloaded. A reason for choosing 2000 as the starting date is that the TOA net radiation fluxes from CERES (Loeb et al., 2009), which are used for calculating the vertical net energy flux by an atmospheric column, are only available from 2000.
The data is downloaded globally, in section 5 also individual cutouts are treated separately.

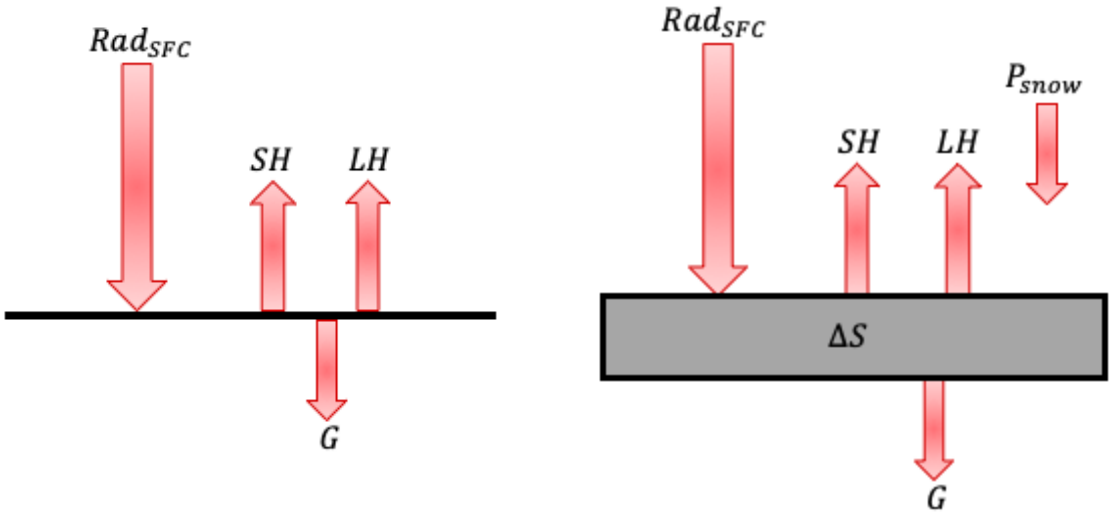
2. Theoretical background

2.1 Simplified energy budget of surface

A starting point for the considerations explained below is the simplified energy budget at Earth's surface. This includes net radiation on surface (Rad_{SFC}), latent heat flux (LH), sensible heat flux (SH) and ground heat flux (G). A slightly more detailed description of the components is given in sections 2.1.1, 2.1.2 and 2.1.3.

The surface gains energy through Rad_{SFC} and loses energy through LH, SH and G. This leads in total to the formula expression given in equation 1. The terms in equation 1 are energy fluxes and are expressed in Wm^{-2}.

$$Rad_{SFC} = -LH - SH - G \tag{1}$$



(a) Simplified energy budget.

(b) Simplified energy budget with storage (ΔS) and snowfall (P_{snow}).

Figure 2: Energy budgets of Earth's surface. Based on Kraus (2008)

Fig. 2 visualizes the initially considered options of the energy balance on the surface. The sign convention is chosen in a way that Rad_{SFC} is calculated as positive downward and that fluxes from the surface are considered as negative. (Kraus, 2008, p.33)
In Fig. 2a the visualisation of equation 1 is given, Fig. 2b shows an improved version of Fig. 2a and will be discussed in section 2.1.4.

2.1.1 Net Radiation

The net radiation (Rad_{SFC}) is the sum of the net solar radiation ($Rad_{SFC_{solar}}$) and the net thermal radiation ($Rad_{SFC_{solar}}$), as shown in equation 2.

$$Rad_{SFC} = Rad_{SFC_{solar}} + Rad_{SFC_{thermal}} \tag{2}$$

$Rad_{SFC_{solar}}$ is the sum of incoming (positive) and outgoing (negative) short wave radiation on surface, $Rad_{SFC_{thermal}}$ is calculated as the difference of incoming and outgoing long wave radiation on surface. (Kraus, 2008, p.128f)

2.1.2 Latent and sensible heat flux

The latent heat flux (LH) and sensible heat flux (SH) are turbulent fluxes and can be expressed as in equation 3.

$$LH = \frac{L_v}{g}\overline{q'\omega'} \qquad SH = \frac{c_p}{g}\overline{T'\omega'} \tag{3}$$

Both fluxes emerge in the atmospheric boundary layer directly above the surface. LH results from evaporation, denoted by $\overline{q'\omega'}$, the vertical turbulent flux of moisture. Heat conduction leads to the formation of SH, expressed by $\overline{T'\omega'}$, the vertical turbulent flux in association with variations in temperature. (Hantel, 2013, p.230)

2.1.3 Ground heat flux

The ground heat flux G is the sum of the energy flux densities through the considered surface. The material can be conventional soil but also rock, sand, water, snow and ice.
Two transport mechanisms are important in solid ground, on the one hand heat conduction, that means the energy transport between nearby particles. On the other hand electromagnetic radiation, the transfer of energy by free electrons.

The ground heat flux directly under the surface can be described by its effect in the soil. This is composed of the achieved heating and possible phase changes Z as shown in equation 4.

$$G = \int_{SFC}^{BOT} \rho_s c_s \frac{\partial T_s}{\partial t} dz + Z \tag{4}$$

In equation 4, the upper limit of the integral is denoted by SFC, which means the land surface. BOT is the bottom of the land column and may be designated as the depth, where the ground heat flux can be assumed as zero. There is a flux from within the Earth in volcanically active areas, but in general the flux is neglectable. A positive sign of G leads to a warming of the soil and snow and ice will melt. If such a phase transition occurs and the resulting water runs off, an ablation of snow and/or ice is the consequence. The ablation is effected by evaporation and melting, the latter combined with the latent heat of fusion L_f. (Kraus, 2008, p.129f)

2.1.4 Modification of simplified energy budget

A more precise formulation of the energy budget can be obtained by inserting a storage layer ΔS and a term for snowfall P_{snow}, as shown in Fig. 2b. ΔS can represent for example, radiation between leaves, plants or buildings. The net effect of these terms is summarized in ΔS. (Stull, 1988, p. 253ff)

Another effect that can be taken into account is the snowfall (P_{snow}). This term represents the cooling of the surface in consequence of falling snow and will be important in the following work. (Mayer et al., 2019)

The mathematical formulation of the improved energy budget can be taken from equation 5.

$$Rad_{SFC} = -LH - SH - G + P_{snow} + \Delta S \tag{5}$$

2.2 Order of size

Fig. 3 shows the annual mean global energy balance. Based on the values in Fig. 3 the budget at the top of the atmosphere (TOA) is closed. The net radiation at Earth's surface is the combination of global radiation plus long wave downward radiation minus reflected radiation minus long wave upward radiation. As a result a value of $-104\,\mathrm{Wm}^{-2}$ is gained, the radiation budget of the atmosphere alone is negative. So a value of $104\,\mathrm{Wm}^{-2}$ has to be added by the turbulent fluxes SH and LH to close the budget. (Brasseur et al., 2005)

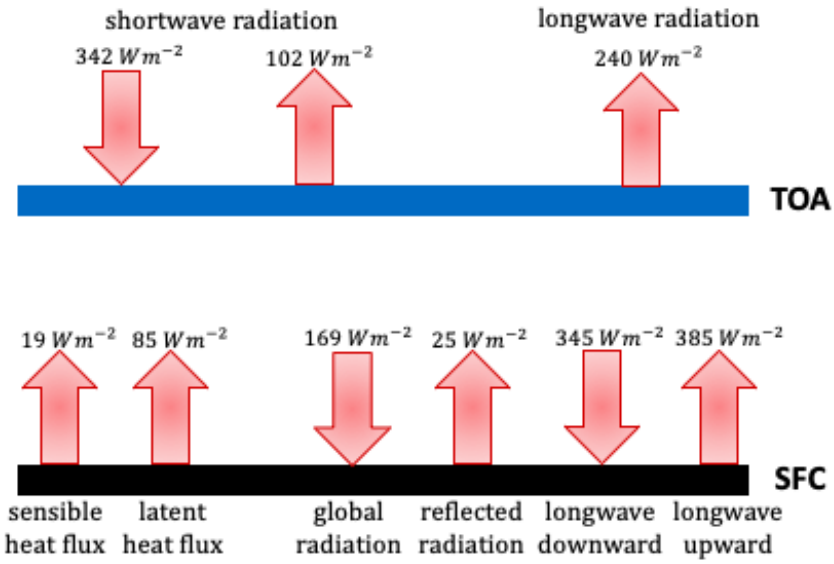


Figure 3: Annual mean global energy balance. Based on Brasseur et al. (2005)

2.3 Energy budget of atmospheric column

2.3.1 General information on enthalpy

The enthalpy h corresponds to the heat required to raise the temperature from zero to any value in Kelvin at constant pressure. If the temperature is zero, the enthalpy is assumed to be zero.
Enthalpy is typically not given in absolute terms. Instead differences from a given reference state e.g. liquid at 0 °C are considered.
For an ideal gas the enthalpy may be given as $h = c_p T$. General gases are not ideal but in the absence of phase transitions under atmospheric conditions the following differential equation holds.

$$dh = c_p dT \tag{6}$$

In equation 6 dh is the differential of enthalpy, c_p the specific heat capacity at constant pressure with a values of $1005\,\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$ and dT the differential of temperature.

If a layer of air is heated in hydrostatic balance, e.g. by radiation heating, the weight of the air above it remains the same as the air pushing down. So the heating process is performed at constant pressure. The energy which is delivered to the air is indicated by an increase in enthalpy. Enthalpy is often also called sensible heat. (Wallace and Hobbs, 2006, p.75f)
The latent heat of vaporization L_v is the difference between the (high) specific enthalpy of water vapour and the (small) specific enthalpy of water. (Hantel, 2013, p.98f)

2.3.2 Derivation of the atmospheric moist static energy equation

This section is strongly based on Mayer et al. (2017), where the derivation below and also more detailed explanations can be found.

In the climate system, water is available in all three states of aggregation. For the consideration of the atmospheric energy budget including enthalpy transports associated with water, the definition of a reference temperature is necessary. The reference temperature T_{00} is 273,15 K.
We consider the atmosphere as a multiface system, it consists of dry air, water vapour, liquid water and ice. In the following, m is used for definition of the moist static energy as given in equation 7.

$$m = (1-q_g-q_l-q_s)c_p(T_a-T_{00})+(q_l+q_g+q_s)c_w(T_a-T_{00})+L_v(T_a)q_g-L_f(T_a)q_s+(q_l+q_g+q_s)h_0+\Phi \tag{7}$$

In equation 7 q_g is the specific water vapour content, q_l the specific liquid water content and q_s the specific ice content in kgkg^{-1}. The specific heat of dry air at constant pressure is expressed by c_p ($1003\,\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$), c_w is the specific heat capacity of liquid water ($4190\,\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$). T_a is the air temperature in K and T_{00} a reference temperature. The latent heat of vaporization is denoted by L_v and depends on the air temperature T_a, but a value of $2.501\cdot10^6$ Jkg^{-1} is assumed. Furthermore h_0 is the specific enthalpy of water at 0 °C, as shown in equation 8.

$$h_0 = c_i \cdot 273.15 + L_f(T_{00}) \tag{8}$$

Equation 8 consists of c_i, the specific heat capacity of ice ($2060\,\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$), L_f is the latent heat of fusion, given by $-0.3337\cdot10^6$ Jkg^{-1}. Finally, the last term in equation 7 is the geopotential Φ. Φ is the potential energy of an air parcel lifted z m above a reference level. It is denoted as $\Phi = g\cdot z$, with g as the gravitational acceleration and z as meters above ground.

If we want to study the energy budget of a column of the Earth, the column consists of atmosphere, ocean, land and ice. This system is presented in Fig. 5. We consider a land column to a depth of about 3 m. As a first step, the integration over an atmospheric column is carried out and discussed in the following.

The moist static energy equation reads as presented in equation 9. With the tendency of moist static and kinetic energy, the advection of moist static and kinetic energy (eq. 10) and the fluxes across the borders, denoted by the energy input in the system Rad_{TOA}.

$$\frac{\partial(m+k)}{\partial t} + N + Rad_{TOA} = 0 \tag{9}$$

The initial point for deriving the energy budget for the atmospheric column is the advection of moist static and kinetic energy given in equation 10.

$$N = \vec{v} \cdot \nabla(m+k) + \omega \frac{\partial(m+k)}{\partial p} \tag{10}$$

N is an immaterial source term, $\vec{v}$ the horizontal wind vector, m the moist static energy as defined in equation 7, k the atmospheric kinetic energy and p the atmospheric pressure. ω is the vertical velocity and can be expressed as dp/dt. The next step is the use of the continuity equation for an atmosphere with water vapour and liquid water (equation 11).

$$\nabla \cdot \vec{v} + \frac{\partial \omega}{\partial p} = 0 \tag{11}$$

In equation 11 the term $\partial\omega/\partial p$ is composed of liquid water and water vapour fluxes.

The next step is the vertical integration of equation 11.

$$\frac{1}{g} \nabla \cdot \int_0^{p_s} \vec{v} q \, dp + \frac{1}{g} \overline{\omega}\, \overline{q}(p_s) + \frac{1}{g} \overline{\omega' q'}(p_s) = 0 \tag{12}$$

$$\frac{1}{g} \nabla \cdot \int_0^{p_s} \vec{v} q \, dp + P + E = 0 \tag{13}$$

The second term in equation 12 denotes the total precipitation P as a sum of rain and snow. According to Hantel and Haimberger (2016, p.284) P is bigger than zero if it rains, that means that the precipitation reduces the precipitation water in the atmospheric column.
The third term in equation 12 is the evaporation rate E at the surface. Based on information from Hantel and Haimberger (2016, p.284), E is the vertical flow of water vapour which is exchanged between Earth's surface and the vertical atmosphere.
Multiplication of equation 11 by $(m+k)$ and adding to equation 10 leads to the flux form of the total atmospheric budget (eq. 14).

$$N = \nabla \cdot [\vec{v}(m+k)] + \frac{\partial}{\partial p} [\omega(m+k)] \tag{14}$$

Vertical integration of equation 14 results in equation 15.

$$N = \nabla \cdot \frac{1}{g} \int_0^{p_s} [\vec{v}(m+k)] \, dp + (m_s + k_s)(P+E) \tag{15}$$

The last term on the right hand side of equation 15 describes the enthalpy flux in relation with the mass transfer through the surface.
There are some assumptions made concerning this term, like no flux of dry air through the surface and that the kinetic energy and the geopotential flux in the context of E and P are neglected. So the term can be expressed by enthalpies of rain, snowfall and evaporation.

$$(m_s + k_s)(P+E) = (P_{rain} + P_{snow})[h_0 + c_w(T_p - T_{00})] - L_f(T_p) P_{snow} + E[h_0 + c_w(T_s - T_{00}) + L_v(T_s)] \tag{16}$$

The term $L_f(T_p) P_{snow}$ in equation 16 is called the snowfall term. It represents the water mass which falls as snow. This water mass leaves the atmospheric column with an enthalpy decreased by the latent heat of fusion $L_f(T_p)$. The term $EL_v(T_s)$ denotes the latent heat flux (LH). LH can be used

together with the sensible heat flux (SH) and the net surface radiation (Rad_{SFC}) as the net surface energy flux (F_S) in form of equation 17.

$$F_S = Rad_{SFC} + LH + SH \tag{17}$$

At the top of an atmospheric column there is only radiation exchange (Rad_{TOA}). Rad_{TOA} is defined as the sum of net incoming and outgoing radiation. It can be measured from spatialized satellite systems like CERES ((Loeb et al., 2009)).
All this effects are now combined in an equation for the total atmospheric energy budget:

$$F_S = Rad_{TOA} - AET - \nabla \cdot F_A + L_f(T_p)P_{snow} \tag{18}$$

In equation 18, F_S is the net surface energy flux as defined in equation 17, Rad_{TOA} is the net radiation at the top of the atmosphere, AET is the atmospheric total energy tendency and $\nabla \cdot F_A$ is the divergence of lateral atmospheric energy transports. A more detailed description of the terms can be found in Mayer et al. (2019).
The last term is the snowfall term with $L_f(T_p)$ as the latent heat of fusion at temperature T_p, which denotes the temperature of the falling snow. The other factor taking part in the snowfall term is the snowfall rate P_{snow} in $\mathrm{kgm^{-2}s^{-1}}$. This term describes the heat released in an atmospheric column in combination with net freezing. (Mayer et al., 2019, 2017)

2.4 Partitioning of the Earth's surface

The result of subsection 2.3.2 provides the input for Earth's surface. If F_S as defined in equation 18 reaches the surface, there are three possibilities of surface types, which occupy certain fractions of the surface: ocean surface fraction f_o, ice surface fraction f_i or land surface fraction f_l. These three fractions add up to 1 (eq. 19).

$$1 = f_o + f_i + f_l \tag{19}$$

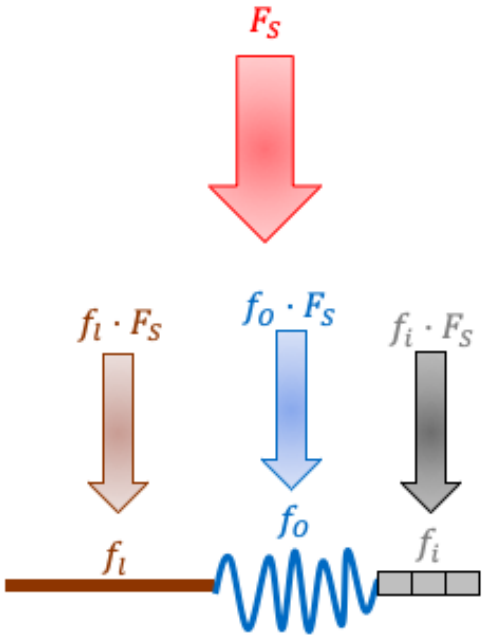


Figure 4: Partitioning of Earth's surface.

Land and ice surfaces may be covered with snow or not.
The F_S from the atmosphere and the F_S of the subsystems are the same. If they were not the same, they would not cancel each other in a common column of atmosphere and the subsystem. The additional flux would increase or decrease the energy in the column. This would be an energy change through internal transformation of energy and thus a conflict with the energy law. (Hantel and Haimberger, 2016, p.263)

Fig. 5 shows all components of the entire system. Below the surface we must distinguish between land ocean and ice. There is a budget equation for each type of surface, presented in sections 2.5 to 2.7. In this work the subsystems atmosphere and land surface are of particular interest.

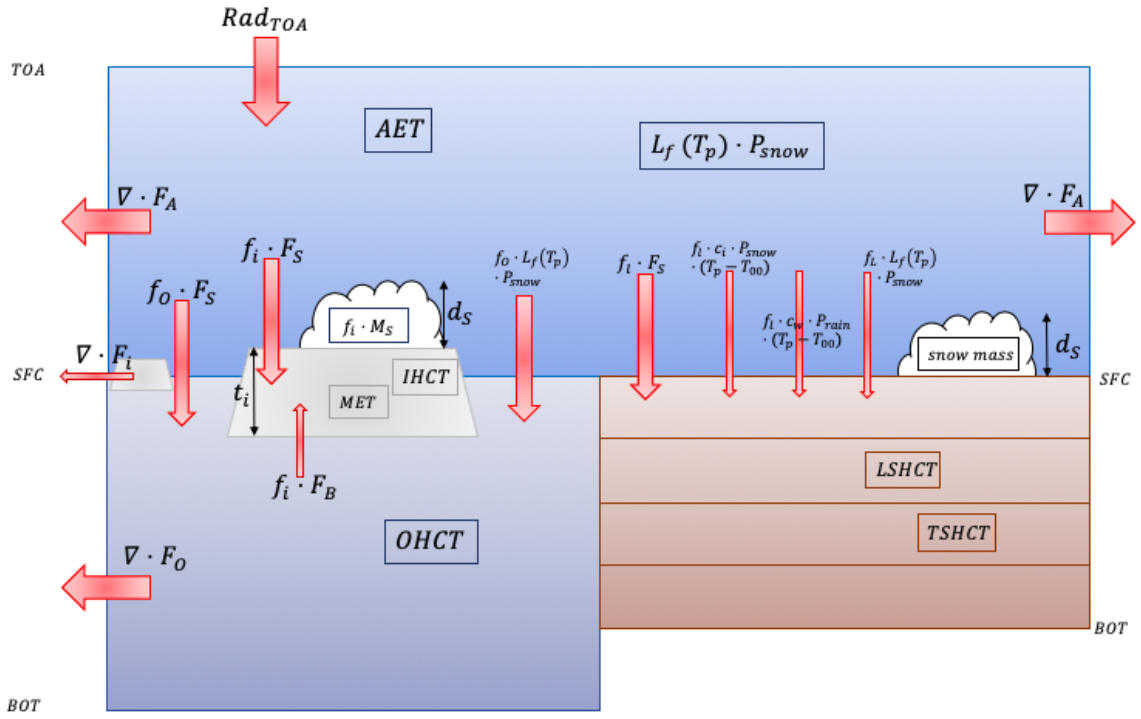


Figure 5: Terms of the energy budget of the system together. Based on Mayer et al. (2019).

2.5 Energy budget of an ocean column

To get the energy budget of an ocean column the field energy equation is integrated from the sea surface (SFC) to the bottom of the ocean (BOT). Due to consistency reasons, the snowfall term is also present in equation 20.

$$f_i \cdot F_B + f_o \cdot F_S = OHCT + \nabla \cdot F_O - f_o \cdot L_f(T_p) P_{snow} \tag{20}$$

In equation 20, f_i denotes the fraction of sea ice in a grid box, F_B is called the basal heat flux, which is directed from the sea surface to the ice that over floods it. This term can be estimated by a balance equation for the flux from water to ice or can be assumed basically as zero. The fraction of ocean in a grid box is given by f_o and F_S is the vertical net energy flux (as defined in equation 17). $OHCT$ is the temporal tendency of ocean heat content, $\nabla \cdot F_O$ is the divergence of ocean heat transport. The third term on the right hand side of equation 20 is the snowfall term and beside the sign and the factor f_o, which denotes the fraction of ocean in a grid box, the same as the last term on the right hand side of equation 18 and describes the cooling effect of snowfall. (Mayer et al., 2019)
According to Hantel and Haimberger (2016, p.263) the energy in the ocean is almost only inner energy, denoted by the temperature of water and the latent energy due to phase transformations is negligible.

2.6 Energy budget of an ice column

Ice and water differ basically by the state of aggregation, so a budget equation for an ice layer can be derived on the same way as for the ocean. The main difference is that the budget equation for the ice column (eq. 21) does not use vertical integrals. Instead of vertical integration, multiplication with t_i, the average thickness of the ice layer, is used.

$$f_i \cdot F_S - f_i \cdot F_B = MET + IHCT + \nabla \cdot F_I + f_i M_{snow} \tag{21}$$

The left hand side of equation 21 represents the vertical convergence of the energy flux within the ice layer. The fraction of sea ice is indicated by f_i. MET is the tendency of energy required to melt sea ice. This means in concrete terms the latent heat of fusion (L_f) which is needed for melting, and released during freezing sea ice.
In agreement with Hantel and Haimberger (2016, p.303), ice has almost only internal energy, so kinetic energy is neglected. Thus a measure of energy is given by the definition of the sensible ice heat content tendency ($IHCT$).
The divergence of the transport of latent heat in association with sea ice is considered by $\nabla \cdot F_I$ in the budget equation. The divergence of the sensible heat in the ice is neglected because this effect is very small compared to the divergence of latent heat.
When snow is lying on the sea ice, the energy required for snow melt is taken into account by $f_i M_{snow}$. Energetic effects in combination with transported snow are neglected due to their very small size. (Mayer et al., 2019)

2.7 Energy budget of a land column

This section is based on Haimberger (2019).
For the definition of the energy budget of a land column, a more detailed formulation of equation 4 is requested. Therefore a combination of equation 1 and 17 is made and leads to equation 22.

$$F_S = -G \tag{22}$$

The underlying idea is to specify the left side of the equation 22 more precisely. F_S is multiplied by f_l to make sure that the ground heat flux is only considered over land surfaces. The integral of the ground heat flux is made over a layer that is defined from the bottom of the land column (BOT) to the snow depth on the surface, denoted by d_s.

$$f_l \cdot F_S = -\int_{BOT}^{d_s} \frac{\partial G}{\partial z} dz \tag{23}$$

The right hand side of equation 23 is expressed more precise in equation 24, in imitation of equation 21.

$$-\int_{BOT}^{d_s} \frac{\partial G}{\partial z} dz = TSHCT + LSHCT + ST - SF - CSF + RF \tag{24}$$

2.7.1 Tendency of thermal soil heat content

The first term in equation 24 is the tendency of soil heat content $TSHCT$. The more precisely formulation of $TSHCT$ is given in equation 25.

$$TSHCT = \int_{BOT}^{SFC} \left((1-\Theta_w-\Theta_i)\rho_s c_s + \rho_w c_w \Theta_w + \rho_i c_i \Theta_i\right) \frac{\partial T_s}{\partial t} dz \tag{25}$$

$(1-\Theta_w-\Theta_i)\rho_s c_s$ is the thermal heat capacity of dry soil, $\rho_w c_w \Theta_w$ is the thermal heat capacity of moist soil with Θ_w as the volumetric soil water content, and $\rho_i c_i \Theta_i$ is the one of frozen soil, whereas Θ_i is the volumetric soil ice content. All these thermal heat capacities are multiplied by the tendency of soil temperature $\partial T_s/\partial t$. Then the integral of this expression is formed over the considered soil layer. If the soil gets energy form a positive net radiation flux towards the surface, $TSHCT$ is also positive. That means that the surface gains energy what leads to a rise of soil temperature, which means that the temperature tendency is also positive. If the net radiation flux is negative, the surface looses energy and $TSHCT$ is also negative.

2.7.2 Tendency of latent soil heat content

The detailed configuration of the tendency of latent soil heat content $LSHCT$ is denoted in equation 26.

$$LSHCT = \int_{BOT}^{SFC} L_f \rho_i \frac{\partial \Theta_i}{\partial t} dz \tag{26}$$

This term describes the change of the ice content in the soil. If F_S is positive, the surface gains energy and the ice melts. Specifically, this means that the tendency of the soil ice content ($\partial\Theta_i/\partial t$) is negative. Due to the fact, that the latent heat of fusion (L_f) is defined negative, the whole term is then positive. If the ice content increases, the tendency is positive and the whole term turns negative.

2.7.3 Tendency of latent heat content of snow on the surface

With ST the effect of snow lying on the surface is taken into account. In equation 27 two options for calculating the term are given.

$$\begin{aligned} ST &= L_f \frac{\partial}{\partial t} \int_{SFC}^{d_s} \rho_{snow} dz \\ &= L_f \rho_w \frac{\partial SWE}{\partial t} \end{aligned} \tag{27}$$

The first option is, to do the integration over the snow depth. Therefore the density of snow must be known. According to DWD (2020a), the density of snow varies significantly depending on the type snow. For an accurate calculation of this term, a differentiation of the snow types has to be made, which is very costly.
Therefore a second option exists, which leads to the same result. Instead of integration over the density of snow, the snow water equivalent can be used. The snow water equivalent is the height of the water layer that would result after melting the snowpack if the melt water remains on a horizontal surface without infiltration or evaporation (DWD, 2020b). To get the same result as in option one, the multiplication by the density of water (ρ_w) has to be made.

Is F_S positive, a decrease of snowpack is the consequence and $\partial SWE/\partial t$ is negative. A multiplication with L_f leads to the same effects as discussed in section 2.7.2.

2.7.4 Snow Fall

The snowfall term is given in equation 28. This is the same as in equation 18.

$$SF = f_l L_f(T_p) P_{snow} \tag{28}$$

The only difference to equation 18 is that the term is multiplied by the fraction of land (f_l).
The negative contribution of SF to the total energy budget of the land surface is because of the fact that the land column looses energy if snow is falling on the surface. The loss of energy is in consequence of the fact that the snow first has to be warmed to the reference temperature T_{00} (water at 0 °C) for which energy is required.

In equation 24, $ST - SF$ must be exactly zero if snow falls and does not melt.
If the radiation budget is neutral and snow is falling on the surface, the column loses energy because of the negative enthalpy of the falling snow which is compensated by $LSHCT$. In summer energy is necessary to melt the snow.

2.7.5 Cold Snow Fall

The term for the cold snow fall (eq. 29) is a refining of the snow fall term in section 2.7.4.

$$CSF = f_l c_i P_{snow}(T_p - T_{00}) \tag{29}$$

In this term, the effect of very cold snow falling on the surface is taken into account. It needs to be heated up to 0 °C before snow melt, as described in section 2.7.3, can start.

2.7.6 Rain Fall

The last term on the right hand side of equation 24 is the effect of rain falling on Earth's surface, given in equation 30.

$$RF = f_l c_w P_{rain}(T_p - T_{00}) \tag{30}$$

Rain falling on the surface provides energy to the land column if it is warmer than 0 °C.

2.7.7 Neglected Terms

A term that is not considered in equation 24 is the tendency of the latent heat content due to soil moisture. If the soil moisture decreases, the soil cools down.
If it rains the soil is saturated with water after a certain time. After the rain, the soil water content can decrease either by infiltration or by evaporation. If evaporation occurs, the question arises where the heat comes from to evaporate. There are three possibilities: From the soil energy balance through radiation, from the atmosphere or from the soil. If the heat comes from the soil, it is cooled. This effect is difficult to take into account, because we only know the tendency of the soil moisture and it is unclear where the heat needed for evaporation comes from. (Haimberger, 2020)

2.8 Energy budget of the whole system

The addition of equations 18, 20, 21 and 24 gives the mathematical composition of the whole system and is provided in equation 31. A graphical representation is given in Fig. 5.

$$F_S = f_o \cdot F_S + f_i \cdot F_S + f_l \cdot F_S \tag{31}$$

$$Rad_{TOA} - \nabla \cdot F_A - AET = OHCT + \nabla \cdot F_O + MET + IHCT + \nabla \cdot F_I + f_i \cdot M_{snow} - \int_{BOT}^{d_s} \frac{\partial G}{\partial z} dz$$

3. Data

3.1 ERA5 reanalysis

Numerical weather models (NWPs) assimilate millions of observations and thus provide a three dimensional, consistent description of the atmospheric state. This result is called analysis. Analysis combine the advantages of the model with the advantages of the observations. With time, NWPs are developed further, so a time series of operational Analysis are not homogeneous over time.
For gaining a consistent time series, all available observations of overall decades are used to recalculate the Analysis with the same model. The output of this process is denoted as reanalysis. (DWD, 2020c)

ERA5 is the fifth generation of ECMWF atmospheric reanalysis. Reanalysis data is available from 1979 to present, an extension to 1950 is in progress. (Hersbach et al., 2020)
The ERA5 atmosphere and land reanalysis were produced by the 4DVAR data assimilation in the CY41R2 IFS cycle. It contains 137 pressure levels in the vertical, from surface at 1000 hPa to 0,01 hPa. There exist also many single level fields. The soil is resolved into four vertical levels. The soil model is further discussed in section 3.1.1.
The ERA5 data is available on a Gaussian grid with a global spatial resolution of 31 km, ten ensemble members are also available with a reduced spatial resolution. The temporal resolution is hourly, but monthly means are also available.
The data can be downloaded form C3S (2020), interpolated to a regular latitude and longitude grid as NetCDF or GRIB file. (ECMWF, 2020a)

The data downloaded for this thesis is listed in section 3.1.2.

3.1.1 Soil Model used in ERA5

The state of the soil is represented in ERA5 by the TESSEL (Tiled ECMWF Scheme for Surface Exchanges over Land scheme). A schematic representation of this scheme is given in Fig. 6.
On the left you can see the six surface types and four soil layers. The six surface types are high and low vegetation, interception water, bare ground, shaded and exposed snow. Water at surface is either frozen or not. All different surface types have different water and energy balances. A Fourier diffusion law is used for the description of the soil heat budget. For the energy balance equation the top boundary condition is a net ground heat flux, and at the bottom a flux of zero is used. The interception layer absorbs water until saturation is reached, then the remaining precipitation water is split up to surface runoff an infiltration. In TESSEL, a single loamy soil is used.
On the right hand side of Fig. 6 the improvements made in the HTESSEL (Hydrology TESSEL) scheme are illustrated. Here, a spatially varying soil type is used. Two adjoining grid boxes with equal land surface conditions which receives the same amount of precipitation will produce a different amount of runoff, proportional to the complexity of terrain with a soil water drainage depending on the soil texture class. (Balsamo et al., 2008)

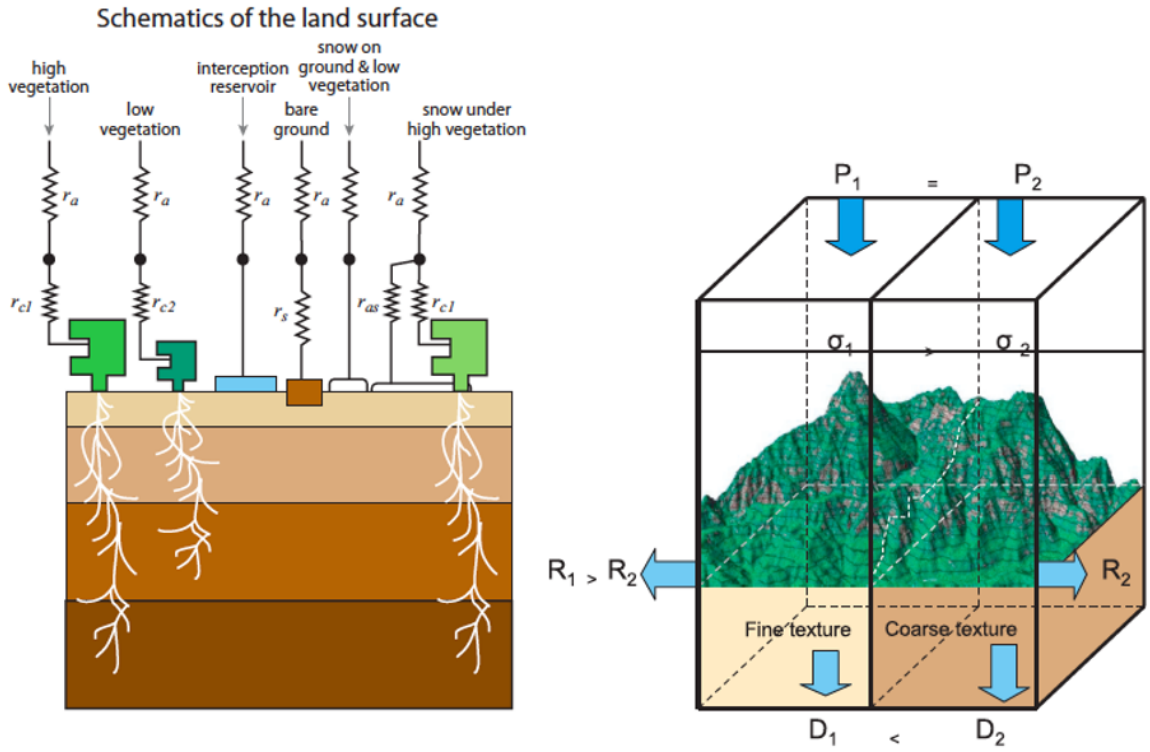


Figure 6: Visualisation of TESSEL (left) and HTESSEL (right) scheme. (Balsamo et al., 2008)

Surface heat and evaporation fluxes

The turbulent fluxes are calculated by using a resistance parametrisation depending on the type of surface. Equation 32 and 33 present the turbulent fluxes for heat and water vapour for ocean, sea ice and snow on low vegetation.

$$H_i = \rho_a c_p |U_L| C_{H,i} (T_L + \frac{g z_L}{c_p} - T_{sk,i}) \tag{32}$$

$$E_i = \rho_a |U_L| C_{H,i} (q_L - q_{sat}(T_{sk,i})) \tag{33}$$

The wind speed is denoted by U_L, q_L is the humidity, z_L the height of the lowest atmospheric model level, $T_{sk,i}$ is the skin temperature and $C_{H,i}$ is the turbulent exchange coefficient. $C_{H,i}$ varies from tile to tile due to the different atmospheric stabilities. The different tiles are marked by i. The air density is given by ρ_a, c_p is the heat capacity of air and T_L is the air temperature.
For dry high and low vegetation, the interception reservoir and for bare soil and also for high vegetation with underlying snow, a slight different formulation of 33 is used. The total fluxes within a grid box are denoted by an area weighted average over all tiles. (ECMWF, 2016)

Surface energy balance

For a determination of the surface energy balance the skin temperature $T_{sk,i}$ has to be defined. This is done by solving the surface energy balance equation for each tile. In HTESSEL, the feedback of $T_{sk,i}$ on the net radiation and the ground heat flux is integrated. The surface energy balance is given by equation 34.

$$(1 - f_{Rs,i})(1 - \alpha_i)R_S + \epsilon(R_T - \sigma T_{sk,i}^4) + H_i + L_{v,s}E_i = \Lambda_{sk,i}(T_{sk,i} - T_1) \tag{34}$$

Here too, the different tiles are marked by i. The partial absorption of net short wave radiation is taken into account by $(1 - f_{Rs,i})$. R_S is the downward short wave and R_T is the downward long wave radiation. The Boltzmann constant is σ, T_1 is the temperature of the upper snow layer, H_i is the sensible heat flux, L_vE_i is the latent heat flux from the skin layer. If there is evaporation of snow, L_s is used instead of L_v. Finally, $\Lambda_{sk,i}$ denotes the skin conductivity for each tile. (ECMWF, 2016)

Soil heat transfer

As mentioned in section 3.1.1, the soil heat transfer is following the Fourier law of diffusion (eq. 35).

$$(\rho C)_{soil}\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left[\lambda_T \frac{\partial T}{\partial z}\right] \tag{35}$$

The volumetric soil heat capacity is assumed as constant with a value of $2.19 \cdot 10^6 \mathrm{Jm^{-3}K^{-1}}$. T is the soil temperature, z denotes the distance from the surface in m and λ_T is the thermal conductivity in $\mathrm{Wm^{-1}K^{-1}}$ depending on the soil water content.
Equation 35 ignores the effect of phase changes an the heat transfer in combination with vertical movement of water in the soil. The depths of the soil layers are chosen in an approximate geometric relation. The depths of the four soil layers used are given in table 2. (ECMWF, 2016)

Soil Layer	**Depth**
Layer 1	0 - 7 cm
Layer 2	7 - 28 cm
Layer 3	28 - 100 cm
Layer 4	100 - 289 cm

Table 2: Depths of the Soil Layers. (ECMWF, 2016)

According to Warrilow et al. (1986) the soil layers used as described in table 2 are accurate enough, for the consideration of data with temporal resolutions from daily to yearly.

3.1.2 Parameters used

In this section the parameters used in this thesis are described. In table 3 the abbreviation and parameter identification of the parameters used in ECMWF (2020a) are listed. All the parameters are directly downloadable form the C3S (2020) data store.

Parameter name	Parameter ID	Abbreviation	Abbreviation used
Surface latent heat flux	147	slhf	LH
Surface sensible heat flux	146	sshf	SH
Surface net solar radiation	176	ssr	$\text{Rad}_{\text{SFC}_{\text{solar}}}$
Surface net thermal radiation	177	str	$\text{Rad}_{\text{SFC}_{\text{thermal}}}$
Volumetric soil water content layer 1	39	swvl1	
Volumetric soil water content layer 2	40	swvl2	
Volumetric soil water content layer 3	41	swvl3	
Volumetric soil water content layer 4	42	swvl4	
Soil temperature layer 1	139	stl1	
Soil temperature layer 2	170	stl2	
Soil temperature layer 3	183	stl3	
Soil temperature layer 4	236	stl4	
Convective rain rate	228218	crr	
Large scale rain rate	228219	lsrr	
Convective snow fall rate	228220	csfr	
Large scale snow fall rate	228221	lssfr	
Snow depth	141	sd	
2m temperature	167	2t	
2m dew point	168	2d	

Table 3: Parameters used with abbreviated name and parameter identification. (ECMWF, 2020a)

The parameters used are described in more detail below. They can be used to calculate F_S either directly from the forecast flux fields above, or indirectly from the changes of the analysed state quantities.

Parameters for calculation of net energy flux

The first option is the direct calculation of the net energy flux via equation 17 with the use of the parameters used in ERA5, as presented in equation 36. Therefore the surface latent heat flux, the surface sensible heat, the surface net solar radiation and the surface net thermal radiation are added. All this parameters are available as accumulated values in Jm^{-2}. The accumulation is dependent on the time period selected. To convert the values in Wm^{-2}, the downloaded data has to be divided to the accumulation period in seconds. (ECMWF, 2020b)
According to ECMWF (2020a), the fluxes have to be divided by 86 400 s.

$$F_S = slhf + sshf + ssr + str \tag{36}$$

This fluxes are parameterized and they are quite uncertain. The parameters described in the following are analysed fields and assimilated by satellite data, so they are closer to the observations as the forecast fluxes. The state quantities can be used to evaluate the soil budget equation. This is the core of this work and is described in section 4.

Volumetric soil water content

The volumetric soil water content describes the volume of water in the respective layers in m^3m^{-3}. The volumetric soil water content is available in four layers, as listed in table 2. It is an average value for each layer and linked to soil parameters as soil texture or the underlying ground water level. (ECMWF, 2020b)

Soil temperature

For all layers in table 2 there exists a soil temperature in K. The value is assigned to the middle of the corresponding layer. (ECMWF, 2020b)

Rain and snow fall rate

Equation 30 requires a rain fall rate in $kgm^{-2}s^{-1}$. The rainfall rate (eq. 37) consists of the convective (crr) and the large scale rain rate ($lsrr$).

$$P_{rain} = crr + lsrr \tag{37}$$

The convective rainfall rate is calculated by the convection scheme of the IFS, which represents convection on smaller scales than grid size. The cloud scheme of the IFS generates the large scale rainfall. This constitutes the formation and dissipation of clouds and large scale rainfall in consequence of changes in atmospheric pressure, temperature or moisture.

Everything that applies to the rainfall rate can be directly applied to the snowfall rate. The snowfall rate (eq. 38) consists too of the convective ($csfr$) and the large scale snow fall rate ($lssfr$). (ECMWF, 2020b)

$$P_{snow} = csfr + lssfr \tag{38}$$

The total precipitation amount is denoted as P (eq. 39) and is the sum of rainfall and snowfall rate.

$$P = P_{rain} + P_{snow} \tag{39}$$

Snow depth

The snow depth is used in m of water equivalent. If all the snow of the snow-covered area of a grid box is melted and distributed over the whole grid box the snow depth in water equivalent results. (ECMWF, 2020b)

Temperature and dew point

For an estimation of the precipitation temperature T_p, the temperature and the dew point in 2 m above the surface are necessary. Both variables are given in K. (ECMWF, 2020b)
The method for the estimation can be found in section 4.6.

3.2 Indirect calculated net energy flux

To get F_S from the atmospheric budget equation (eq. 18) the divergence of the lateral atmospheric ($\nabla \cdot F_A$) has to be calculated. This is a classical method and is explained in detail in Mayer and Haimberger (2012) or Holopainen and Fortelius (1986). The total energy flux divergence is then combined with the CERES observation data of the net radiation at the top of the atmosphere (Loeb et al., 2009). The only difference to equation 18 is that in this calculation the term concerning snow fall was omitted.
The data calculated with this method was provided by Johannes Mayer, staff at the Institute of Meteorology and Geophysics at the University of Vienna.

4. Methods

This section describes the necessary steps to indirectly calculate F_S from the soil budget equation 24.

4.1 Regridding

Due to the fact, that the ERA5 data is available on another grid than the CERES TOA flux, regridding is required and necessary to ensure the comparability of results.

Regridding is the interpolation process from one grid resolution to another grid resolution. It may include temporal, vertical or horizontal interpolations. In this work, horizontal regridding is used in order to interpolate the resolution of the ERA5 data from 0,25 ° to 1 ° spatial resolution.
There are various regridding methods available, therefore it is a difficult task to find the interpolation methods that fits the problem best. (Shea, 2020)

The first part of regridding is the calculation of a weight matrix, which includes information of how the grid points of the source grid contribute to the target grid. After this step, the data of the source grid is multiplied by the weight matrix to generate the interpolated values of the target grid. (DeLuca, 2018)

In this thesis, the Python module xesmf (Zhuang, 2017b) is used for regridding. This module uses the Earth System Modeling Framework (ESMF), which is a software infrastructure for building and coupling earth science related applications. (DeLuca, 2018)
There are some different methods of regridding, whereas bilinear and conservative are the most frequently used. For the upscaling of data, conservative regridding is recommended. The upscaling method takes averages over small source grid boxes while the bilinear method throws away most of the source grid boxes. (Zhuang, 2017a)

4.2 Soil ice content

For the determination of the soil latent heat content tendency, the soil ice content is required. In ERA5 only the soil water content is available. Therefore an approximation for the soil ice content has to be made.

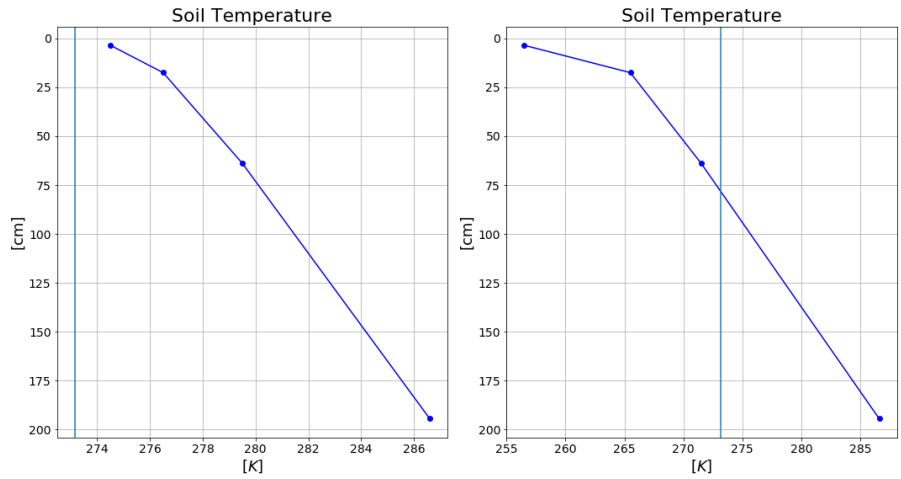


Figure 7: Two example profiles of soil temperature for showing the determination of the soil ice content.

Fig. 7 shows two example profiles of soil temperature at two different points in time. The vertical line in light blue denotes the freezing temperature of 273,15 K.
The left picture shows a soil, where the temperature in every layer is greater than 273,15 K. The right picture gives a profile where the temperature in the upper part of the soil to approximately 75 cm in depth is less then 273,15 K.
If the temperature of soil is lower than 273,15 K, the soil water available is assumed to be frozen, which is assumed to be the soil ice content. Otherwise, if the temperature is greater than 273,15 K, it is assumed that the soil contains no ice.

4.3 Zero degree levels

From the profiles shown in Fig. 7 the zero degree level can be derived. That means the depth in soil, where a temperature of 273,15 K is reached. There can be none if the soil temperature in every layer is smaller or if the soil temperature is always greater than 273,15 K. With four data points, at a maximum there can be three zero degree levels in one profile. The zero degree levels are needed for the integration with the trapezoidal rule, described in section 4.5.2.

4.4 Calculating tendencies

4.4.1 Calculating tendencies from monthly data

In equation 24 the tendency of thermal heat content (eq. 25), the tendency of latent heat content (eq. 26) and the tendency of latent heat content of snow mass on the surface (eq. 27) need to be evaluated. The calculation of tendencies is made by the use of finite centered differences, based on Brand (2017).

$$\frac{\partial X}{\partial t} = \frac{X(t_{j+1}) - X(t_{j-1})}{2\Delta t} \tag{40}$$

Equation 40 describes the calculation method for a parameter X, which can be the soil temperature, the soil ice content or the water equivalent of snow in this thesis. The computation of the tendency is made by subtracting the value of the previous month ($X(t_{j-1})$) from the value of the next month

($X(t_{j+1})$). The difference is divided by two times the time range Δt, which is one month.
If the variable for which the tendency is to be calculated is present in several layers, the tendency is calculated separately for each layer.

This method is not very accurate and is only recommended if only monthly data but no daily data are available.

4.4.2 Calculating tendencies from instantaneous fields

ERA5 provides the above mentioned state quantities with hourly resolution. In this thesis the instantaneous field for soil temperature, soil water content and snow depth can therefore be used to calculate the tendencies as given in equation 41.

$$\frac{\partial X}{\partial t} = X_{t_{end}} - X_{t_{start}} \tag{41}$$

$X_{t_{end}}$ denotes the value of the end of the month and $X_{t_{start}}$ is the value at the beginning of the month, typically a daily average. In this work the tendency for example April is calculated by first of May at 0000 UTC ($X_{t_{start}}$) minus first of April at 0000 UTC ($X_{t_{end}}$).

4.4.3 Calculating tendencies from forecast data

A further point for tendency calculation is the calculation from forecasts. For the determination of $TSHCT$ and ST, two times a day the difference between forecast and analysis is calculated and then the average for one month is determined. This data was provided by Johannes Mayer, staff at the Institute of Meteorology and Geophysics at the University of Vienna.
For the calculation of $TSHCT$ and ST it does not matter whether the tendencies are calculated first and then the integrals of the tendencies are calculated, or the reverse. First, the tendencies were calculated with equation 42.

$$\frac{\partial X}{\partial t} = \frac{1}{N} \sum_{i=1}^{N} \frac{X_{fc}^{i} - X_{an}^{i}}{\Delta t} \tag{42}$$

In equation 42, N is the number of forecasts in a month (e.g. 62 for January), X_{fc}^{i} are the 12 hour forecasts and X_{an}^{i} are the analysis in a month. Δt is the forecast time, in this work 12 hour forecasts are considered.

For the calculation of $LSHCT$, the calculation order is important, because the soil ice content depends on the soil temperature. Therefore the integrals must be determined first and then the tendencies of the integrals are calculated. Therefore monthly averages for the 12 hour forecasts and the analysis were calculated. The integrals are calculated for the monthly means of analysis and forecasts. Afterwards the difference is formed and divided by 43200 to get monthly means of $LSHCT$ from forecast data. (Haimberger, 2020)

4.5 Methods of integration

In equations 25 and 26 the discretization of the vertical integrals is required. The discretization methods used are discussed in this section.
Soil temperature and soil moisture (btw. soil ice content) are available in four layers. Therefore these integrals are calculated for each layer separately and are added together after integration to obtain the total heat content of the soil.

4.5.1 Integration with Riemann integral

The simplest approximation to the integral is the use of rectangles for calculating the area under the curve. According to Strampp (2012), the soil is divided into layers, the depth of each layer is denoted by $x_k - x_{k-1}$ in equation 43. The depth of the layer is then multiplied by the value of the function in that layer ($f(\xi_{k-\frac{1}{2}})$). For the total value of the whole function, all calculated values of the layers are summed up.

$$\int_a^b f(x) \approx \sum_{k=1}^{n} f(\xi_{k-\frac{1}{2}})(x_k - x_{k-1}) \tag{43}$$

4.5.2 Integration with trapezoidal rule

A slightly more accurate discretization method than the Riemann integration is the integration with trapezoids. The formula used for trapezoidal integration is given in equation 44, as given in Hanke-Bourgeois (2009). $(b-a)$ denotes the depth of the layer, $f(a)$ is the value at the begin of the layer, $f(b)$ the one at the end.

$$\int_a^b f(x) \approx \frac{b-a}{2}(f(a) + f(b)) \tag{44}$$

This integration method uses the calculated zero degree levels as markers for additional layers. As an example, consider 2 layers, where layer 1 extends from 0 to 7 cm into the ground and layer 2 from 7 to 28 cm. For the trapezoidal rule, the temperature in the layer is assigned to the middle of the layer. So the temperature of the first layer is assigned to a height of 3,5 cm, the temperature of the second layer to a height of 17,5 cm. The first sub-integral is calculated with the Riemann integral from the surface (0 cm) to the first layer centre (3,5 cm). If the middle of the first layer is reached it has to be distinguished if the temperature in the second layer has the same sign as the first layer. If the temperature of the second layer has the same sign as the first layer, formula 44 can be applied and the calculation is done for a layer from 3,5 to 17,5 cm. If there is a sign change of temperature between the layers, the zero degree level (as described in section 4.3) with a temperature of 0 °C is used as an additional layer. For this purpose equation 44 is applied for the range from 3,5 cm to the zero degree level and then additionally from the zero degree level to 17,5 cm. As the last sub-integral, the Riemann integration from 17,5 cm to the lower limit of the layer (28 cm) is carried out. Then all partial integrals are summed up.

A special feature is the calculation of $LSHCT$. The Riemann integration also uses the zero degree levels as layer boundaries.

4.5.3 Comparison of integration methods

To see if the integration with the trapezoidal rule delivers plausible values, some tests are performed. All tests are performed in a small area, from 54 °N to 60 °N and form 30 °E to 40 °E.
Each test was performed for both $TSHCT$ and $LSHCT$.

Test 1

In the first test, in every layer the same soil temperature/moisture is assumed. The expected output is that both integration methods show the same result. For this test, the area mean of the small area is calculated to show the time series. In order to better identify possible differences, only an extract of 3 years, from 2000 to 2003, is shown.
In Fig. 8, the result for the tendency of thermal heat content is shown. Already at first view one can see that there is hardly any difference between the integration methods. The differences, which can be seen especially in the area of extreme values, are so small that they can be neglected.

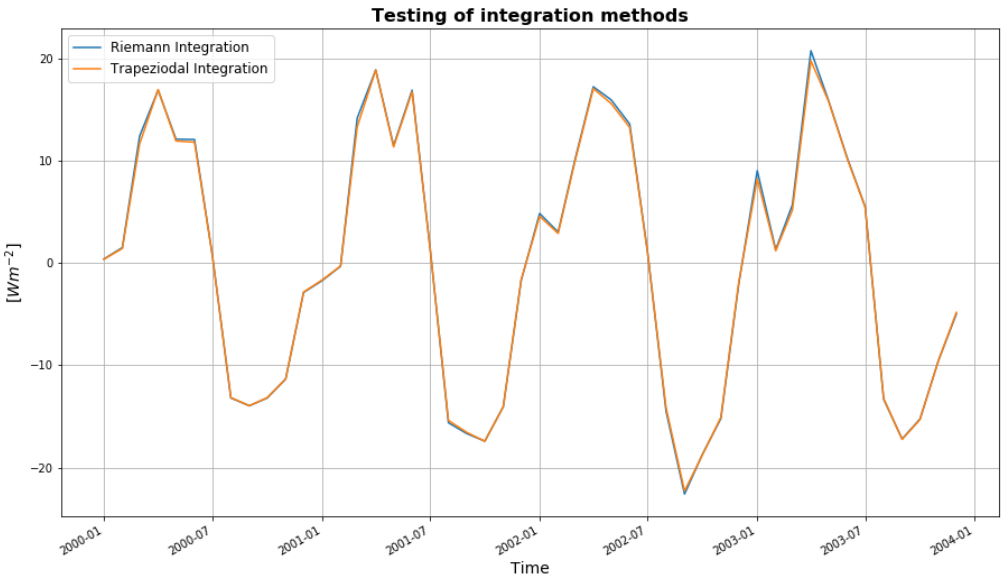


Figure 8: Result of integration for $TSHCT$ with Riemann integration (blue) and trapezoidal integration (orange) with assumption of the same soil temperature.

This test was also performed for $LSHCT$, with a very similar result.

Test 2

A second test compares the integration methods for the individual layers. Therefore the data in the chosen extract where grouped by the months and then averaged. This means an average for all January mean values and so on for all months.

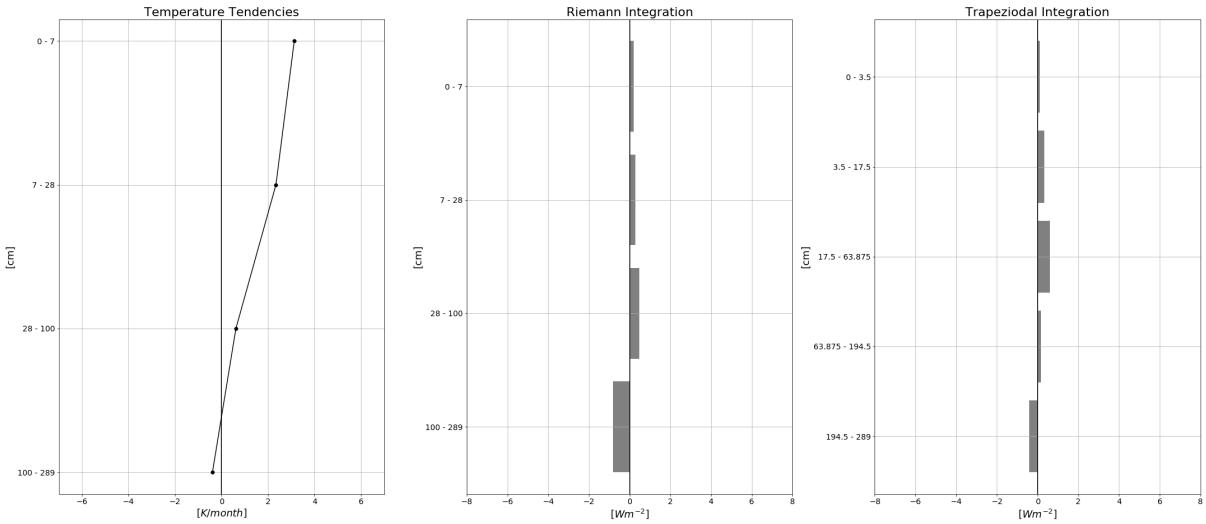


Figure 9: Temperature tendency, $TSHCT$ via Riemann integration and $TSHCT$ via trapezoidal integration as an average for a small area, averaged for all months March.

Fig. 9 shows averages for all months march for a small area of temperature tendency and the calculation of $TSHCT$ with both integration methods. Apart from the slightly different distribution of layers for the two integration methods, it can be seen that the values of integration are very similar. They are positive for positive temperature tendencies and negative when the temperature tendency turns negative.
Fig. 10 shows the course of the year for the averaged monthly means. For the given curves, the values of all layers are summed up. It appears that both curves are very similar, the peaks also occur almost simultaneously and are in the same value range from about $13\,\mathrm{Wm}^{-2}$ to $-11\,\mathrm{Wm}^{-2}$.

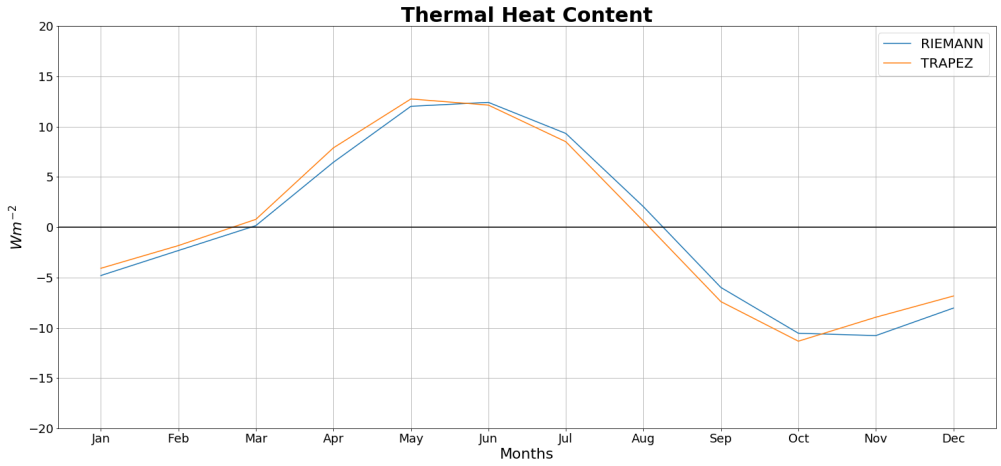


Figure 10: Time series for the averaged monthly means of $TSHCT$ for a small area are calculated by Riemann integration (blue) and by trapezoidal integration (orange).

Test 2 was also performed for $LSHCT$, with the same conclusion as for $TSHCT$: The two integration methods provide results in the same value range and show a similar course.

4.6 Precipitation temperature

The terms in equation 24 associated with precipitation need the precipitation temperature T_p. T_p can be estimated by the wet bulb temperature. According to AMS (2012), the wet bulb temperature T_f is the temperature an air parcel would have if it is cooled adiabatic to saturation and then compressed to the original pressure. Precipitation temperature is always lower than the 2 m temperature, the wet bulb temperature is also always lower than the 2 m temperature and so it is a good approximation to T_p.
There are many different empirical formulas for calculating the wet bulb temperature. In equation 45 a formula is given, that only uses the 2 m temperature in $°C$ and the relative humidity RH in %. It returns the wet bulb temperature in $°C$. (Stull, 2011)

$$\begin{aligned} T_f = T \cdot atan\left(0.151977(RH+8.313659)^{0.5}\right) + atan(T+RH) - atan(RH-1.676331) \\ + 0.00391838 RH^{1.5} atan(0.023101 RH) - 4.686035 \end{aligned} \tag{45}$$

According to Stull (2011) equation 45 is applicable for temperatures between $-20\,°C$ and $50\,°C$ and for relative humidities from 5 % to 99 %. The error made ranges from $-1\,°C$ to $0{,}65\,°C$.

The air temperature is available in ERA5 and can be directly used, but the relative humidity is not accessible. Therefore the dew point temperature in ERA5 is downloaded. With the use of the Clausius Clapeyron equation for the saturation vapour pressure (eq. 46), the relative humidity (eq. 47) can be calculated.

$$e_s(T) = e_s(T_0) \cdot exp\left[\frac{L_v}{R_D}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right] \tag{46}$$

The reference value for the saturation vapor pressure $e_s(T_0)$ is 101 325 Pa, the reference temperature T_0 is 373,15 K, the latent heat of vaporization is denoted with $2{,}501 \cdot 10^6$ $\mathrm{Jkg^{-1}}$ and the gas constant for dry air R_D is 287 $\mathrm{Jkg^{-1}K^{-1}}$.

$$RH = \frac{e_s(T_d)}{e_s(T)} \tag{47}$$

The relative humidity is the relation of saturation vapour pressure at the dew point to saturation vapour pressure at temperature. For values of RH in %, equation 47 has to be multiplied by 100. (Hantel, 2013, p.68,99ff)

5. Results and discussion

5.1 Evaluation of individual terms of the soil energy budget equation

5.1.1 Tendency of thermal soil heat content

As discussed in section 2.7.1, the tendency of thermal soil heat content is the integration over the depth of the soil of the temperature tendency multiplied by the specific heat capacities for soil, soil water content and soil ice content.

Seasonal Average (Riemann Integration)

Fig. 11 shows the means over all seasons for the considered time period from 2000 to 2018 for integration with Riemann integration as discussed in section 4.5.1.

What is noticeable in advance is that some areas have been masked out for all evaluations. An example of this is a large part of Greenland. As the snow water equivalent in this area has not changed over the period under consideration, it is assumed that glaciers have been assigned a constant value.This affects the calculation of ST and would falsify the result and is discussed in section 5.2.2.

On the top left (seasonal mean for spring), it is to see that $TSHCT$ is for a large part of the chart positive, but in the north of Canada and Asia it is negative. The negative values of this term mean that the soil loses energy because of the fact that the temperature tendency is negative. If it is positive, the soil gains energy, as is the case with the red areas. It can be seen that the heat content increases towards the south, which is to be expected in spring. The area mean in spring is $3{,}7 \pm 13{,}9\,\mathrm{Wm}^{-2}$. These values support the thesis that in spring most of the depicted area gains energy.

On the top right of Fig. 11 the $TSHCT$ as a seasonal mean for the summer months is given. It is obvious that the whole area gains energy in summer which leads to a positive temperature tendency in soil and so a positive $TSHCT$ follows. The values are slightly smaller in the west of Canada and in the east of Asia. Which is probably due to the fact that there more energy is needed to thaw the soil and therefore the temperature increases slower than in areas further south. The area mean is $7{,}1 \pm 13{,}5\,\mathrm{Wm}^{-2}$.

The chart on the bottom left of Fig. 11 shows the seasonally averaged $TSHCT$ of autumn. In autumn the values are negative throughout. This indicates that the soil temperature decreases and the surface loses energy, which is what is gradually expected in autumn. It is noticeable here that areas further south show lower values than areas further north. The average of the area is $-6{,}7 \pm 14{,}8\,\mathrm{Wm}^{-2}$.

In the chart on the bottom right of Fig. 11 the thermal heat content for winter is displayed. Areas further north show lower values than areas in the south, excepted parts of northern Canada and Asia. On a seasonal average, the energy loss in the north is lower for the winter months than for areas in the south. The area average of the thermal heat content in winter is $-4{,}2 \pm 11{,}7\,\mathrm{Wm}^{-2}$.

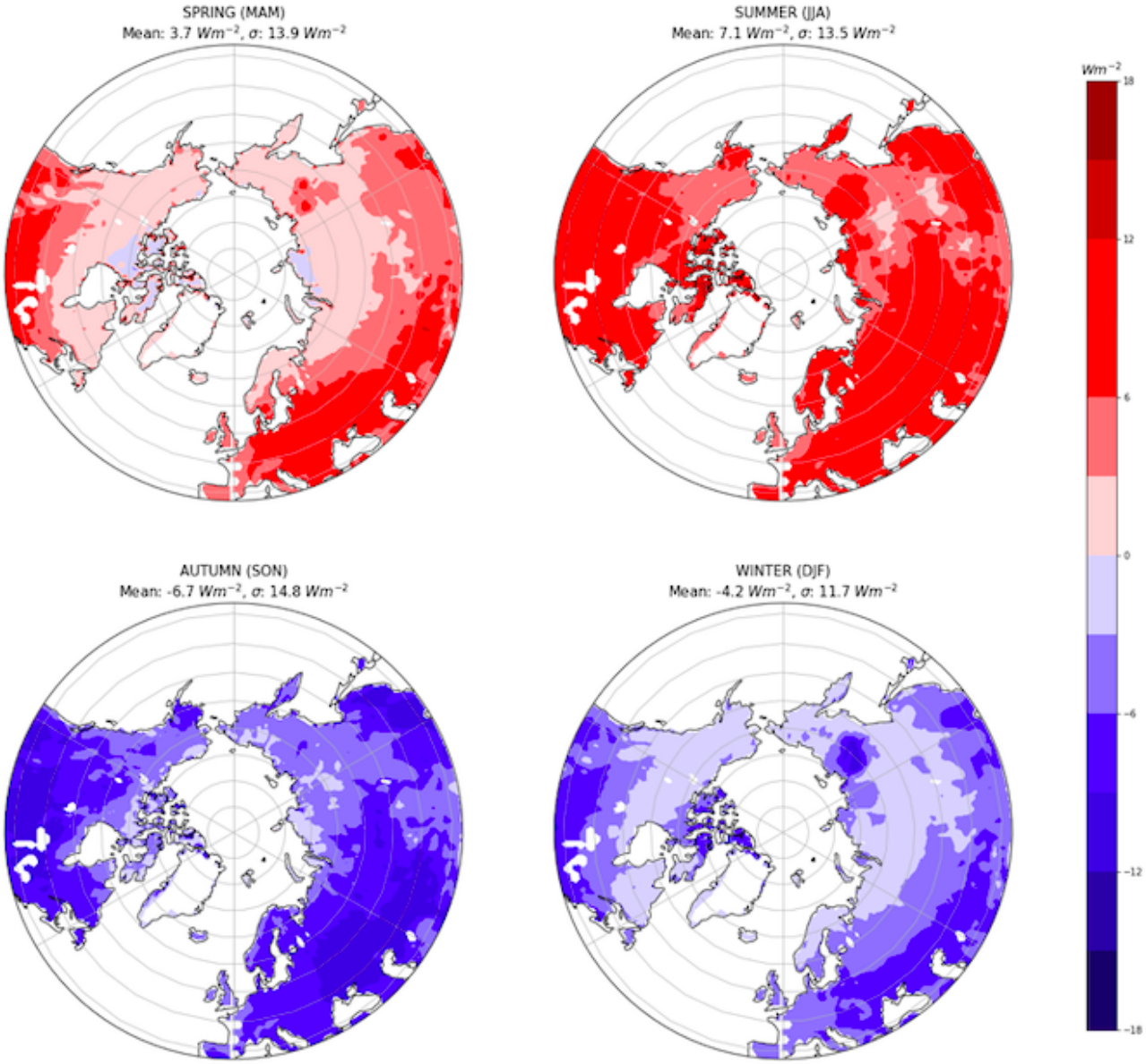


Figure 11: Seasonal means for TSHCT for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The integration method used is the Riemann integration. The units are Wm^{-2}*, zero is shown in white.*

Seasonal Average (Trapezoidal Integration)

Fig. 12 shows the same as Fig. 11, with the difference that instead of the Riemann integration, the integration was carried out using the trapezoidal integration. To better illustrate the differences between the integration methods, Fig. 13 shows the difference between Fig. 11 and Fig. 12.

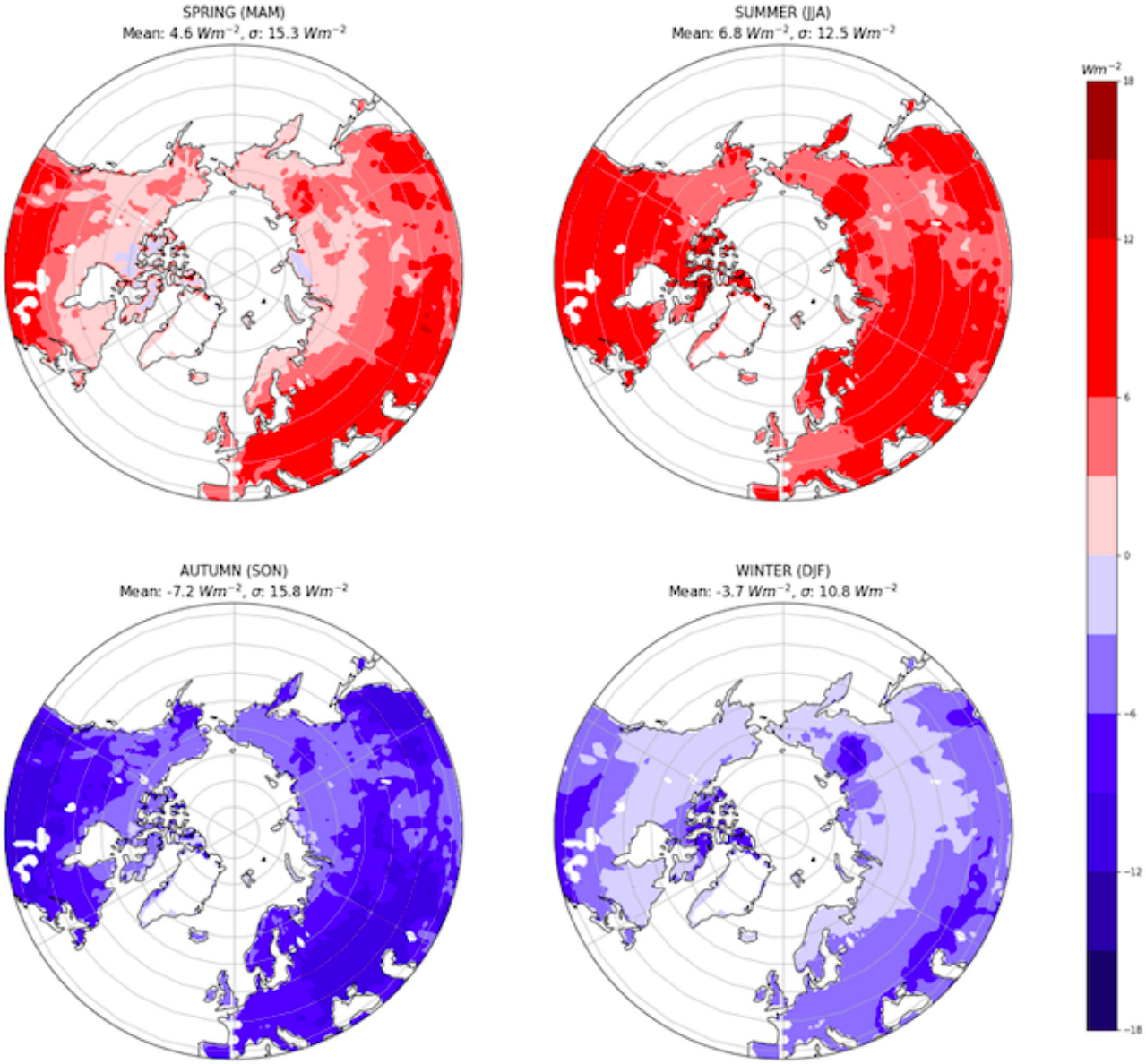


Figure 12: Seasonal means for $TSHCT$ for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The integration method used is the trapezoidal integration. The units are Wm^{-2}*, zero is shown in white.*

Differences between integration methods

In Fig. 13 the difference of calculating $TSHCT$ with Riemann integration and trapezoidal integration is provided. The result of trapezoidal integration is subtracted from the result gained by Riemann integration. Positive values indicate that the $TSHCT$ calculated by Riemann integration returns higher values than trapezoidal integration. If the trapezoidal integration delivers higher values than the Riemann integration, the result of the subtraction is negative.
Considering the value range of figures Fig. 11 and Fig. 12, which extend from $-18\,\mathrm{Wm}^{-2}$ to $18\,\mathrm{Wm}^{-2}$, the difference between the integration methods (Fig. 13) is about $\pm 3\,\mathrm{Wm}^{-2}$, which corresponds to a deviation of 16,67 %.

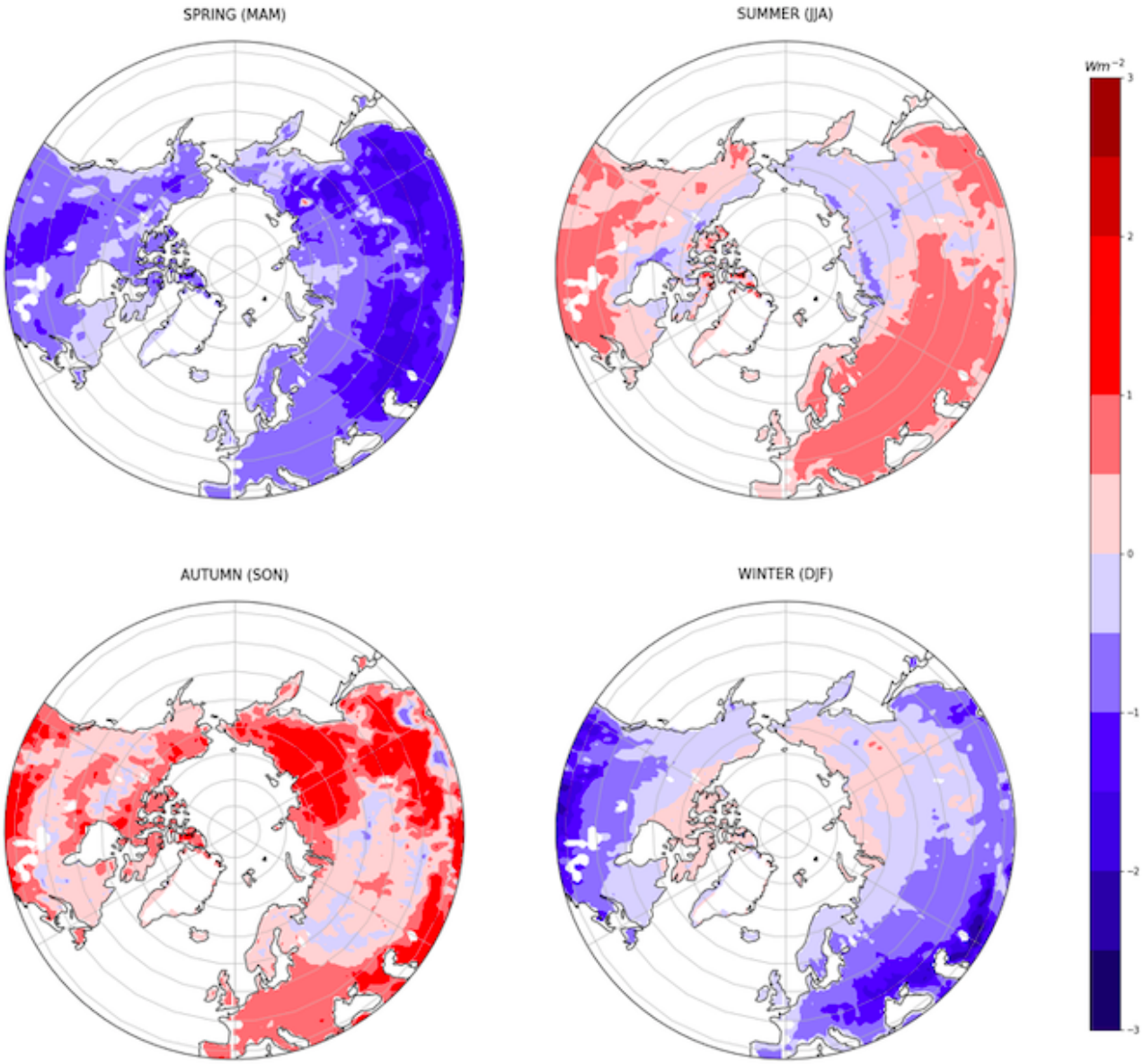


Figure 13: Riemann integration minus trapezoidal integration for $TSHCT$. Seasonal means for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white. Note the different color scale compared to Fig. 11 and Fig. 12.*

The top left part of Fig. 13 shows the difference of the integration methods in spring. It is recognizable that Riemann integration provides lower values than trapezoidal integration. The biggest differences occur in the middle of northern America and in Asia.

In summer (top right part of Fig. 13), Riemann integration provides higher values in the major part of the area. Negative differences occur in the north- west part of Asia and in the north of Canada.

The biggest differences in autumn (bottom left part of Fig. 13) occur in the southern part of northern America and in the west part of Asia. In the middle of the continents, Riemann integration provides lower values than trapezoidal integration.

In winter (bottom right chart of Fig. 13), in the southern part of the chart, the highest negative values are present. Further to the north, the differences are slight positive.

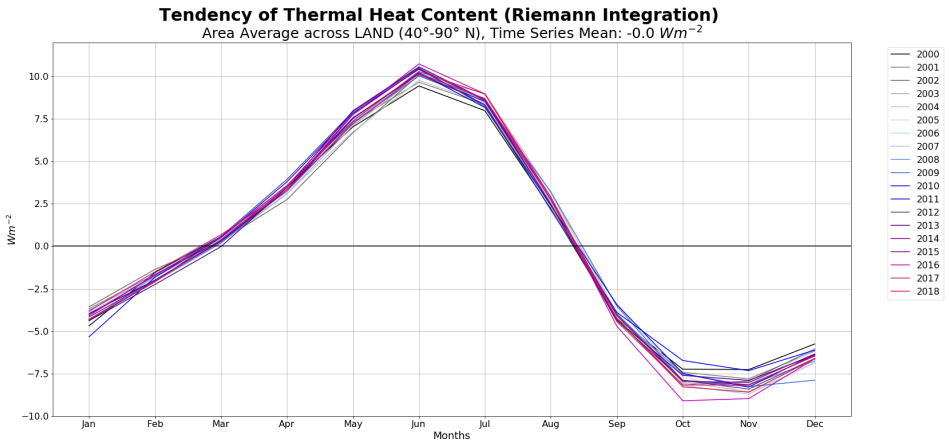
Possible reasons why one integration method yields higher values than the other is probably related to the zero degree levels in the soil. For example: It may happen that the temperature in the soil layer is slightly lower than 0 °C and in the layer underneath it is positive. So the zero degree level for trapezoidal integration is shortly after the middle point of the layer. This leads to the fact, that in Riemann integration the layer is counted as negative but in trapezoidal integration it is already counted as positive. The effect also occurs in the opposite direction and may lead to some differences.

Annual Cycle

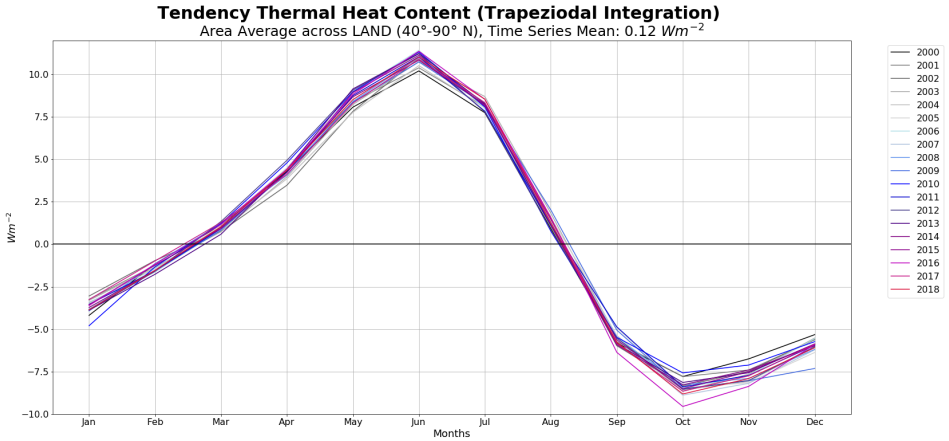
Fig. 14 shows the annual cycle of the tendency of thermal soil heat content. Each line shows one year, with the months of the year on the horizontal axis. The values on the vertical axis range from about $-10{,}0\,\mathrm{Wm^{-2}}$ to $12{,}0\,\mathrm{Wm^{-2}}$.
A spatial mean was made for the land surface from 40 °N to 90 °N. The result for Riemann integration is given in Fig. 14a that for trapezoidal integration in Fig. 14b.

Figures Fig. 14a and Fig. 14b show a similar course. At the beginning of the year, the thermal heat content is negative. In spring, at the beginning of march the sign changes from negative to positive, which means from this point on the soil gains energy. The peak is reached in June and is slightly higher for trapezoidal integration in Fig. 14b. After reaching the maximum, the $TSHCT$ of both integration methods decreases and changes the sign in August, whereas the change is a little faster in Fig. 14b than in Fig. 14a. In Fig. 14a the minimum of the thermal heat content cannot be determined exactly and varies between October and November depending on the year. The minimum of the thermal heat content in Fig. 14b can be more accurately detected than in Fig. 14a and is in October.
In addition, it can be recognized that the individual years are generally similar within a range of variation of about $3\,\mathrm{Wm^{-2}}$. In the end of the year the range is about $5\,\mathrm{Wm^{-2}}$, which represents a trend towards an enhanced seasonal cycle, which is consistent with Mayer et al. (2016).

On an annual average the tendency of thermal heat content should disappear. For Riemann integration the mean of the whole time series is $0\,\mathrm{Wm^{-2}}$, what the expected result is. Also for trapezoidal integration, the average of the whole time series is with $0{,}12\,\mathrm{Wm^{-2}}$ nearby the expected result, and still acceptable.



(a) Annual cycle of $TSHCT$ calculated by Riemann integration.



(b) Annual cycle of $TSHCT$ calculated by trapezoidal integration.

Figure 14: Annual cycles of $TSHCT$ for both integration methods. Every year is one curve. A spatial mean was made for the land surface from 40 °N *to* 90 °N. *The mean of the time series was calculated for the whole time series from 2000 to 2018.*

Uncertainties

For the calculation of this term some approximation were made. The first thing approximated is the value for the heat capacity of soil. As listed in section 2.7.1 and given in section 3.1.1, the heat capacity of soil is assumed as constant with a value of $2.19 \cdot 10^6$ $Jm^{-3}K^{-1}$. According to Abu-Hamdeh (2003), the thermal properties of soil are dependent on several factors, for example on the soil moisture content and soil density.
Another approximation made is that equation 25 allows soil, water and ice in the same layer. In this work the assumption is made that there is either water or ice.
A further fact of uncertainty is the calculation method of the tendency term, which will be discussed later in section 5.5.

5.1.2 Tendency of latent soil heat content

As discussed in section 2.7.2, the tendency of latent soil heat content is the integration over the depth of the soil of the soil ice content tendency multiplied by the the latent heat of fusion and the density of ice.

Seasonal Average (Riemann Integration)

Fig. 15 shows the seasonal averages for the tendency of latent soil heat content for all years from 2000 to 2018, calculated with Riemann integration.

The chart on the top left of Fig. 15 displays the average for spring. The area mean is given by 9,1 $\pm$ 61,5 Wm^{-2}. The positive values mean that the tendency of the soil ice content is negative, so the net energy flux is positive on area average and the soil ice content decreases. The highest values appear in northern America, and in center Asia. In Europe the white coloring means that there are no tendencies of ice content due to the assumptions made available (see assumptions in section 4.2). In the northern part of Canada the tendency of latent heat content is negative, so this means that there the soil ice content is still increasing in spring.

In summer (Fig. 15, top right) it is apparent that the ice content is retreating further to the north. The area is also positive with 10,8 $\pm$ 55,2 Wm^{-2}. $LSHCT$ is decreasing. On the edges of Greenland there are areas with a positive $LSHCT$. Which means that the soil ice content decreases on the edges, what is expected.

The chart on the bottom left of Fig. 15 presents the seasonal average for autumn. The area average of $-11,1 \pm 55,7\,Wm^{-2}$ reflects what is visible at first sight: The area mean of $LSHCT$ is negative in autumn. The lowest values are located in northern Canada and for the most part of Asia. In the edge area of Greenland, the latent heat content is slightly positive, which indicates the decrease of soil ice content.

Fig. 15 on the bottom right shows the seasonal mean for winter with an area mean of $-8,7 \pm 53,8\,Wm^{-2}$. This shows a mixed picture of positive and negative tendencies of latent heat contents, but most of the map show a negative $LSHCT$, so an increase of soil ice content.

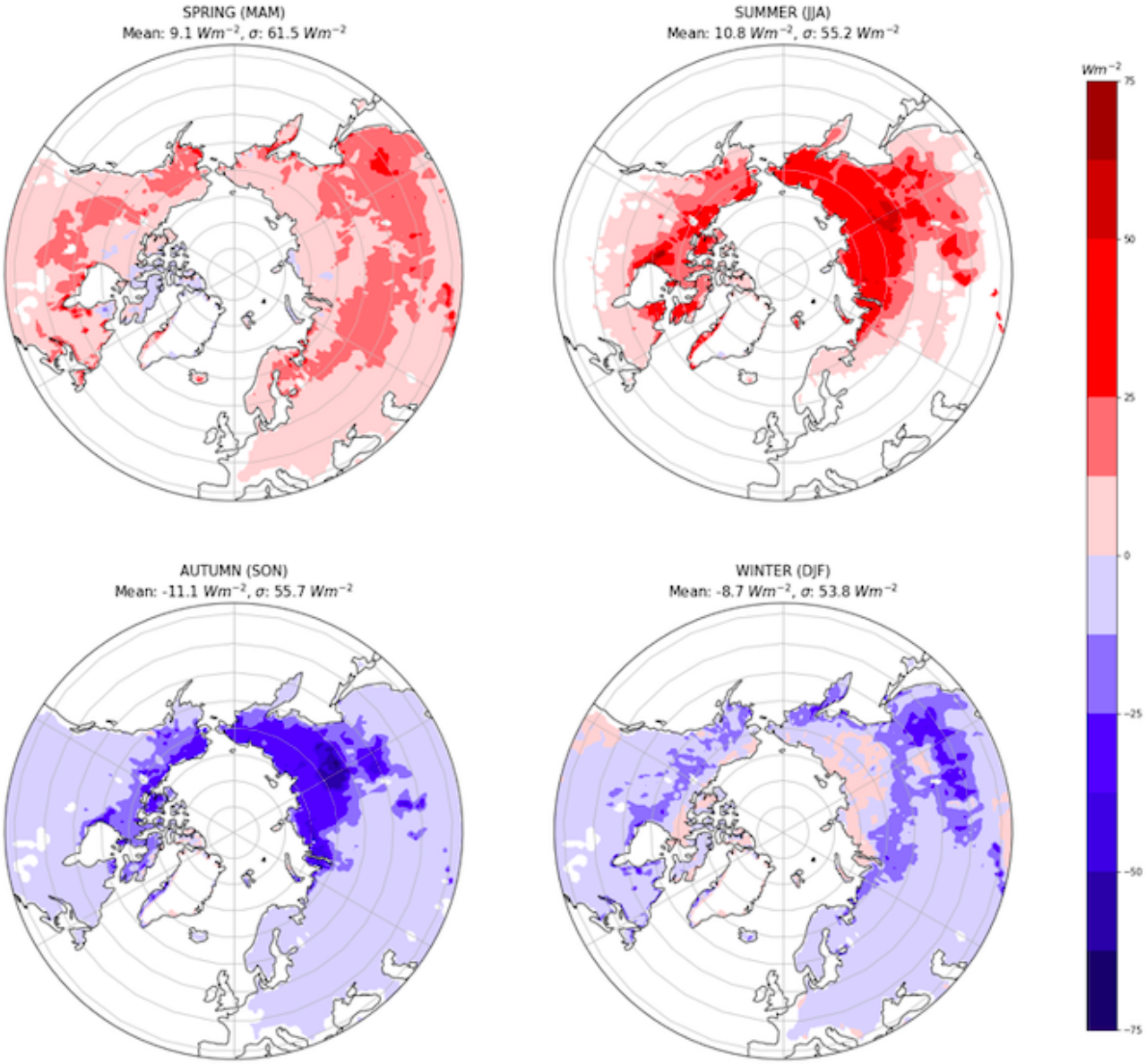


Figure 15: Seasonal means for LSHCT for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The integration method used is the Riemann integration. The units are Wm^{-2}*, zero is shown in white.*

Seasonal Average (Trapezoidal Integration)

Fig. 16 shows the seasonal averages for *LSHCT* for all years from 2000 to 2018, calculated with trapezoidal integration.

On a first sight, figures Fig. 15 and Fig. 16 show the same structures and distributions of positive and negative values for the tendency of latent heat content. In order to better identify the differences between the integration methods, the difference between the two methods is calculated and shown in Fig. 17.

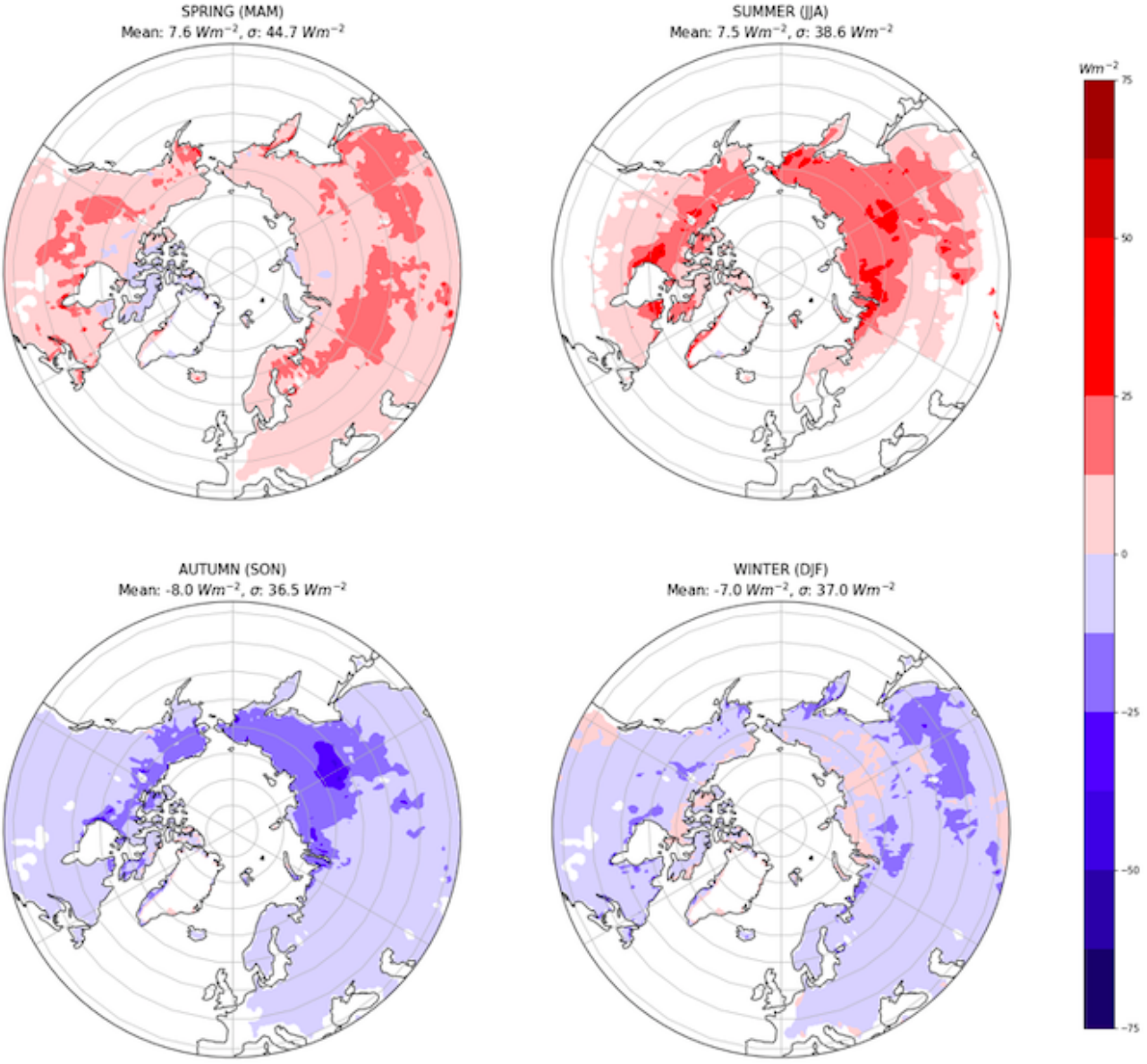


Figure 16: Seasonal means for LSHCT for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The integration method used is the trapezoidal integration. The units are Wm^{-2}*, zero is shown in white.*

Differences between integration methods

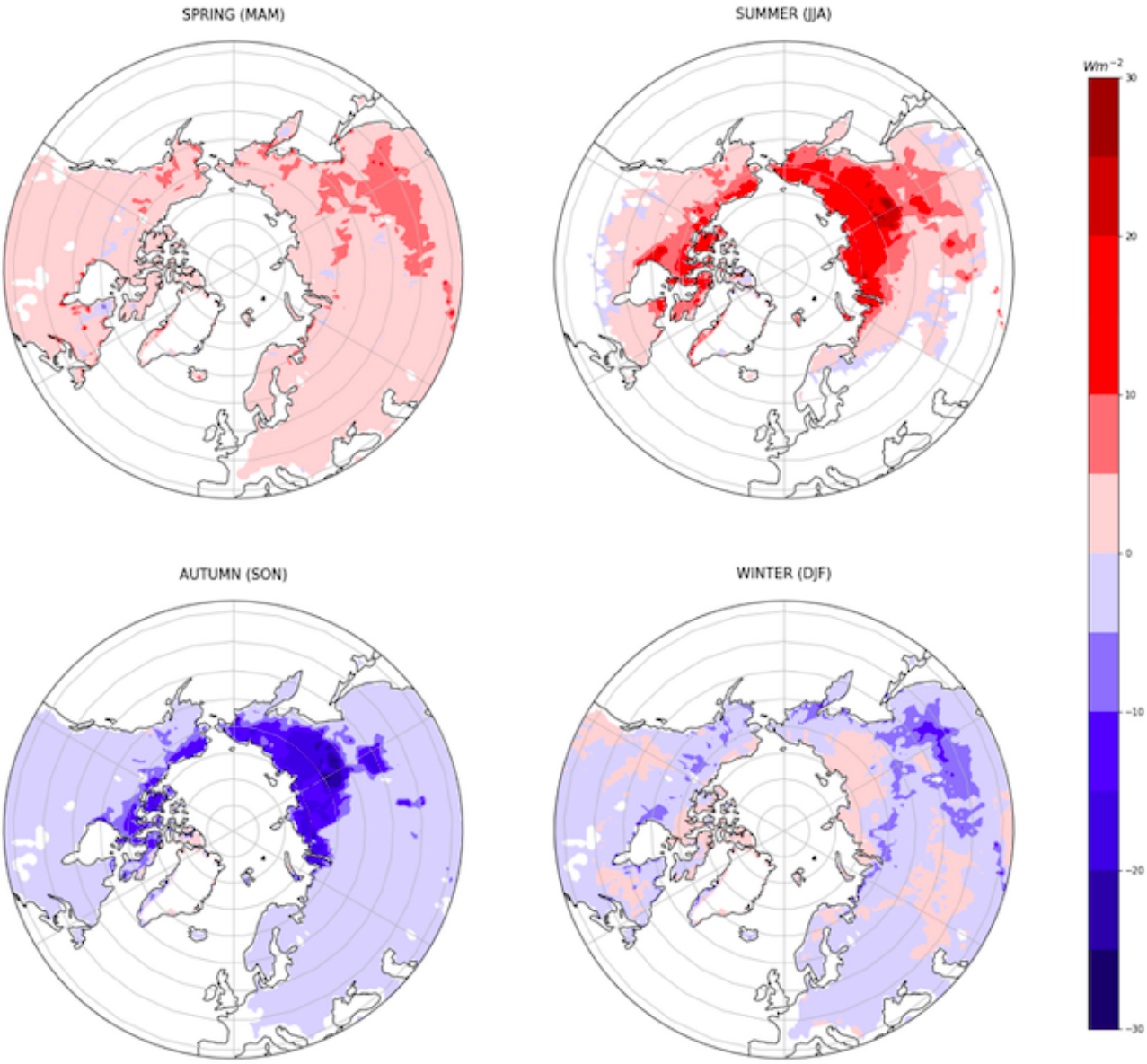


Figure 17: Riemann integration minus trapezoidal integration for $LSHCT$. Seasonal means for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white. Note the different color scale compared to Fig. 15 and Fig. 16.*

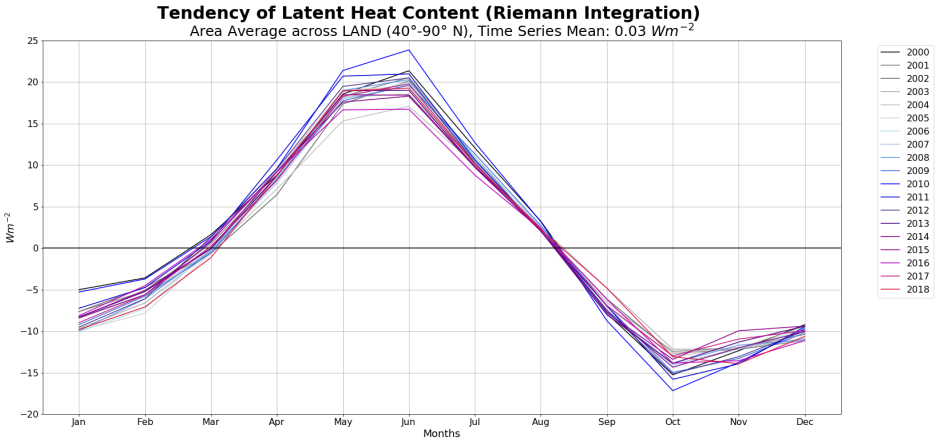
Fig. 17 shows the result of subtracting the seasonal means for $LSHCT$ of trapezoidal integration (Fig. 16) from Riemann integration (Fig. 15). Positive values indicate that Riemann integration leads to higher values than trapezoidal integration and negative values show that trapezoidal integration delivers higher values than Riemann integration.

In spring and summer (Fig. 17, top left and top right) the difference are mainly positive. The greatest differences between integration methods occur in summer and autumn (Fig. 17, bottom left) in Asia. In winter and spring (Fig. 17, bottom right and top left), the value range of the differences is the smallest. In winter a mixed picture of positive and negative values appears.

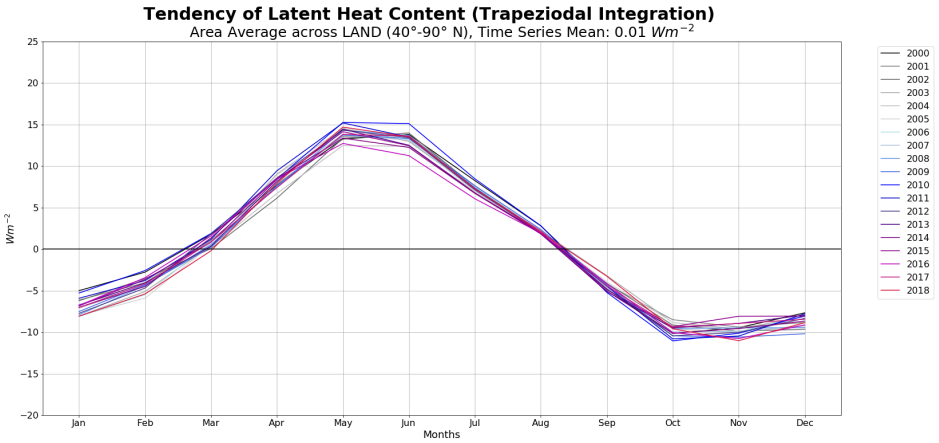
Considering the value range of figures Fig. 15 and Fig. 16, which extend from $-75\,\mathrm{Wm}^{-2}$ to $75\,\mathrm{Wm}^{-2}$, the difference between the integration methods (Fig. 13) is about $\pm 30\,\mathrm{Wm}^{-2}$, which corresponds to a deviation of about 40,0 %.

Annual Cycle

Fig. 18 shows the annual cycle of $LSHCT$. Each line shows one year, with the months of the year on the horizontal axis. The values on the vertical axis range from about $-18\,\mathrm{Wm}^{-2}$ to $25\,\mathrm{Wm}^{-2}$ for the Riemann integration and from about $-10\,\mathrm{Wm}^{-2}$ to $15\,\mathrm{Wm}^{-2}$.
A spatial mean was made for the land surface from 40 °N to 90 °N. The result for Riemann integration is given in Fig. 18a that for trapezoidal integration in Fig. 18b.



(a) Annual cycle of $LSHCT$ calculated by Riemann integration.



(b) Annual cycle of $LSHCT$ calculated by trapezoidal integration.

Figure 18: Annual cycles of $LSHCT$ for both integration methods. Every year is one curve. A spatial mean was made for the land surface from 40 °N to 90 °N. The mean of the time series was calculated for the whole time series from 2000 to 2018.

It is noticeable in Fig. 14 that the annual cycles start in a range between $-10{,}0\,\mathrm{Wm}^{-2}$ and $-5\,\mathrm{Wm}^{-2}$ for both integration methods. The switch from negative to positive heat content is performed in the end of February for both integration methods. It should be noted that the transition is not the

same every year, but fluctuates in the range of about 2 weeks. The maximum of $LSHCT$ with about 25 Wm^{-2} occurs from May to June for Riemann integration and with about 15 Wm^{-2} for trapezoidal integration. After June in both figures the latent heat content decreases. Form October to November in both figures a not clearly defined minimum is reached. In Fig. 18a the minimum is lower than in Fig. 18b, the spread lies within a range of about 5 Wm^{-2}.
Both, Fig. 18a and Fig. 18b, show similar courses and the variation range of about 5 Wm^{-2} for the different years is also comparable for both figures.
The latent heat content should, like the thermal heat content, also disappear in the mean. With a time series mean of 0,03 Wm^{-2} (Riemann integration) and 0,01 Wm^{-2} (trapezoidal integration), the result is very close to zero and acceptable.

Uncertainties

The biggest source of uncertainty for the tendencies of latent soil heat content is the soil ice content. The soil ice content is determined as described in section 4.2 by the assumption that the soil water content is the soil ice content if the temperature is below 0 °C. Water and ice can be present at the same soil depth, which is also excluded in this thesis. There is either water or ice.
For determining $LSHCT$ the order of integration is important. First the integrals have to be calculated and afterwards the tendencies, due to the fact that the soil ice content is dependent on the soil temperature as discussed in section 4.2.

5.1.3 Tendency of the latent heat content of snow

The term for the tendency of the latent heat content of snow is calculated as described in section 2.7.3. The tendency of the snow water equivalent is calculated and multiplied by the density of water and the latent heat of fusion.

Seasonal average

The seasonally means for the seasons for all years from 2000 to 2018 is calculated and the result is displayed in Fig. 19.

In Fig. 19 on the top left, the seasonal average for spring is given. In spring the area average of 2,3 $\pm$ 12,8 Wm^{-2} shows that in most parts of the chart, the snow mass term is positive. A positive snow mass term means, that the tendency of snow water equivalent is negative, and so snow melts. Only in a small area in the north of Hudson Bay is negative, where snow mass increases. It is noticeable that over Greenland ST is zero. So in Greenland, changes of snow mass only occur at the edges for all seasons. This is a reason why this area is masked in the other plots as mentioned above. The biggest values appear over the Rocky Mountains and Norway.

In summer (Fig. 19, top right), also in most of the areas the snow is melting, despite some negative areas the area mean is positive with 0,7 $\pm$ 5,0 Wm^{-2}. Along an area stretching from Europe to the west of Asia, it can be seen that there is no change in ST. This is probably because of the fact that the snow in low altitudes is already molten.

On the bottom left of Fig. 19 it is to notice that the area average is negative with $-1{,}2 \pm 4{,}5\,Wm^{-2}$. It would be expected that the snow mass starts to increase in autumn.

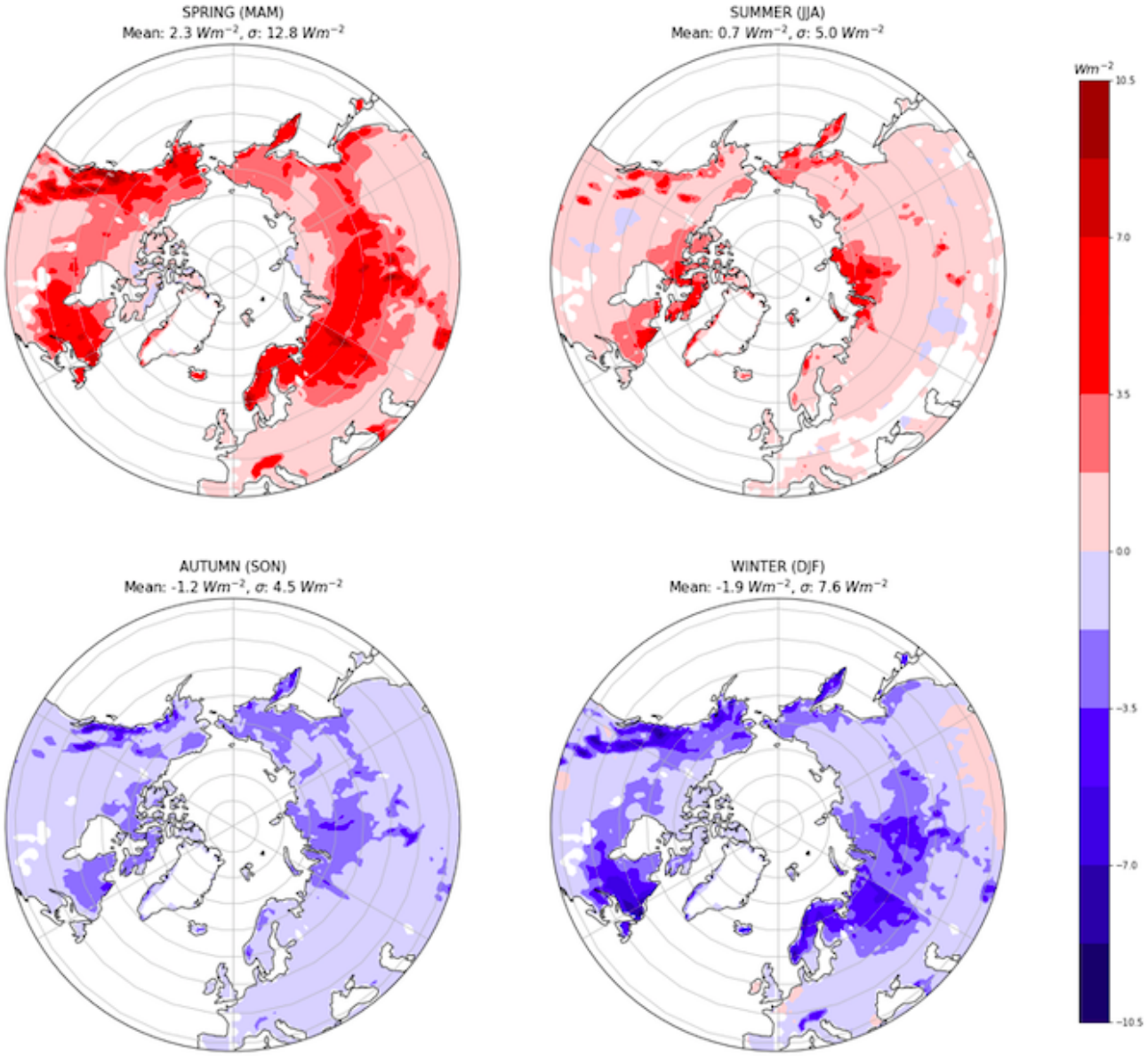


Figure 19: Seasonal means for ST for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white.*

In winter the area average of the chart on the bottom right in Fig. 19 is with $-1{,}9 \pm 7{,}6\,\mathrm{Wm}^{-2}$ still negative. The biggest changes in ST occur in the Rocky Mountains, the area around Hudson Bay, Scandinavia and in the northern parts of Asia.

Annual cycle

Fig. 20 shows the annual cycle of snow mass for all considered years. The range of values is between $-2{,}2\,\mathrm{Wm}^{-2}$ and $4{,}1\,\mathrm{Wm}^{-2}$, the spread between the years is up to $1{,}8\,\mathrm{Wm}^{-2}$ around the peak.
The annual cycle starts in January with negative values around $2{,}0\,\mathrm{Wm}^{-2}$. The sign changes in a range from end of February to mid of March from negative (increase of snow mass) to positive (decrease of

snow mass). The peak is in May, whereby the fluctuation range of the individual years is greatest here. In August the decrease of snow mass turns into an increase and reaches its minimum in December.

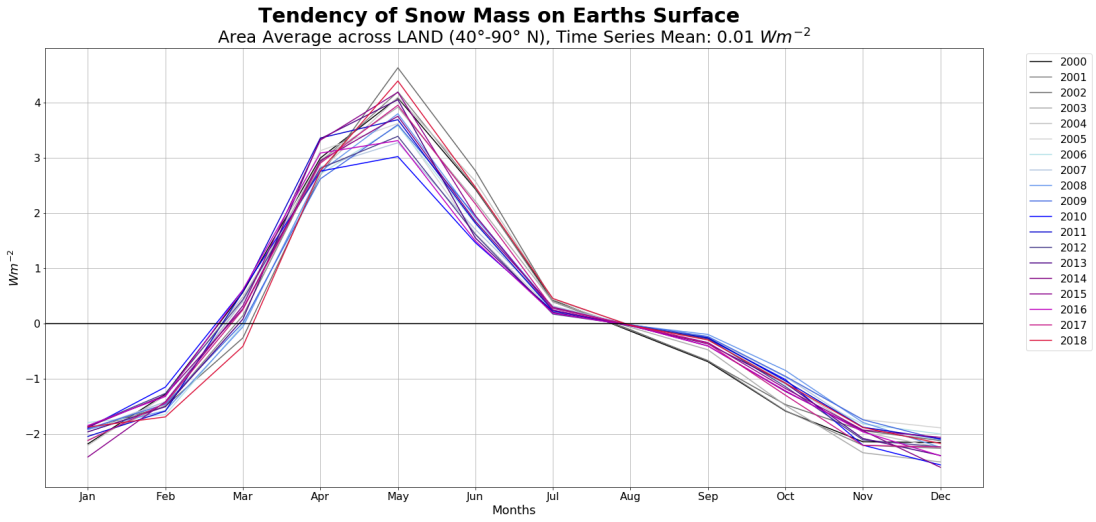


Figure 20: Annual cycle of ST. Every year is one curve. A spatial mean was made for the land surface from 40 °N *to* 90 °N. *The mean of the time series was calculated for the whole time series from 2000 to 2018.*

5.1.4 Snowfall

The term for the snowfall is calculated as described in section 2.7.4. To get the snowfall rate is multiplied by the latent heat of fusion (see section 2.7.4).

Seasonal average

The seasonally means for the seasons for all years from 2000 to 2018 is calculated and the result is displayed in Fig. 21.

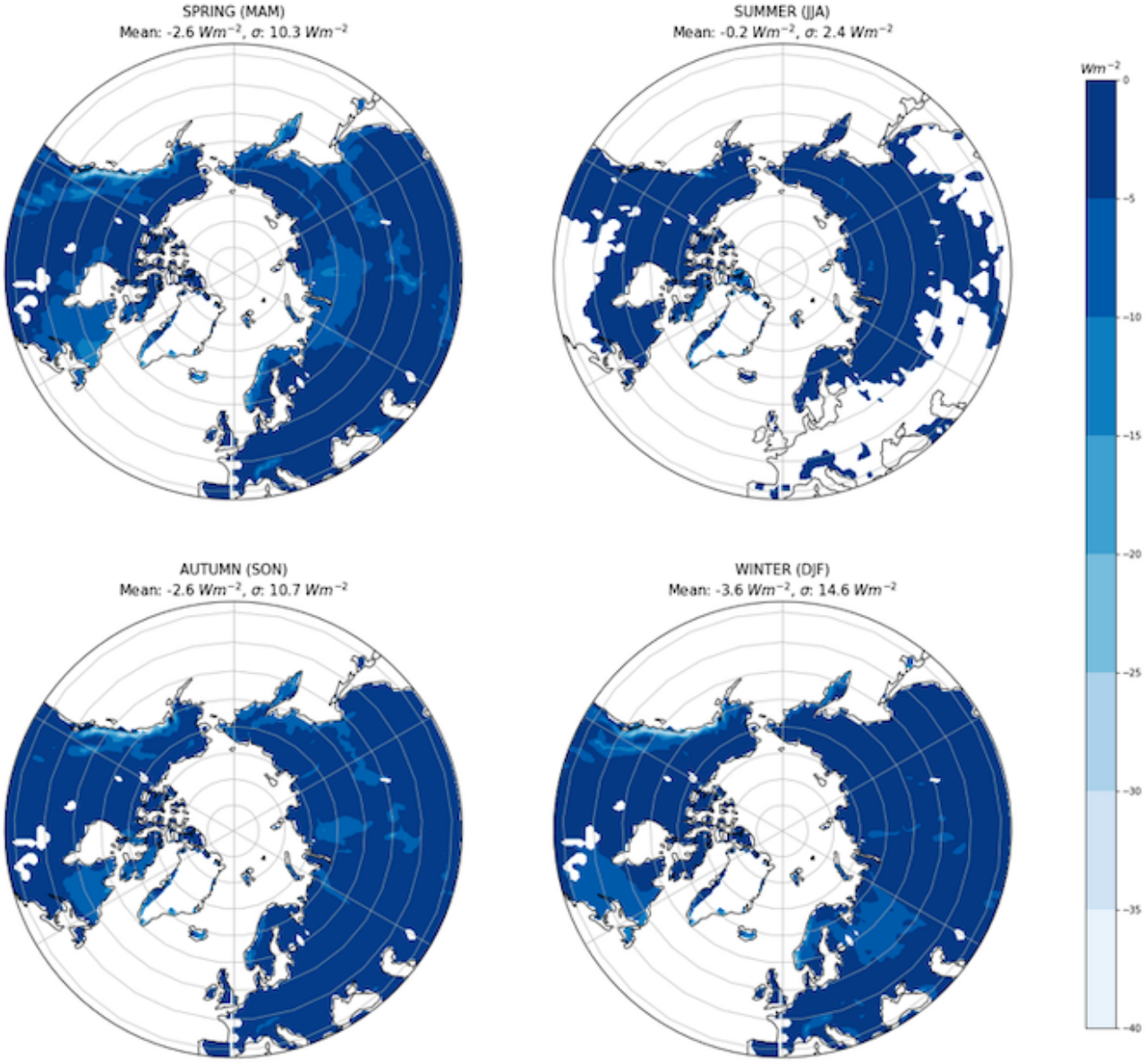


Figure 21: Seasonal means for SF for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white.*

The snowfall term in Fig. 21 is always negative, due to the chosen definition in section 2.7.4. When looking at the seasons, it can be seen that the lowest values occur over the Rocky Mountains, in the west of Hudson Bay and in some parts of Asia and in winter in parts of Europe. It is noticeable that in summer, the snowfall term disappears (as expected) further to the south of the chart.

Annual Cycle

Fig. 22 shows the annual cycle of snowfall term. At the beginning and the end of the year the spread between the years is the greatest with about 1,0 Wm^{-2}. The snowfall term reaches its maximum between June and August, dependent on the year under consideration.

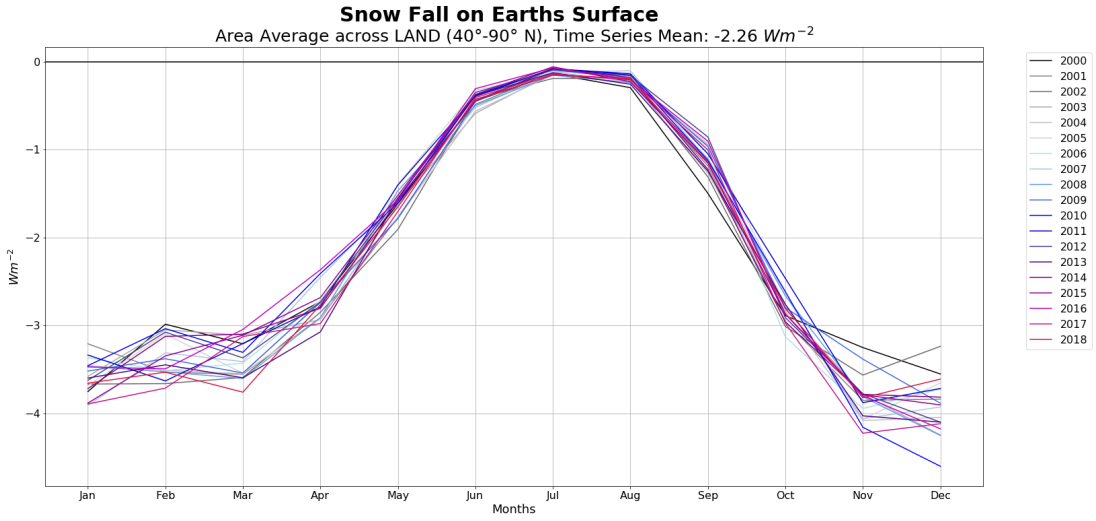


Figure 22: Annual cycle of SF. Every year is one curve. A spatial mean was made for the land surface from 40 °N *to* 90 °N. *The mean of the time series was calculated for the whole time series from 2000 to 2018.*

5.1.5 Cold Snowfall

The term for the cold snowfall is calculated as described in section 2.7.5. The snowfall rate is multiplied by the difference from temperature of the falling snow to a reference temperature of 0 °C and then multiplied by the specific heat capacity of snow (see section 2.7.5).

Seasonal average

Fig. 23 shows the seasonal means for the cold snowfall term for all years from 2000 to 2008. The distribution is similar to snowfall term (Fig. 21), but the term itself is significantly less. In summer it is only present in the most northern parts of northern America and Asia. In autumn it is apparent that in Europe the alpine region the surface also loses energy due to cold snowfall.

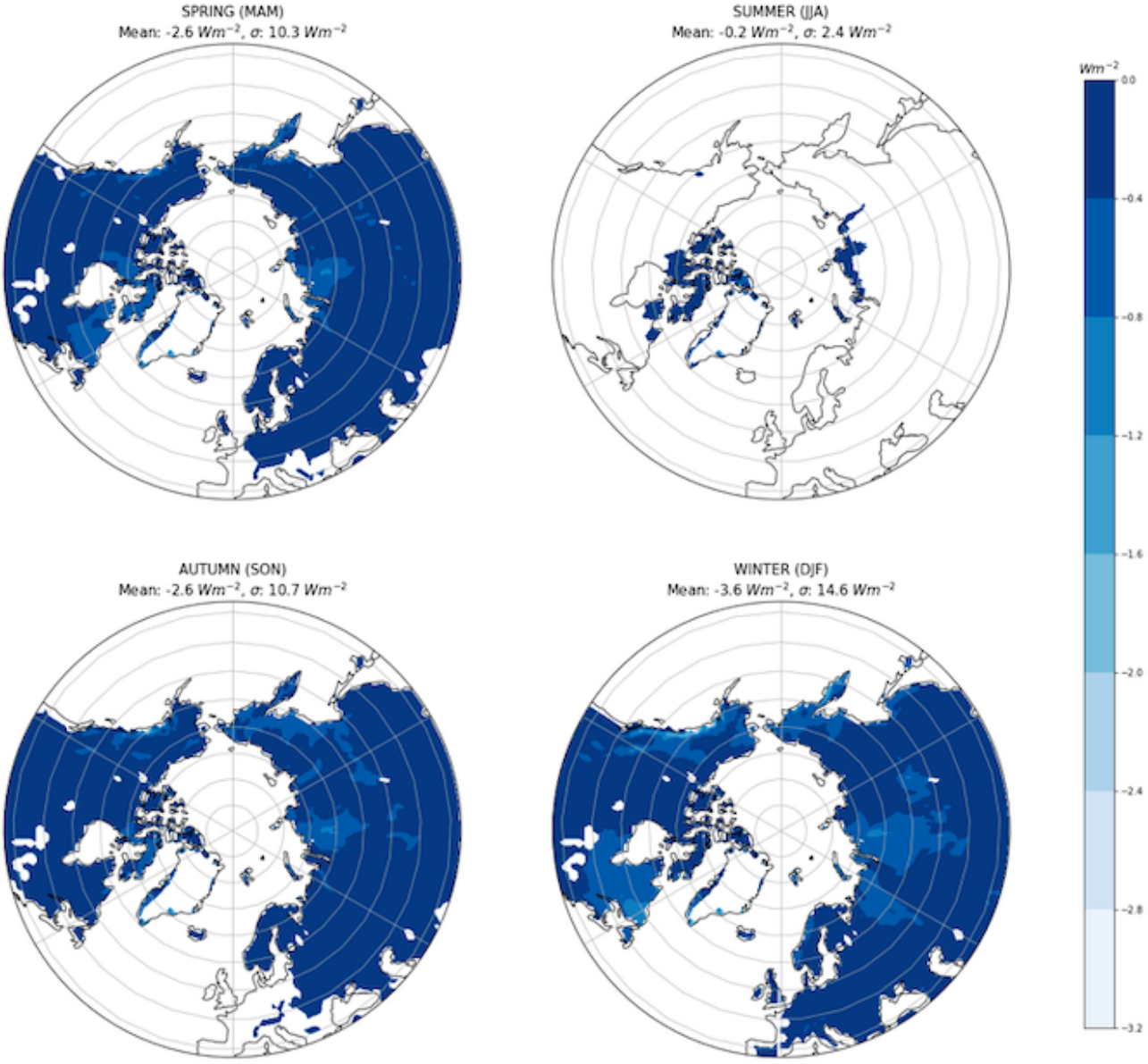


Figure 23: Seasonal means for CSF for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white.*

Annual Cycle

Fig. 24 views the annual cycle of cold snowfall with a curve for every year. It is recognizable that at the beginning and end of the year the spread is the largest. However, it is important to not forget that the value range is only between 0,0 Wm^{-2} and $-0{,}4\,\mathrm{Wm}^{-2}$ anyway. There is no significant maximum, it ranges form end of May to August.

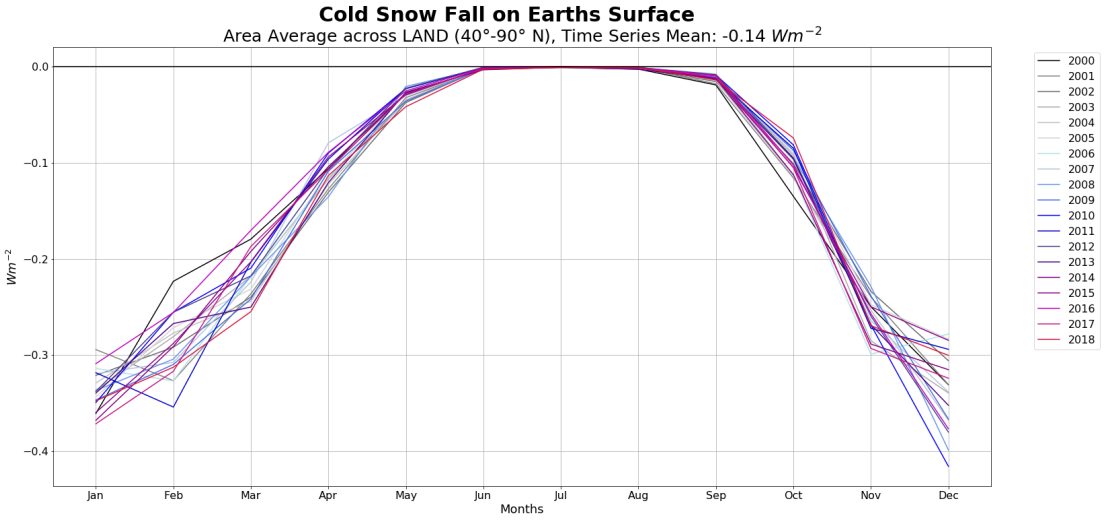


Figure 24: Annual cycle of CSF. Every year is one curve. A spatial mean was made for the land surface from 40 °N *to* 90 °N. *The mean of the time series was calculated for the whole time series from 2000 to 2018.*

5.1.6 Rainfall

The effect of rain falling on the surface can be denoted by multiplying the specific heat capacity of water by the rainfall rate and the difference between precipitation temperature and a reference temperature of 0 °C, as mentioned in section 2.7.6.

Seasonal average

The seasonal means for the rainfall term for all years from 2000 to 2018 are displayed in Fig. 25. In spring, nearly everywhere rain falls to the surface. Northern parts of Asia remain unaffected of rain fall. A maximum can be found in Europe, the Rocky Mountains and southern parts of Canada. In summer, the rainfall has its maximum in the southeastern parts of Asia. In autumn, the rainfall term decreases and disappears in the most parts of the area. Europe and parts of northern America remain effect by rainfall in winter. Here it should also be mentioned that this term is small compared to others.

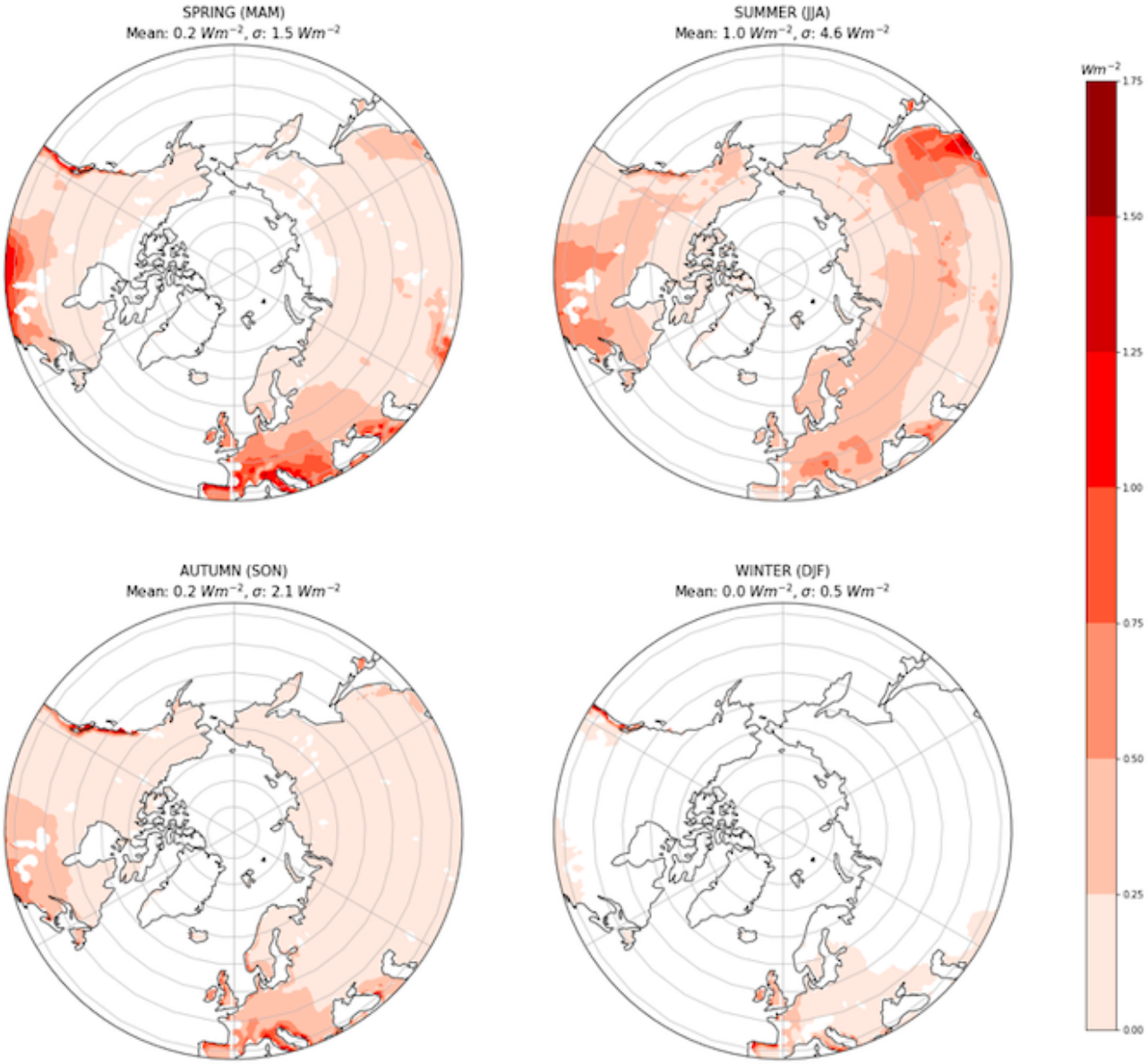


Figure 25: Seasonal means for RF for all years from 2000 to 2018. On the top left the seasonal mean for March-April-May is given, on the top right for June-July-August, on the bottom left for September-October-November and on the bottom right for December-January-February. The units are Wm^{-2}*, zero is shown in white.*

Annual Cycle

The annual cycle of the rainfall term is presented in Fig. 26. The maximum is June, with the greatest variation range of the years within about 0,2 Wm^{-2}. In the other months the fluctuation range is smaller. The minimum of rainfall is as expected during winter.

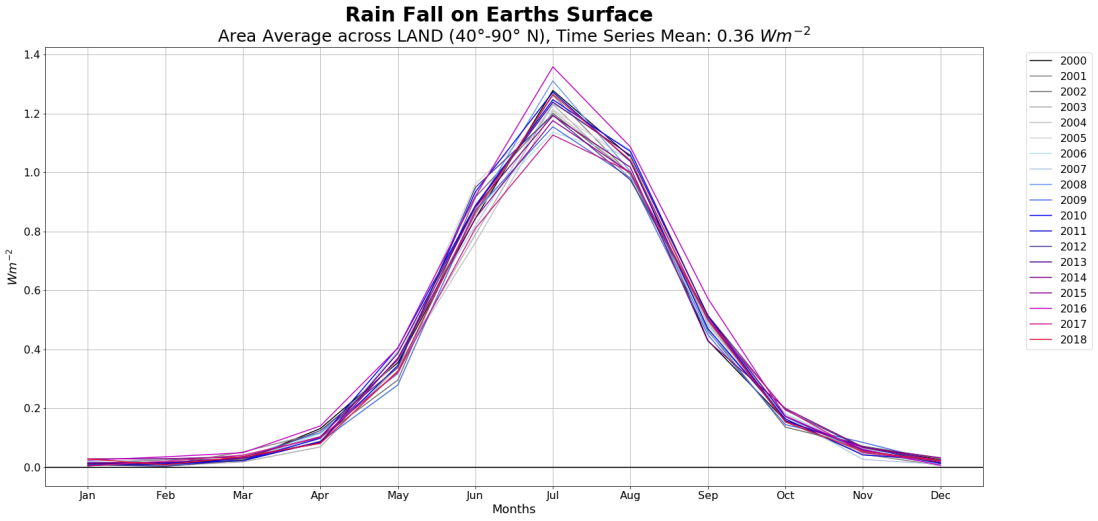


Figure 26: Annual cycle of RF. Every year is one curve. A spatial mean was made for the land surface from 40 °N *to* 90 °N. *The mean of the time series was calculated for the whole time series from 2000 to 2018.*

5.2 Comparison with reference curves

After discussing the order of magnitude, distribution and years of the single terms of equation 24, the focus is now on comparing the calculation methods of the net energy flux.
Fig. 27 shows the time-related course of F_S with all four calculation methods. It can be seen that F_S calculated by the Riemann integration covers the largest value range, followed by F_S calculated by trapezoidal integration. The values calculated with equation 24 are the smallest.

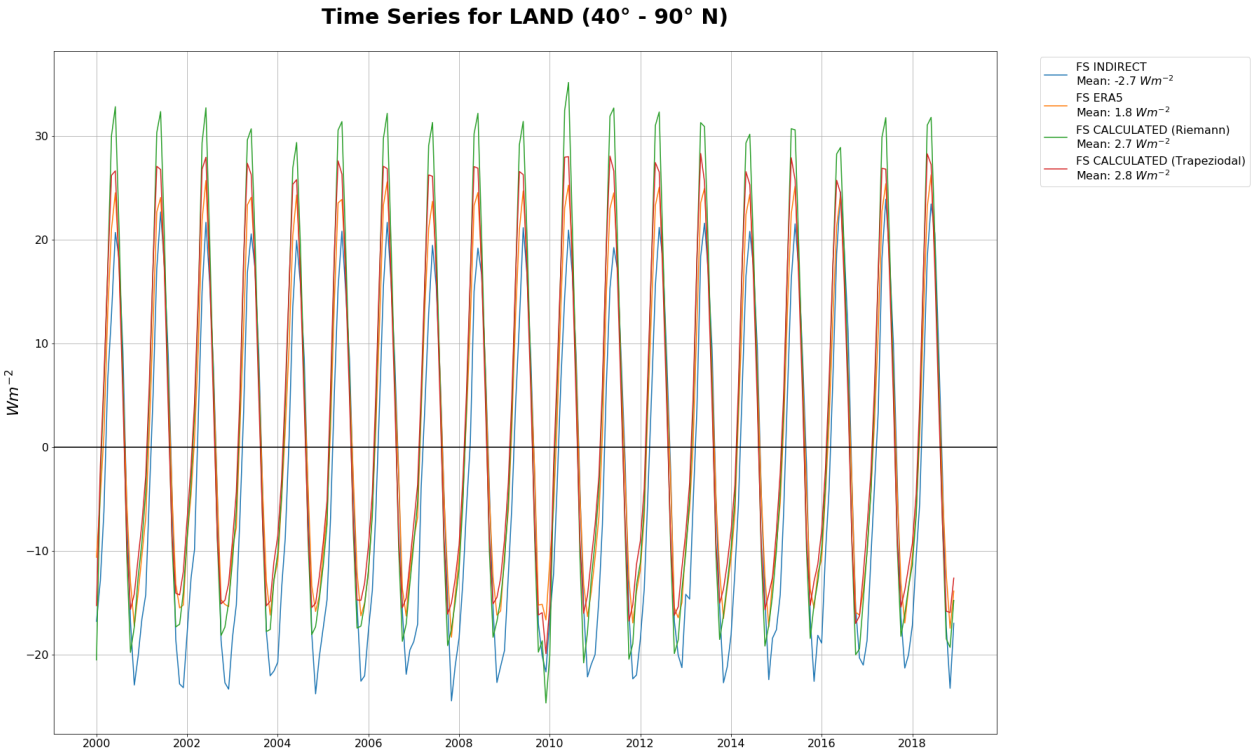


Figure 27: Time series of F_S calculated by Riemann integration (green), trapezoidal integration (ref), directly form ERA5 and calculated by the atmospheric column.

It is difficult to derive results from the representation in Fig. 27. But what can be seen is that it seems that the time series is stationary without trends. Therefore in the following, annual cycles with monthly means for all years from 2000 to 2018 can be considered.

5.2.1 Annual cycle for land surface from 40 °N to 90 °N

Fig. 28 shows the annual cycle of the calculated F_S (icons in black) with the corresponding terms (colored bars) used for the calculation of F_S by equation 24. The signs next to the colored bars in the legend denote if the bar is added or subtracted in equation 24. The integration method used is the Riemann integration. To get one curve, the average for every month for all years (2000 to 2018) was calculated.

The tendency of thermal soil heat content ($TSHCT$) is the red colored bar. At the beginning of the year this term is negative, the soil temperature tendency is negative and thus the soil cools down. In March it is noticeable that $TSHCT$ is very small, but already positive. From March to August, $TSHCT$ is positive, with its maximum impact to the whole equation in June with nearly 10 Wm^{-2}. In September the soil temperature tendency is negative and so the soil cools down again until the

end of the year. Just by taking a look one can recognize that $TSHCT$ provides the second largest contribution to F_S.

The largest contribution to the calculated F_S is ascribed to the tendency of latent soil heat content colored ($LSHCT$) in royal blue. $LSHCT$ is negative (meaning a positive tendency of soil ice content leading to an increase of soil ice content) in January and February. In March this term is infinitesimal and does not affect the calculation of F_S. Form September to December the soil ice content is increasing.

The term concerning the change of the snow mass (ST, grey colored bar) on Earth's surface behaves in the same way as the tendency of latent soil heat content. ST is very small in July and September and disappears in August. In winter, the mass of snow is increasing and in summer, the snow mass is decreasing because of melting.

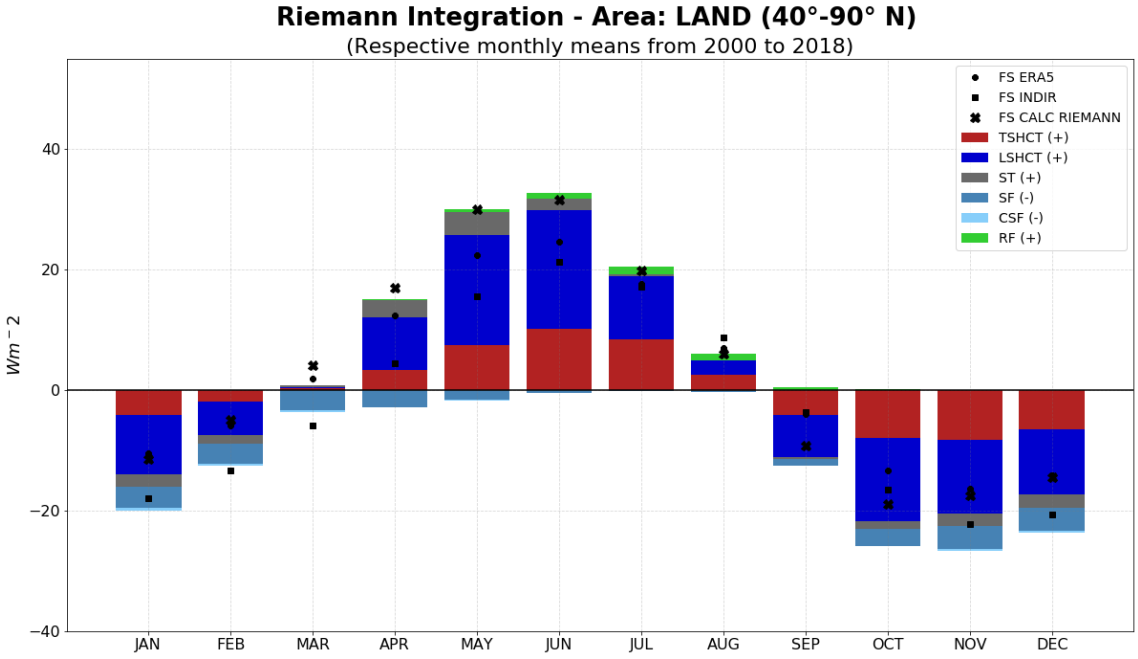


Figure 28: Annual cycle of F_S for different calculation methods (icons in black) including the influence of each term in equation 24 in colored bars. The mean for the months was calculated for the whole time series from 2000 to 2018. The integration method used is the Riemann integration.

Another term with great impact in Fig. 28 is the the tendency of the latent heat content of snow (ST, grey-blue colored bar). It is the biggest in winter and becomes smaller in summer, with its minimum impact in July and August.
The term for the cold snowfall (CSF, light blue bar), is very small and has no significant influence in this figure.
Rainfall (RF, green bar) is only available in summer, with a maximum in July. It is recognizable that this term is also very small.

The 'x' icon in black shows the result for calculating F_S via equation 24 (hereafter $F_{S_{CALC,Riemann}}$), the '+' displays the result of calculation directly from the parameters given in ERA5 with equation 17 (hereafter $F_{S_{ERA}}$). A circular symbol represents the F_S calculated via atmospheric column as described in section 3.2 (hereafter $F_{S_{INDIR}}$).

It is apparent that the symbols do not show a perfect match. It is difficult to draw more precise conclusions in Fig. 28, so Fig. 29 is used for a closer look.

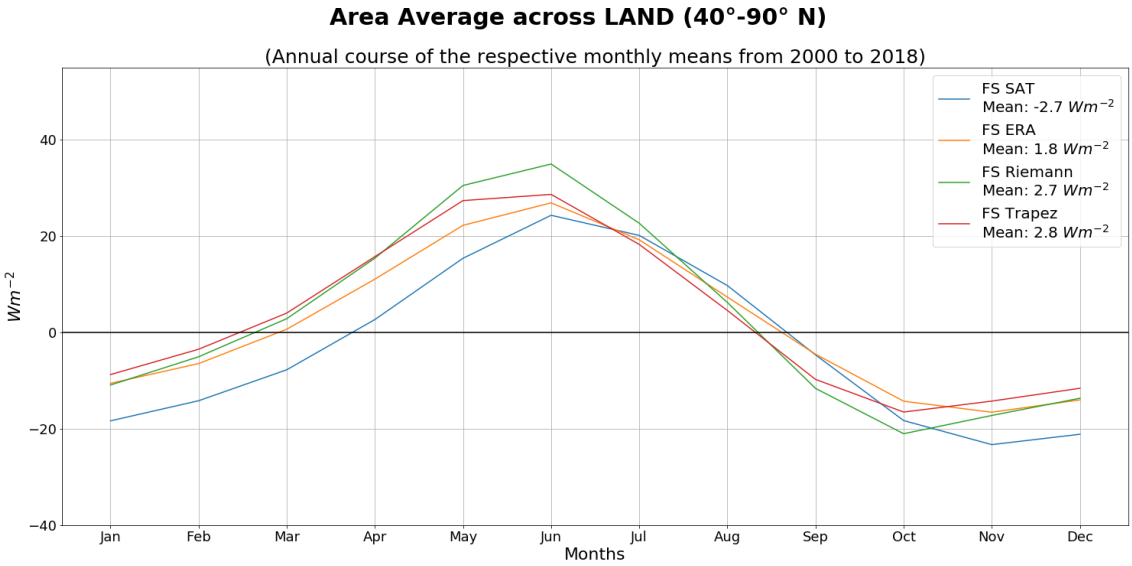


Figure 29: Annual cycle of F_S for different calculation methods. The mean for the months was calculated for the whole time series from 2000 to 2018.

Fig. 29 shows the icons presented in Fig. 28 connected by a line, with the difference that Fig. 29 also includes the result for equation 24 calculated by trapezoidal integration (hereafter $F_{S_{CALC,Trapez}}$, red line). The green line shows the result for $F_{S_{CALC,Riemann}}$, the yellow line is $F_{S_{ERA}}$. The blue line delivers the result for $F_{S_{INDIR}}$.

It is obvious that $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ show significant differences, whereas $F_{S_{CALC,Trapez}}$ fits better to the reference lines of $F_{S_{ERA}}$ and $F_{S_{INDIR}}$.
In general, all lines are within the same fluctuation range, with a minimum in winter of about $-20\,\mathrm{Wm}^{-2}$ to a maximum in summer of $30\,\mathrm{Wm}^{-2}$.
Furthermore the peaks of $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ seem to be in phase compared to $F_{S_{ERA}}$ and $F_{S_{INDIR}}$, where the peak $F_{S_{CALC,Trapez}}$ is less marked than the others. What can also be seen is that the change from negative to positive for $F_{S_{INDIR}}$ takes place in mid-March, in the other curves this change occurs earlier (mid to late February). This can occur due to the chosen method of tendency calculation and will be discussed in section 5.5.
Another fact that should be mentioned is the mean of the annual cycle. While it is negative for $F_{S_{INDIR}}$, positive values occur for $F_{S_{ERA}}$, $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$, where those calculated by equation 24 with $2{,}7\,\mathrm{Wm}^{-2}$ and $2{,}8\,\mathrm{Wm}^{-2}$ are the greatest. The value for $F_{S_{ERA}}$ lies with $1{,}8\,\mathrm{Wm}^{-2}$ in between.

5.2.2 Annual cycles for cutouts

Asia

Asia is defined as an area from 50 °E to 190 °E and from 10 °N to 90 °N.

Fig. 30 depicts the annual cycle of each term of equation 24 for the region Asia. In the beginning of the year the tendency of thermal heat content is negative, in March the $TSHCT$ is positive, so Asia gains

energy until August, with a maximum in June of about 8 Wm^{-2}. The $LSHCT$ shows a similar course. It is apparent that the greatest effects of soil ice melting are in May and June with about 20 Wm^{-2}. The tendency of snow mass (ST) influences the equation most during snow melt periods from April to June. The impact of snowfall (SF) is the biggest during winter, the one from the rainfall RF has its maximum in July and disappears in winter.

Looking at the icons, which represent the values for F_S, it is noteworthy that $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$ are very similar in March and August and differ in the remaining year. To have a closer look at this, Fig. 31 will be consulted.

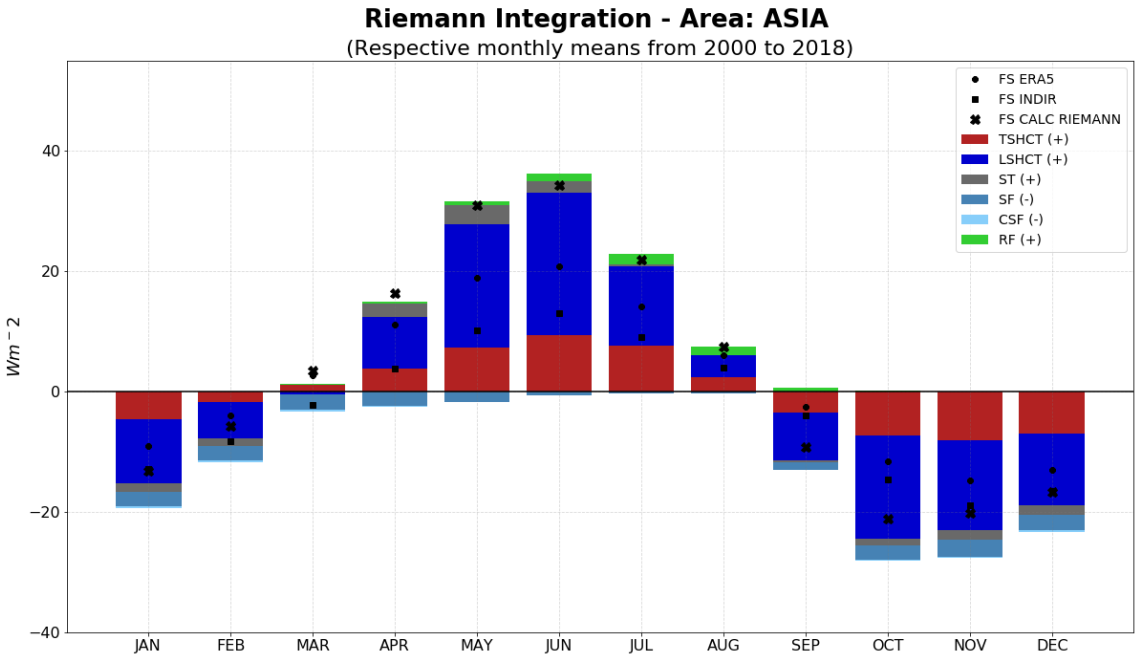


Figure 30: Annual cycle of F_S for different calculation methods (icons in black) including the influence of each term in equation 24 in colored bars for Asia. The mean for the months was calculated for the whole time series from 2000 to 2018. The integration method used is the Riemann integration.

The lines in Fig. 31 show $F_{S_{INDIR}}$ in blue, $F_{S_{ERA}}$ in yellow, $F_{S_{CALC,Riemann}}$ in green and $F_{S_{CALC,Trapez}}$ in red. It is noticeable that $F_{S_{ERA}}$, $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ are similar at the beginning of the year, but in summer the difference between $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$ is about 20 Wm^{-2}, which leads to the assumption that the $LSHCT$, which is the major part of the equation (Fig. 30), may be overestimated. $F_{S_{INDIR}}$ shows a phase shift, which could already be seen in Fig. 29 and is going to be discussed in section 5.5. The mean of $F_{S_{INDIR}}$ is negative with a value of −3,1 Wm^{-2}, while the means for the other lines are thoroughly positive. $F_{S_{CALC,Riemann}}$ with 2,4 Wm^{-2} and $F_{S_{CALC,Trapez}}$ with value of 2,5 Wm^{-2}, $F_{S_{ERA}}$ is slightly lower with 1,6 Wm^{-2}.

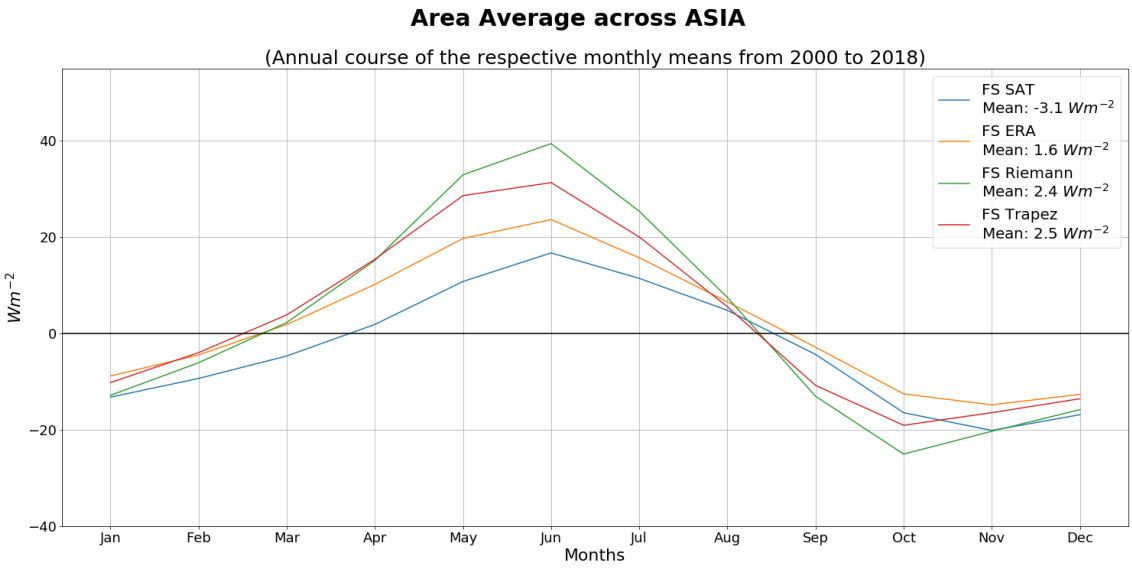


Figure 31: Annual cycle of F_S for different calculation methods for Asia. The mean for the months was calculated for the whole time series from 2000 to 2018.

Europe

Europe is selected by −30 °W to 50 °E and form 35 °N to 90 °N. Parts of Greenland in the selected area where masked and did not influence the area.

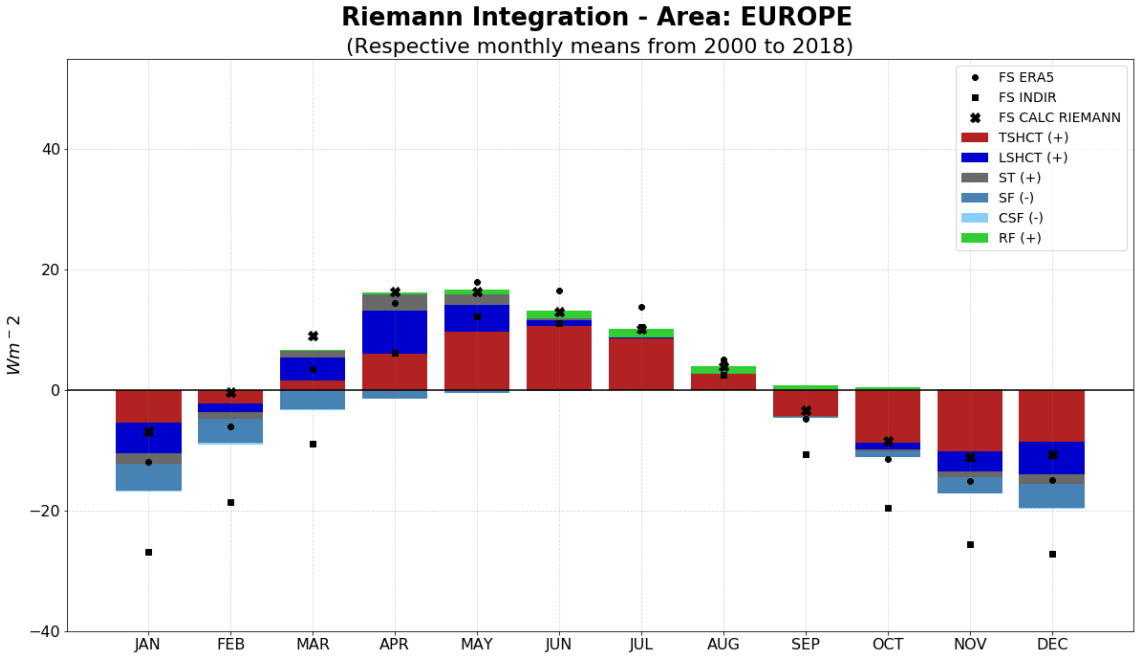


Figure 32: Annual cycle of F_S for different calculation methods (icons in black) including the influence of each term in equation 24 in colored bars for Europe. The mean for the months was calculated for the whole time series from 2000 to 2018. The integration method used is the Riemann integration.

Looking at Fig. 32 in comparison to figures Fig. 28 and Fig. 30, the month April is outstanding. In April the maximum of curve is reached. The course of the terms is similar, but the term ST is of interest. Already in March, snow melt is starting and reaches its maximum in April.

Considering Fig. 33 shows an interesting picture. $F_{S_{ERA}}$ and $F_{S_{INDIR}}$ show a kind of plateau in summer. $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ show a very similar course, with an earlier maximum as the reference curves. Here again a phase shift can be seen, as already shown in Fig. 29 and Fig. 30. The distribution of the time series averages is also relevant. $F_{S_{INDIR}}$ has an average of $-7{,}9\,\mathrm{Wm}^{-2}$, the others are positive, with $0{,}6\,\mathrm{Wm}^{-2}$ for $F_{S_{ERA}}$, $2{,}4\,\mathrm{Wm}^{-2}$ and $2{,}5\,\mathrm{Wm}^{-2}$ for $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$.

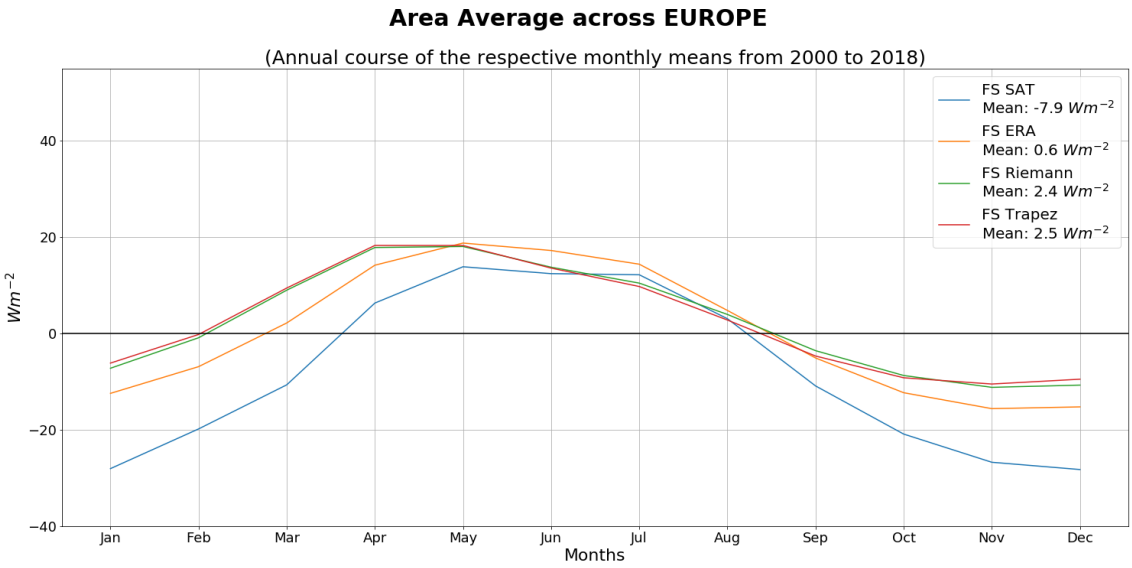


Figure 33: Annual cycle of F_S for different calculation methods for Europe. The mean for the months was calculated for the whole time series from 2000 to 2018.

Northern America

Northern America is defined as an area from 169 °W to 10 °W and from 10 °N to 90 °N. Parts from Africa and Iceland in this extract where masked and removed.

The term with the greatest influence is still $LSHCT$, followed by $TSHCT$. It is noticeable that $TSHCT$ is infinitesimal in March. The effect of the tendency of snow mass on surface (ST) is the greatest in May, but also of importance in April and June. In winter, ST is also present, but the accumulation of snow is not as big as the melting. There is also a large component of SF during winter, which is minimal in July.

Regarding the icons, it is notable that throughout the year the differences between the calculation methods are not as big as for example in Asia.

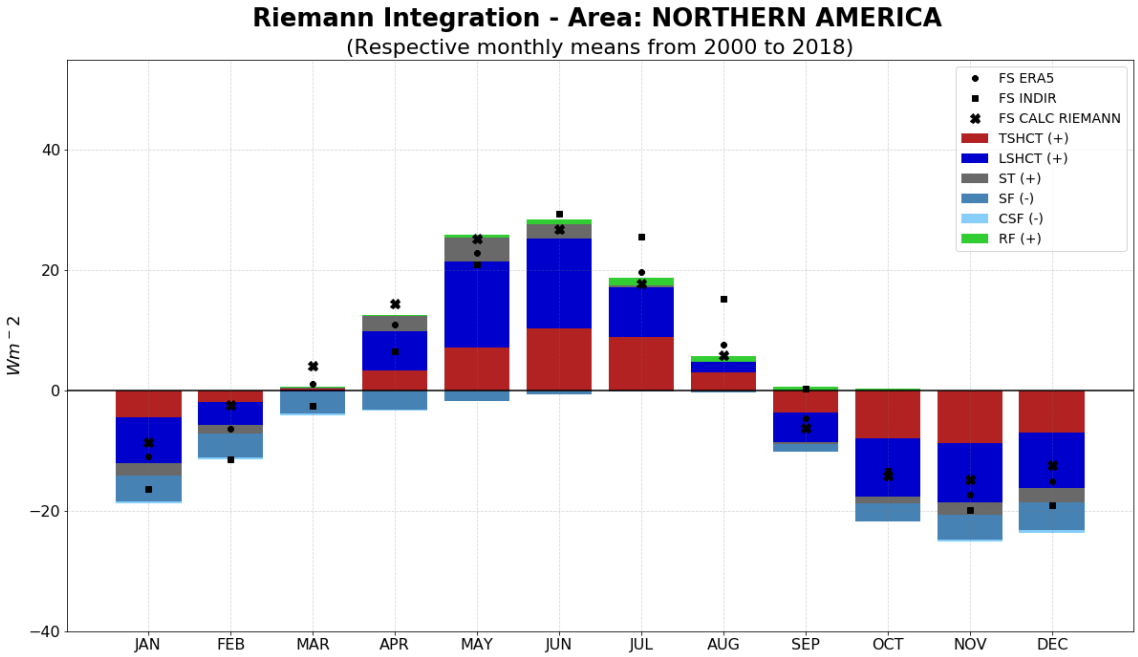


Figure 34: Annual cycle of F_S for different calculation methods (icons in black) including the influence of each term in equation 24 in colored bars for northern America. The mean for the months was calculated for the whole time series from 2000 to 2018. The integration method used is the Riemann integration.

$F_{S_{ERA}}$, $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ fit together very well, $F_{S_{INDIR}}$ shows the phase shift as mentioned above. In Fig. 35 it is remarkable that the maximum of $F_{S_{INDIR}}$ reaches 30 Wm^{-2}, and the average is also positive with 1,3 Wm^{-2}. Also $F_{S_{ERA}}$ is in average positive (1,7 Wm^{-2}), and so do $F_{S_{CALC,Riemann}}$ (3,0 Wm^{-2}) and $F_{S_{CALC,Trapez}}$ (3,1 Wm^{-2}).

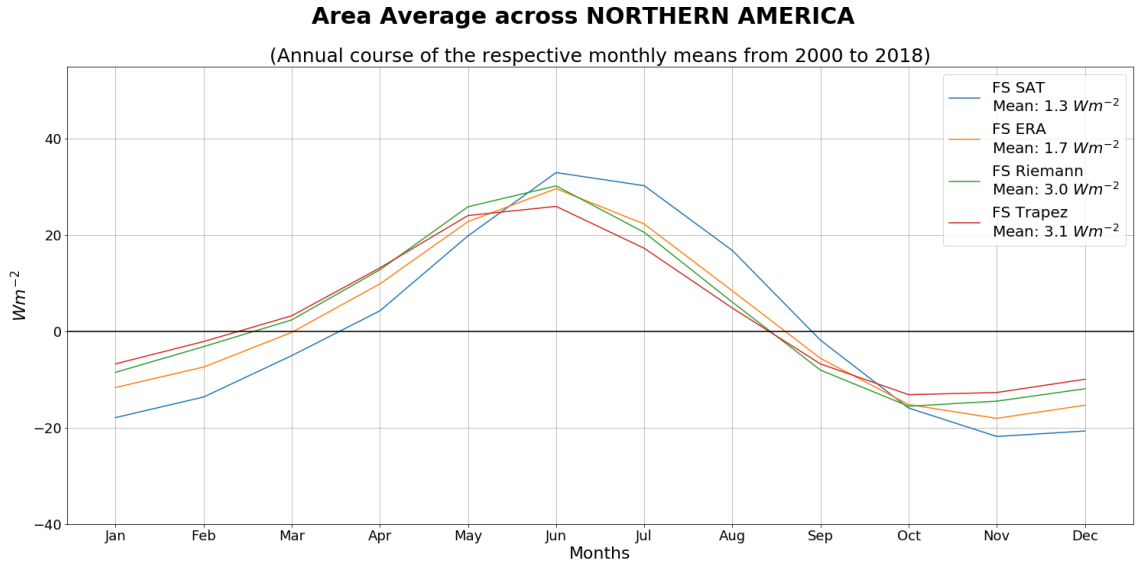


Figure 35: Annual cycle of F_S for different calculation methods for northern America. The mean for the months was calculated for the whole time series from 2000 to 2018.

Greenland

Greenland is cut out in between 10 °W and 45 °W, and in between 55 °N and 90 °N. Parts of land that do not belong to Greenland have been removed and not taken into account.

A closer look at Greenland led to the conclusion that areas that are glaciated should be masked out and excluded in the calculations and evaluations in this work. Here it was shown that these areas were assigned a constant value of SWE over the year. This affects the calculation of ST, which is determined by the temporal change of SWE (see section 2.7.3). This would mean that these ranges would then never show a trend and ST would therefore always disappear, which is not the case.

Taking a look at Fig. 36, it soon becomes clear that the major components of calculated F_S are the snowfall and the tendency of latent heat content. It is surprising that ST is very small in winter, which means that only little snow falls on the surface.

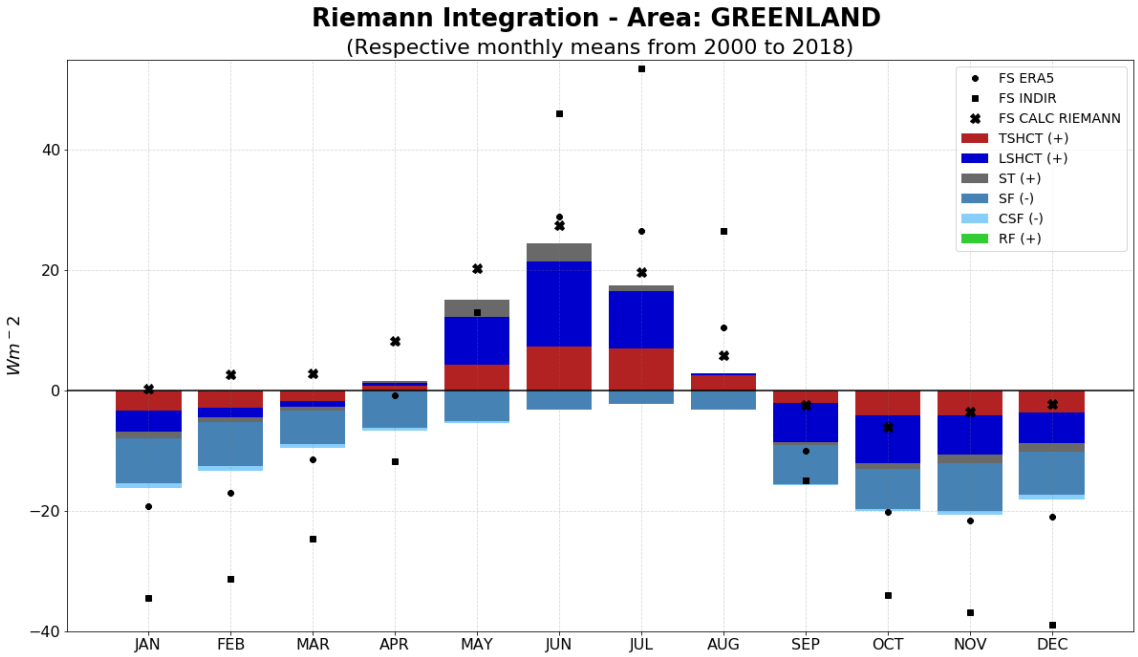


Figure 36: Annual cycle of F_S for different calculation methods (icons in black) including the influence of each term in equation 24 in colored bars for Greenland. The mean for the months was calculated for the whole time series from 2000 to 2018. The integration method used is the Riemann integration.

Due to the fact that $TSHCT/LSHCT$ and SF enter equation 24 with opposite signs, these compensate each other, in particular in the months January to March and from October until December, which can be seen in Fig. 37.

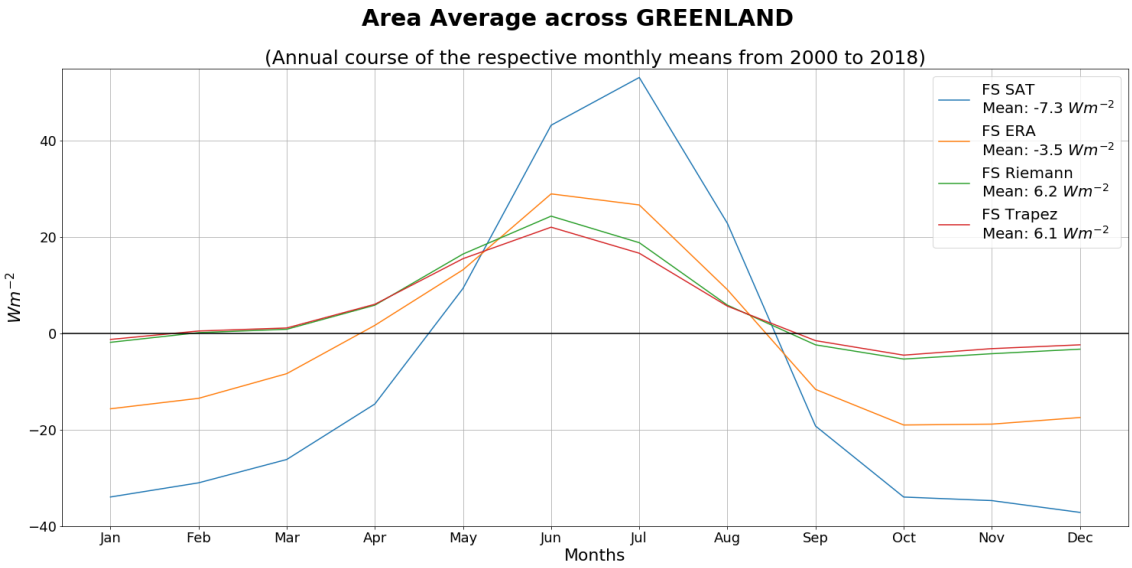


Figure 37: Annual cycle of F_S for different calculation methods for Greenland. The mean for the months was calculated for the whole time series from 2000 to 2018.

Fig. 37 shows the already described situation even more clearly through the line formation. The fact that in winter the net energy flux of the calculated curves is never negative may be caused by the effect that the tendency of thermal and latent heat content are compensated by ST. Another fact is that the greatest part of Greenland is masked, due to reasons given above. So Greenland only consists of a small strip at the edge of the island. Some investigations are going to be carried out in section 5.5.
Considering the average values, $F_{S_{INDIR}}$ provides $-7{,}3\,\mathrm{Wm}^{-2}$ similar to $-3{,}5\,\mathrm{Wm}^{-2}$ delivered by $F_{S_{ERA}}$. The values for $F_{S_{CALC,Riemann}}$ and $F_{S_{CALC,Trapez}}$ rank with $6{,}2\,\mathrm{Wm}^{-2}$ and $6{,}1\,\mathrm{Wm}^{-2}$.

5.2.3 Viewing in the area

In figures 28 to 37, the averages were made over big areas, so it could be helpful to discuss the differences of the calculation methods by viewing the area. Therefore seasonal means were calculated including the whole time series. Figures 38 to 41 show the differences for each season. In each case the upper left picture shows the difference between $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$, the upper right picture the difference $F_{S_{ERA}}$ and $F_{S_{CALC,Trapez}}$, the lower left the difference between $F_{S_{INDIR}}$ and $F_{S_{CALC,Riemann}}$ and finally the lower right picture gives the difference between $F_{S_{INDIR}}$ and $F_{S_{CALC,Trapez}}$.
Pictures in one line always look very similar. This is because of the fact that Riemann and trapezoidal integration produce very comparable results. The results of $F_{S_{INDIR}}$ is very noisy, what can be seen also in the differences.
Positive (red) values mean that $F_{S_{ERA}}$ or $F_{S_{INDIR}}$ provide higher values than $F_{S_{CALC,Riemann}}$ or $F_{S_{CALC,Trapez}}$, reversed for negative values.

In spring (Fig. 38) $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$, respectively $F_{S_{CALC,Trapez}}$, deviate over a large area by approximately $\pm$ 10 Wm^{-2}, which is small compared to the deviations of $F_{S_{INDIR}}$. The difference to $F_{S_{ERA}}$ is the biggest in Norway and in the Rocky Mountains. It is also noticeable that the deviations are often greatest along water-land borders, for example as around Europe. $F_{S_{ERA}}$ provides higher values in larger parts of Asia and northern America, while in Europe $F_{S_{ERA}}$ provides smaller values than $F_{S_{CALC,Riemann}}$ or $F_{S_{CALC,Trapez}}$. A similar picture is provided for the deviation of $F_{S_{INDIR}}$, with much higher values, but also much finer structures.

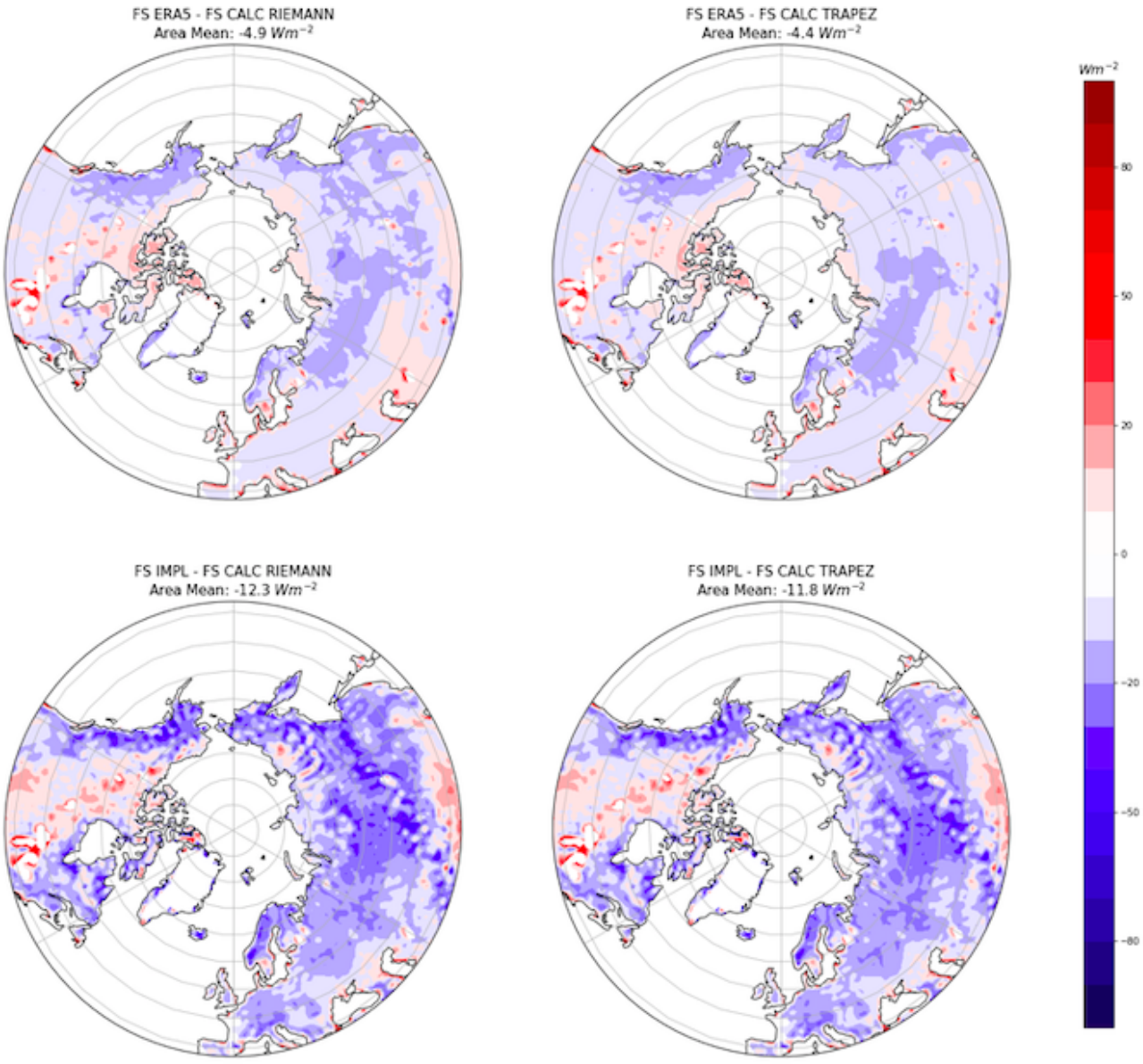


Figure 38: Differences between calculation methods as seasonal means for all years for spring.

For Fig. 39 it is noticeable that for the difference with $F_{S_{ERA}}$ the highest positive differences occur along the land sea borders. Continental the differences are similar to those in spring, with a difference in northern America and central Asia. There are lots of red patches within the content. It is likely that in those areas probably snow melts, which could not be resolved with equation 24, because of reasons given in section 5.4.
The difference to $F_{S_{INDIR}}$ is again very noisy, but shows a similar picture for northern America.

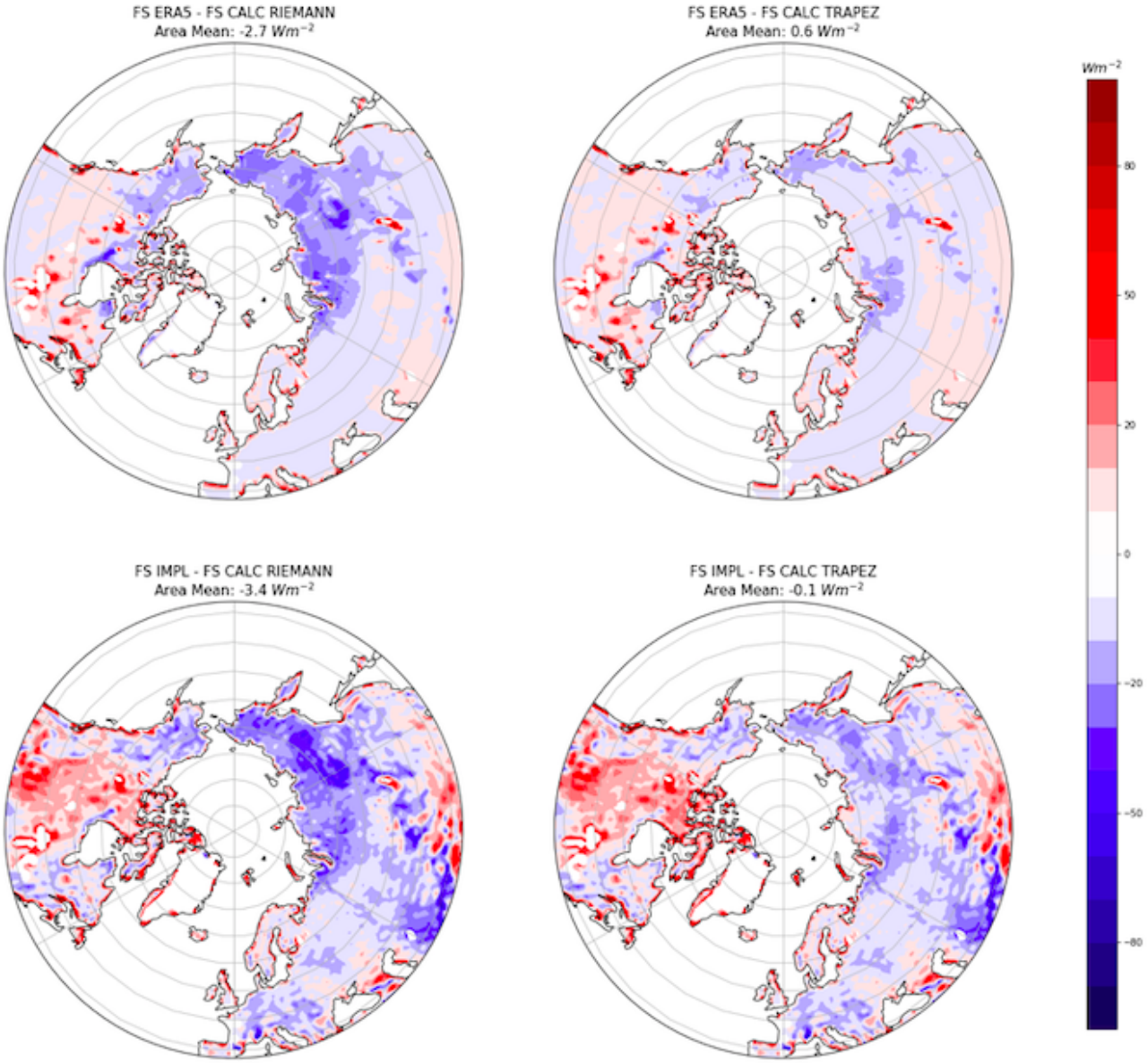


Figure 39: Differences between calculation methods as seasonal means for all years for summer.

Fig. 40 shows the differences in autumn, where, at least at a quick glance, it looks like the differences to $F_{S_{ERA}}$ and $F_{S_{INDIR}}$ seem to be a similar range of magnitude. The presented differences in northern America show a similar pattern, but are more precise in $F_{S_{INDIR}}$. In Norway and the Rocky Mountains, the differences to $F_{S_{ERA}}$ and $F_{S_{INDIR}}$ show equally large negative values at the edges.

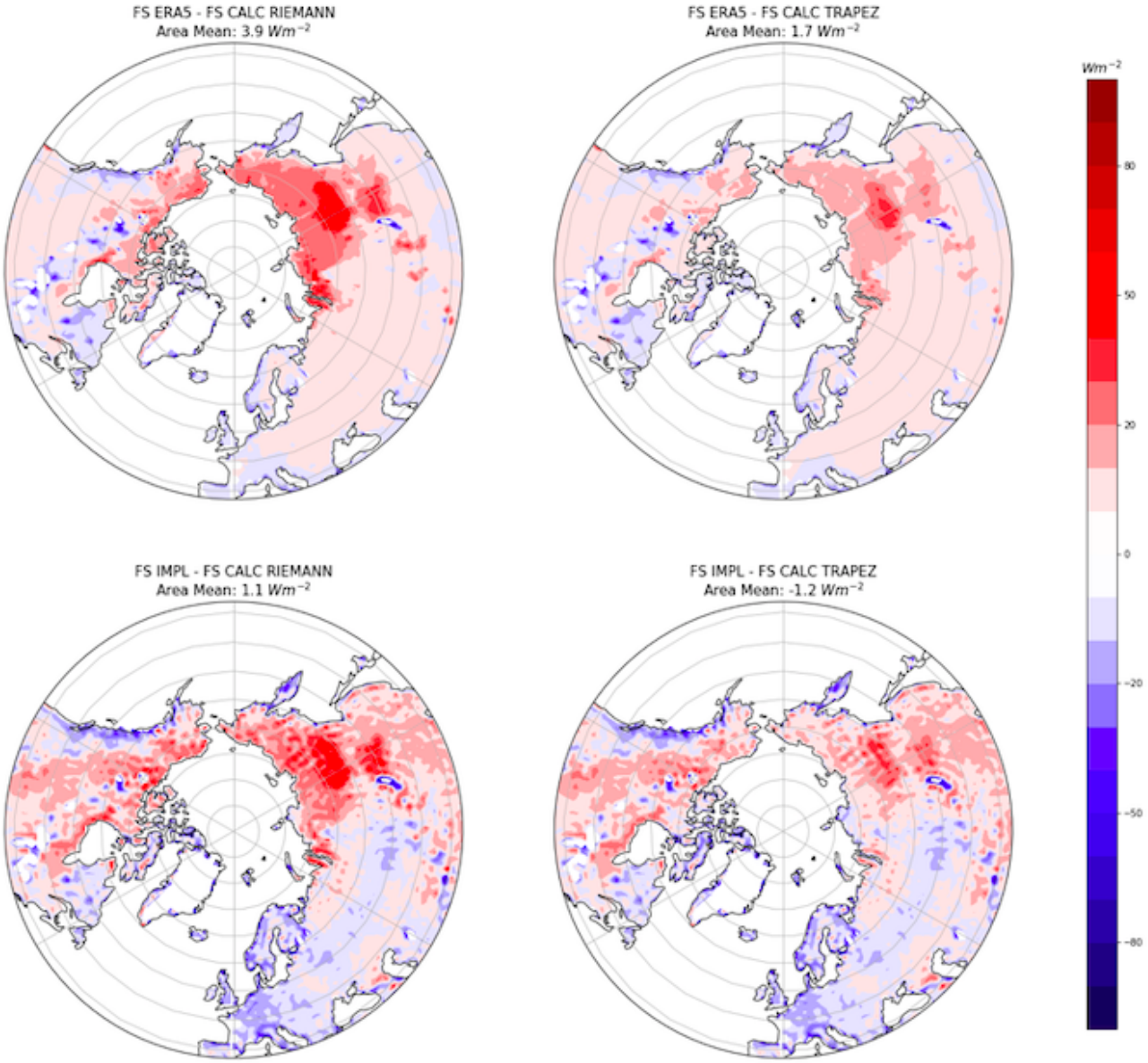


Figure 40: Differences between calculation methods as seasonal means for all years for autumn.

The difference to $F_{S_{INDIR}}$ shows in winter the most noisy pattern. The basic color distribution is similar to the difference to $F_{S_{ERA}}$. In Norway and the Rocky Mountains a comparable picture to Fig. 40 is displayed.

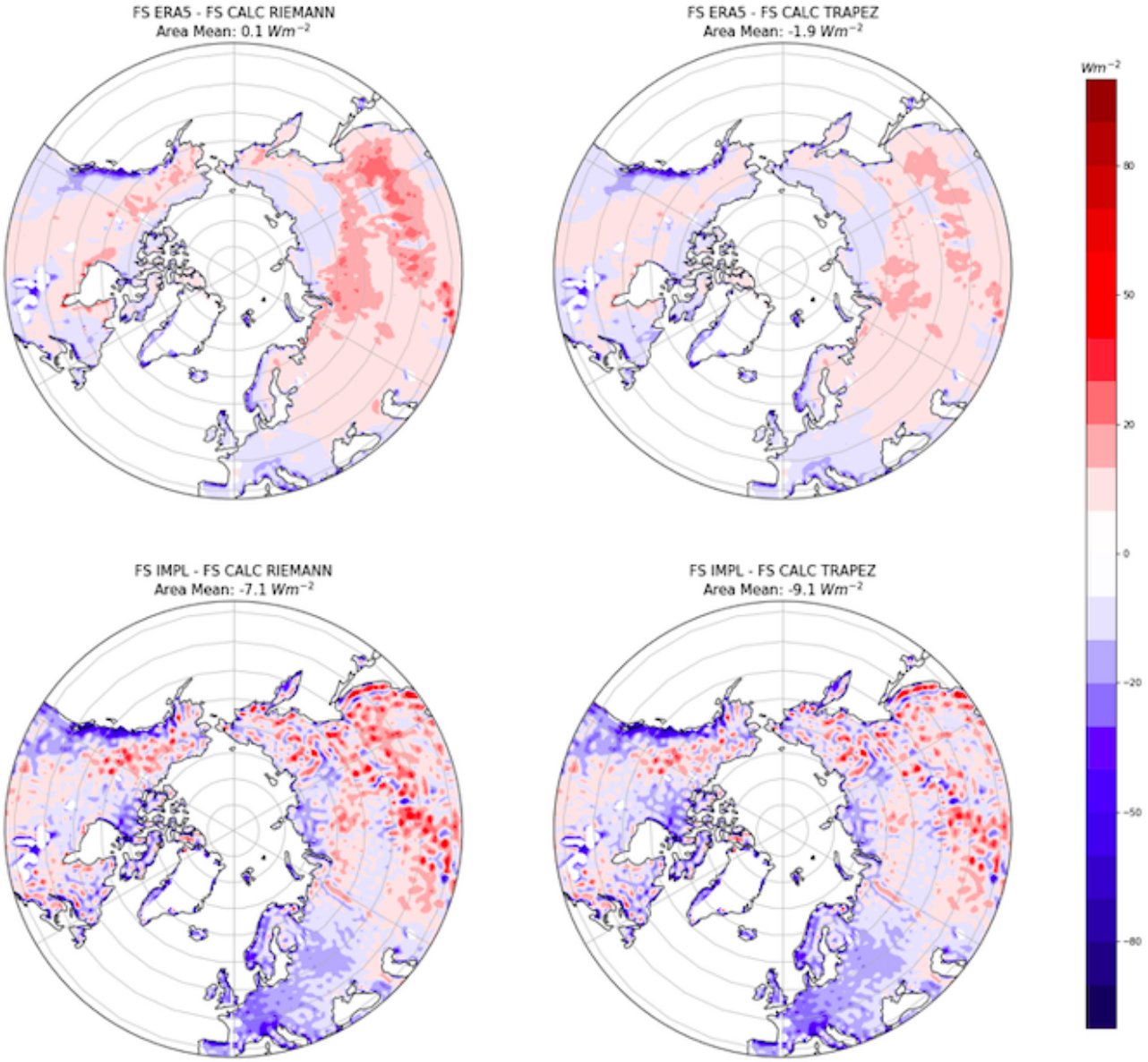


Figure 41: Differences between calculation methods as seasonal means for all years for winter.

5.3 Trends in soil temperature and soil moisture

Remembering the introduction, an imbalance of 0,6 Wm^{-2} was mentioned. Therefore the anomalies of soil temperature and soil moisture for the land surface from 40 °N to 90 °N in all soil layers are considered. The red line represents a linear trend line.

Fig. 42 shows the temperature anomalies for all soil layers from 2000 to 2018. The slope of the red line is not significant at a first sight. Looking at the slopes of the trend lines, a small trend can be seen. It is interesting to note that this is greatest in the deepest layer, but with 0,001 K/19a very small.

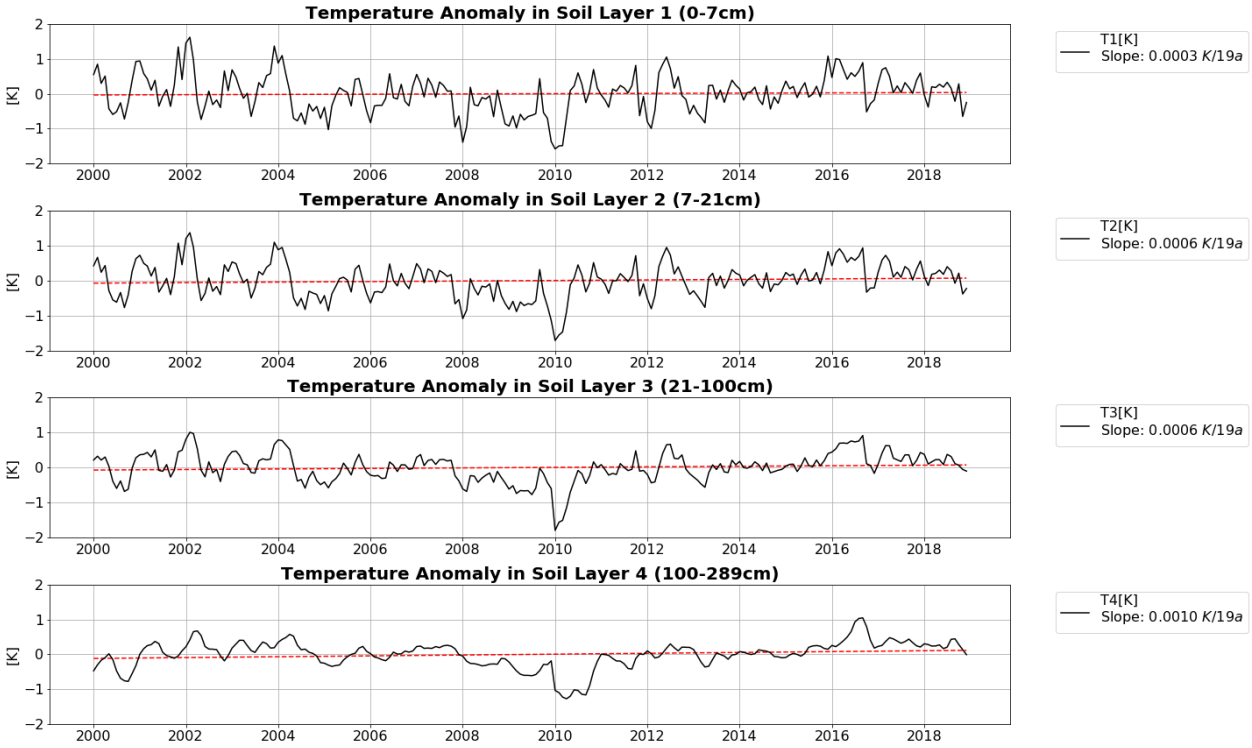


Figure 42: Soil temperature anomaly and trend line for all soil layers from 2000 to 2018.

In Fig. 43 the moisture anomalies for all soil layers in the whole period under consideration are displayed. The value of the slope of the trend line does not show any significant values (rounded to four decimal places it is still 0), but optically a trend can be recognized. The red line has a negative slope, which indicates a decrease in soil moisture in all layers. This is also consistent to a tendency towards drier soils. An interesting feature is the anomaly of soil moisture in layer 4. In 2010 it seems that there is a jump in the data, so in 2015, which could be an indication of inhomogeneities in the time series. This also occurs in the lowest and thickest layer, which may have an effect to the calculation of $LSHCT$.

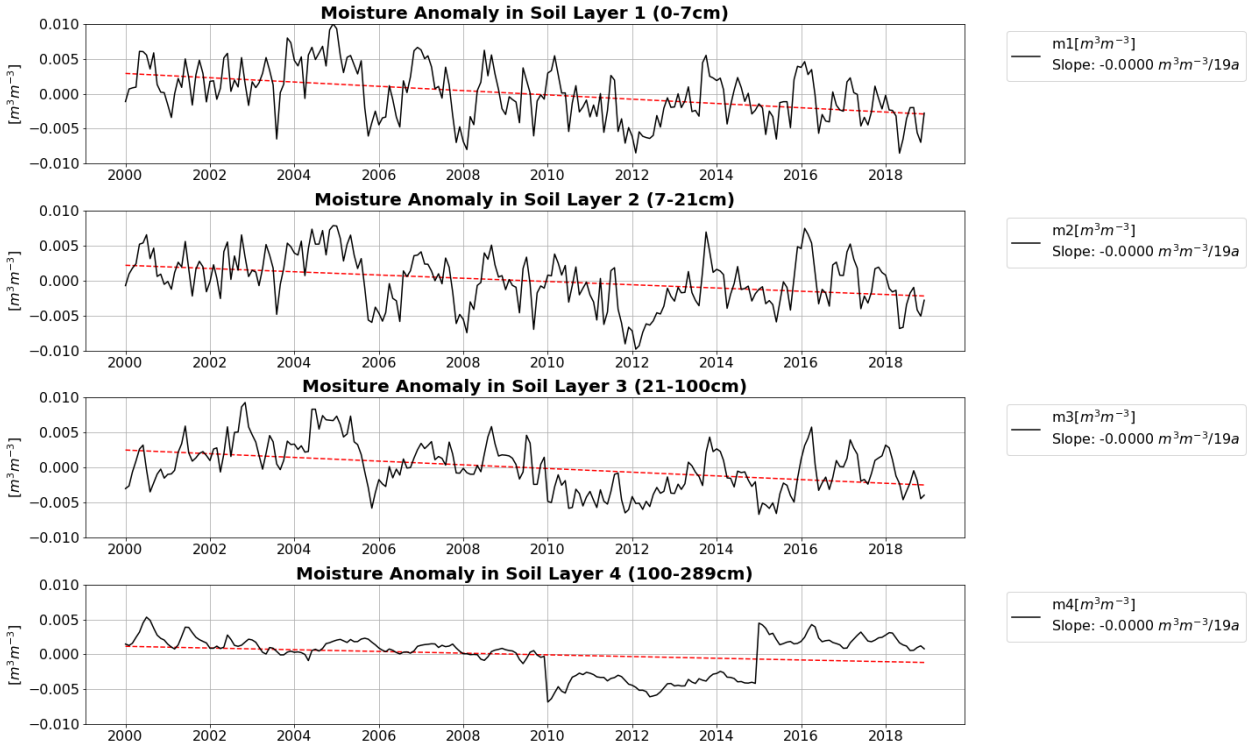


Figure 43: Soil moisture anomaly and trend line for all soil layers from 2000 to 2018.

In Fig. 44 the climatology of soil temperature in all four layers is shown. As expected it can be seen that the amplitude of the first layer which extends close to the ground from 0 to 7 cm has the greatest amplitude. The deeper you go into the soil, the weaker the annual cycle becomes. The maximum for the first layer appears in July, for the fourth layer it occurs in September. Deeper in the ground, it also takes longer before cooling occurs again.

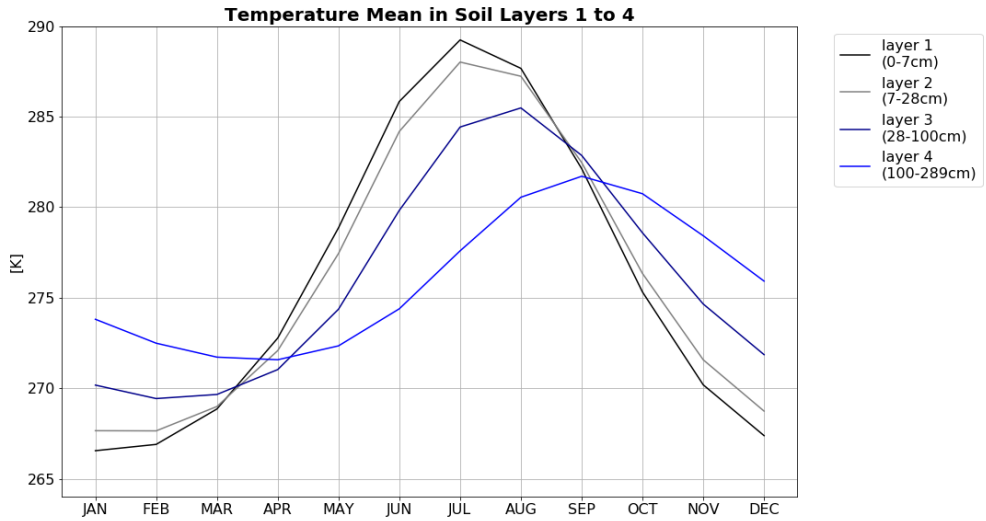


Figure 44: Soil temperature climatology for all soil layers from 2000 to 2018.

5.4 Snow Melt

If $TSHCT$ and $LSHCT$ are 0 Wm^{-2} (fulfilled as explained in section 5.1) on average, $F_{S_{ERA}}$ should be balanced by snow melt, which means the part ST - SF in equation 24. Actually, CSF and RF should also be part of the snow melt, but due to the fact that they are so small compared to the other two components they are neglected.

Fig. 45 shows the annual mean values of $F_{S_{ERA}}$ in orange and ST - SF in blue. The unit is Wm^{-2}, the values range between 1,5 Wm^{-2} and 2,5 Wm^{-2}. In an area average made for the land surface from 40 °N to 90 °N, the two terms are not in balance but very close. The agreement is better at the beginning of the observation period and then gradually becomes worse.

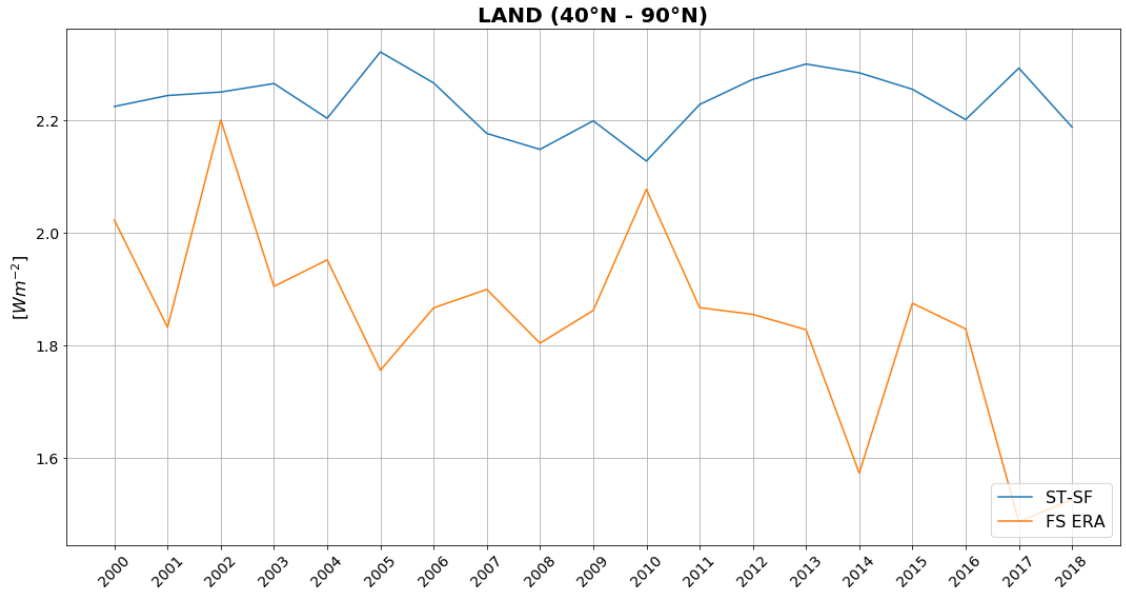


Figure 45: Annual mean values from 2000 to 2018 of $F_{S_{ERA}}$ (orange) and ST -SF (blue). The averaging area is land surface from 40 °N *to* 90 °N.

Taking a look at the difference between the terms in the area considered in Fig. 45 is quite difficult, as large regional differences can occur and these disappear on average. Therefore two cutouts are used to show the regional variations.

The first cutout is northern America (Fig. 46), an area where as discussed above, the result of different calculating methods is satisfying. It should be noted here that the scale differs from that of Fig. 45, and ranges from 1,2 Wm^{-2} to 2,7 Wm^{-2}. The spread between $F_{S_{ERA}}$ and ST - SF is still there, but the differences are in about the same value range as in figure 45.

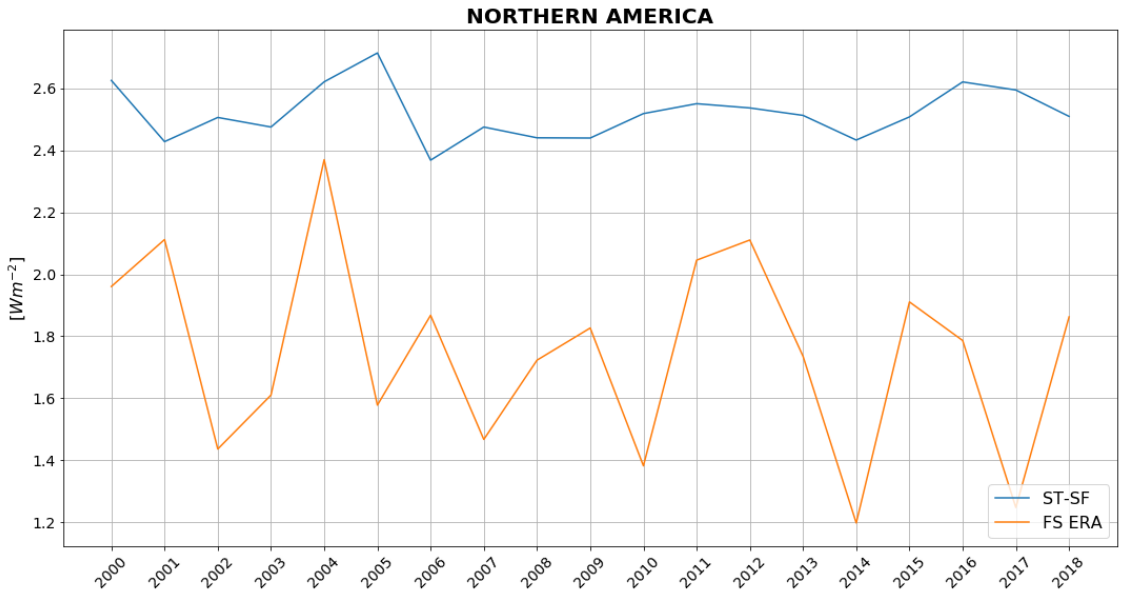


Figure 46: Annual mean values from 2000 to 2018 of FS_{ERA} (orange) and ST-SF (blue). The averaging area is northern America.

The second cutout presents Greenland (Fig. 47), an area where the results gained are unsatisfactory. It should be noted here that the scale differs from that of figures Fig. 45 and Fig. 46, and ranges from $-4{,}0\,\mathrm{Wm}^{-2}$ to $6\,\mathrm{Wm}^{-2}$. It is very clear, that in this area the spread is huge with about $10{,}0\,\mathrm{Wm}^{-2}$ for the whole period. This may indicate that the area under consideration is too small after masking and so equation 24 does not work.
However, the verification of this fact is a major difficulty as observation data for ST and SF are missing.

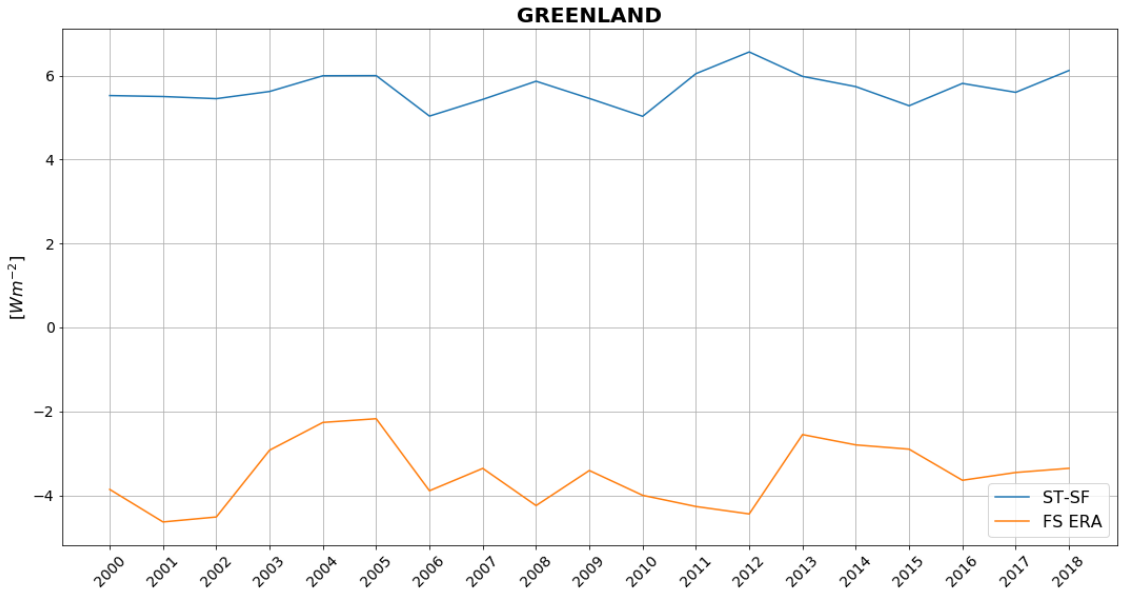


Figure 47: Annual mean values from 2000 to 2018 of FS_{ERA} (orange) and ST-SF (blue). The averaging area is Greenland.

5.5 Calculation of tendencies

Another point that should be considered is the impact of the tendency calculation method on the result. So far only analysed state quantities have been used, as discussed in section 4.4.1. A further option for tendency calculation is the calculation from forecasts, as noted in section 4.4.3. Another opportunity is to calculate the tendency from instantaneous fields as given in section 4.4.2.

Fig. 48 shows basically the same curves as Fig. 29. The brown line represents the result for equation 24 with the tendencies calculated via direct differences from instantaneous fields (eq. 41). Additionally the curve for the tendencies calculated with forecast data (eq. 42) is given in light purple.

It is recognizable that the purple line provides much higher values than the other methods from March to June, which will be discussed below. F_S from instantaneous fields fits to $F_{S_{INDIR}}$ very well from January to March, but shows significant deviations in the rest of the year. $F_{S_{CALC,Riemann}}$, $F_{S_{CALC,Trapez}}$ and the brown line show a good agreement from mid-July to early September.

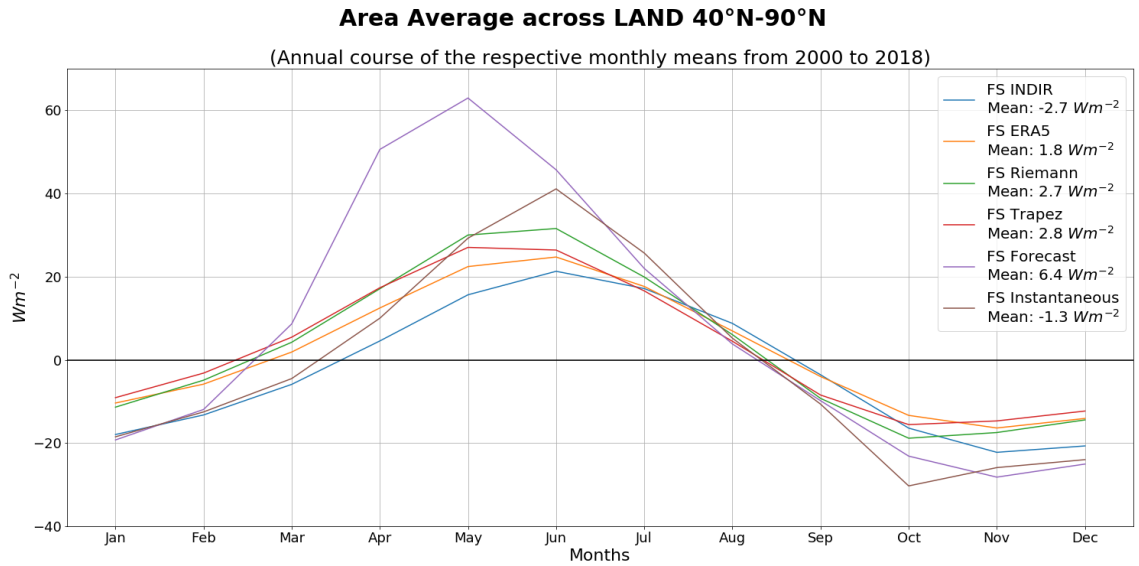


Figure 48: Comparison of annual cycle for different methods for tendency calculation.

Considering the average of the time series, the annual average should be $0\,\mathrm{Wm}^{-2}$ for a closed system. The average values given in Fig. 48 show that none of the methods provides a mean value of $0\,\mathrm{Wm}^{-2}$. It is noticeable that F_S from forecast data delivers the highest mean value of about $6{,}4\,\mathrm{Wm}^{-2}$, which is not surprising considering the large deviation from March to June. F_S from instantaneous fields returns with $-1{,}3\,\mathrm{Wm}^{-2}$ the value nearest to $0\,\mathrm{Wm}^2$, but is like $F_{S_{INDIR}}$ negative which rather speaks for a cooling off.

Fig. 48 shows that different methods for the calculation of tendencies have a major impact on the result. It was assumed that the tendencies provided by forecasts should deliver a better fit to $F_{S_{ERA}}$. If there is an error in the forecasts, it is going to influence the analysis and reanalysis.

Figures 49 to 51 show the terms affected by the method of tendency calculation for an area average of the land surface from 40 °N to 90 °N. The black bars show the tendency calculated from monthly means as described in section 4.4.1, the dark grey bar the tendency calculated from instantaneous fields as displayed in section 4.4.2 and the light grey bar the tendency calculated from forecast data as presented section 4.4.3.

Considering Fig. 49 the black and the dark grey bar show similar values. When averaging the time series, the average should be 0 Wm^{-2}. The tendency calculated from monthly means (black) with $-0{,}38\,Wm^{-2}$ comes closest to 0 Wm^{-2}. Slightly worse is the mean for the tendency calculated from instantaneous fields with $-0{,}47\,Wm^{-2}$. Regarding the tendency of thermal heat content from forecast data, those values differ from the other methods. From January to August the light grey bars deliver lager values as the other methods, which may indicate that the soil is warming up more in forecasts than in analysis. A time series average of $1{,}04\,Wm^{-2}$ also suggests that the soil is warming.

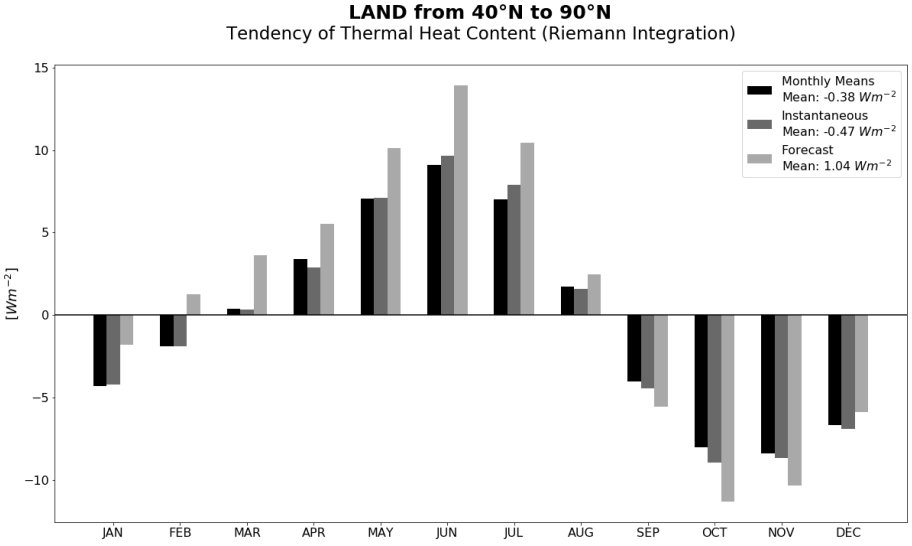


Figure 49: Comparison of $TSHCT$ for different methods of tendency calculation. The integration method used is the Riemann integration.

In Fig. 50 the black and the grey bar behave very similar again. With both methods $LSHCT$ disappears approximately in the temporal average ($0{,}03\,Wm^{-2}$ (black bar) and $0{,}08\,Wm^{-2}$ (dark grey bar)), which is a very good result. The calculation of $LSHCT$ from the forecast data is out of line. The values in March, April and May differ considerably. In March the values are around 10 Wm^{-2}, in April around 30 Wm^{-2} and in May around about 20 Wm^{-2} larger than those of the other methods. This also affects the mean of the time series, which is now at $5{,}69\,Wm^{-2}$, which means that a closed system can no longer be assumed. This explains why the purple line in Fig. 48 from March to May differs significantly from the others. This will be interpreted as more soil ice melting in the forecasts than in the analysis. This assumption is also consistent with the results as discussed for Fig. 49 above.

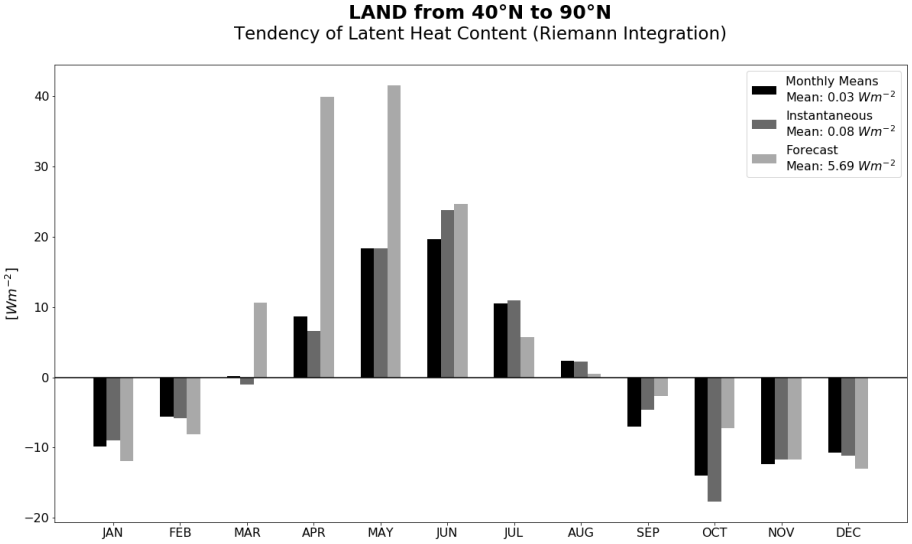


Figure 50: Comparison of $LSHCT$ for different methods of tendency calculation. The integration method used is the Riemann integration.

Regarding Fig. 51, the black and the dark grey bar again show a similar course. The mean for the tendencies from monthly means is 0,1 Wm^{-2}, which is a good result very close to zero. Even better is the result for the tendencies from the instantaneous fields with 0,09 Wm^{-2}. The tendency calculated from forecast data shows greater deviations from the other calculation methods. The values are about 1 to 2 Wm^{-2} less per year than with the other methods. The average of the time series shows with −0,63 Wm^{-2} larger deviations from the desired result.

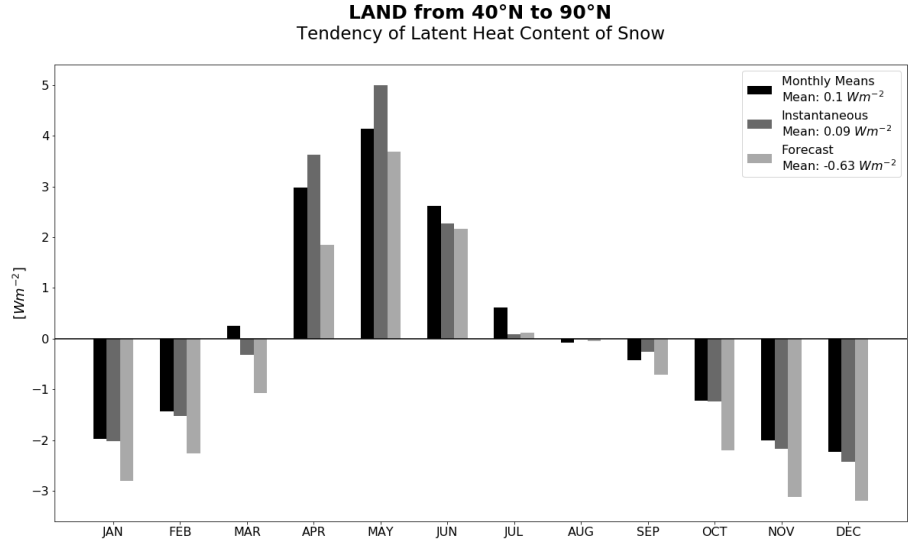


Figure 51: Comparison of ST for different methods of tendency calculation.

5.6 Verification

Visual comparison between the curves is possible, but some statistical measures should be used for qualitative description. The correlation coefficients according to Pearson and Spearman are used in the following, as well as the root mean square error ($RMSE$) and the mean absolute error (MAE).

The Pearson correlation coefficient explains the variance of one variable by another variable. It has the property that it is not robust and not resistant to outliers. If the Pearson correlation coefficient is 1, there is perfect linear correlation, if it is -1, the slope of the regression line is negative. According to Wilks (1995), the Pearson correlation coefficient can be expressed as given in equation 48.

$$CORR - Pearson = \frac{cov(x,y)}{s_x s_y} \tag{48}$$

In equation 48, $cov(x,y)$ denotes the covariance of the considered variables x and y, s_x is the standard deviation of x, s_y the one of y.

For the Spearman correlation coefficient the data is sorted and the rank of the one variable is correlated with the rank of the other variable. The Spearman correlation coefficient is robust to outliers. (Wilks, 1995)

The $RMSE$ is calculated by equation 49, the MAE is given by equation 50, taken from Dorninger (2019).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2} \tag{49}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|F_i - O_i| \tag{50}$$

N is the sample, F_i is the forecast data, where $F_{S_{ERA}}$ is used therefore in the following. The *observation* data O_i is presented by $F_{S_{CALC,Riemann}}$ in the following.
According to Dorninger (2019), $RMSE$ and MAE should always be used together. If both are about the same size so all errors have the same magnitude, if $RMSE$ is way bigger than MAE a high variance of errors is expected.
If the data points were all located on the red line, perfect correlation between $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$ would exist. In addition, the distribution of the values of $F_{S_{ERA}}$ can be seen as a histogram on the right, a histogram for $F_{S_{CALC,Riemann}}$ on the top.

The following figures (Fig. 52, Fig. 53, Fig. 54) present the scatter diagram for an area average of land surface from 40 °N to 90 °N, northern America and Greenland.

Fig. 52 shows under consideration of the correlation coefficients a good agreement of $F_{S_{ERA}}$ and $F_{S_{CALC,Riemann}}$. As discussed above a value of 1 provides perfect correlation. The difference between $RMSE$ and MAE is with 3,41 Wm^{-2} still big, which indicates a large variance of errors.

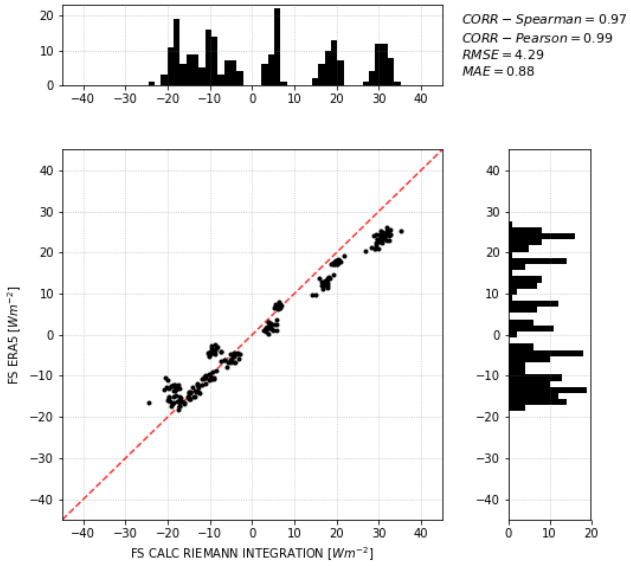


Figure 52: Scatter diagram for an area average of land surface from 40 °N *to* 90 °N.

For northern America (Fig. 53) the data points fit better to the line. The correlation coefficients are slightly increased to 0,98 and 0,99. The difference between $RMSE$ and MAE is also better with 1,55 Wm^{-2}.

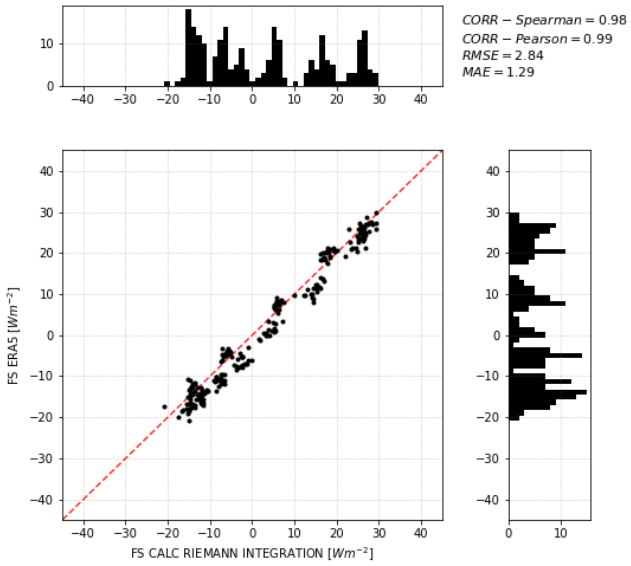


Figure 53: Scatter diagram for an area average of northern America.

For Greenland it can be seen that the data points scatter further around the red line, the correlation coefficients level off at 0,87 and 0,91. If one looks at a regression line through the data points, there is an underestimation, which means that the values for FS_{ERA} are too small.

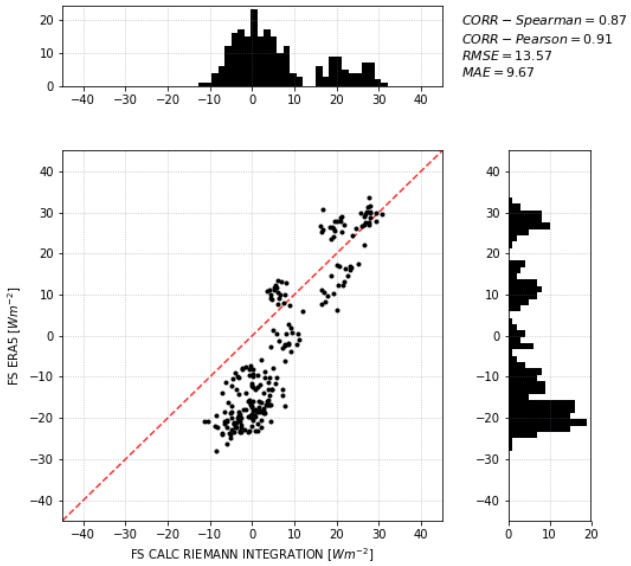


Figure 54: Scatter diagram for an area average of Greenland.

6. Conclusion

In the context of this work an energy budget equation for land surface was developed to get an estimate of the net energy flux on surface.

The comparison of the calculated net energy flux to the net energy flux calculated by other methods showed that the result is generally satisfying, but strongly dependent on the selected area of interest. So for northern America very good results were achieved, while the results for Greenland were very unsatisfying.
The masking of areas with no change in SWE leads to a result that showed how sensitive the land energy budget equation is to the term ST. Using the mask provided an output which was better suited to the reference curves in some areas like northern America. Other areas, for example Greenland, suffered because only a small edge of Greenland remained for the evaluation. It would therefore be desirable to obtain better information on the inland ice in order to be able to consider the land energy budget over these surfaces.

Some data is observation-near, such as the TOA net radiation or the total energy flux divergence. Some surface fields are analysed as well such as snow depth or soil temperature. Forecast fluxes, such as latent and sensible heat fluxes, rainfall and snowfall have highest uncertainty.

Another problem is that different methods for calculating tendencies lead to different results. If the values of the tendencies calculated from the forecasts differ much from the other tendency calculation methods, it is likely that there is an error in the forecast model. An example of this is $LSHCT$ as discussed in section 4.4. An error in the forecast model leads to an error in the analyses, respectively reanalyses.
In order to better evaluate this, observational data would be an advantage. However, this is a major obstacle in the spatial resolution used. What would be helpful and desirable, would be observational data over a small area where, for example, the ground ice content is measured.

However $LSHCT$ represents one of the greatest uncertainties, and has the greatest influence to the energy budget of the land surface. Due to the fact that there is no data available for the soil ice content, only assumptions can be made, like that it is ice if the temperature is less than 273,15 K. Which brings us to the next problem, that at 273,15 K the melting process starts. So is at this temperature the soil still frozen or is only water available? In this thesis the water was assumed to be frozen if the temperature is definitely lower than 273,15 K.

Another issue is that the soil temperature and moisture are only available up to 289 cm soil depth. To obtain more accurate results, the soil values would have to be known down to a greater depth.

Regarding an imbalance of 0,6 Wm^{-2}, it would have been expected that the trends in soil temperature and moisture are more significant. Nevertheless small signs of an increase in soil temperature and a tendency towards drier soils are visible.

Bibliography

Abu-Hamdeh, N. H. (2003), Thermal properties of soils as affected by density and water content. *Biosystems Engineering* **86**, 97–102, DOI:10.1016/S1537-5110(03)00112-0.

AMS (2012), Glossary of meteorology. http://glossary.ametsoc.org/wiki/Wet-bulb_temperature, Accessed on 13.04.2020 12:02.

Balsamo, G., P. Viterbo, A. Beljaars, B. Van den Hurk, M. Hirschi, A. Betts, and K. Scipal (2008), A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *ECMWF Technical Memoranda* **563**, DOI:10.21957/yzyeh0vlw.

Beltrami, H., J. E. Smerdon, H. N. Pollack, and S. Huang (2002), Continental heat gain in the global climate system. *Geophysical Research Letters* **29**, 8.1–8.3, DOI:10.1029/2001GL014310.

Brand, C. (2017), Finite differenzen methode. http://institute.unileoben.ac.at/amat/lehrbetrieb/num2/NUMMII1_17.pdf, Accessed on 23.02.2020 16:26.

Brasseur, G., W. Cramer, M. Ehrendofer, L. Emmons, M. Erhard, D. Gerten, V. Gouretski, C. Granier, U. Haberlandt, L. Haimberger, M. Hantel, A. Hense, P. Huybrechts, R. Jaenicke, K. Koltermann, M. Kottek, J. Meincke, H. Miller, A. Ohmura, E. Raschke, F. Rubel, B. Rudolf, C.-D. Schönwiese, and C. Stubenrauch (2005), *Numerical data and functional relationships in science and technology: new series. Group 5 Vol. 6: Geophysics Observed global climate.*, chap. 10, pp. 10.1–10.30. Springer, Berlin, ISBN:978-3-540-20206-6.

C3S (2020). https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means?tab=form, Accessed on 06.04.2020 12:02.

Church, J. A., N. J. White, L. F. Konikow, C. M. Domingues, J. G. Cogley, E. Rignot, J. M. Gregory, M. R. van den Broeke, A. J. Monaghan, and I. Velicogna (2011), Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophysical Research Letters* **38**, DOI:10.1029/2011GL048794.

DeLuca, C. (2018), Esmpy : Esmf python regridding interface. https://www.earthsystemcog.org/projects/esmpy/, Accessed on 11.04.2020 08:44.

Dorninger, M. (2019). Lecture notes on verification of winter term 2019.

DWD (2020a). https://www.dwd.de/DE/service/lexikon/Functions/glossar.html?lv2=102248&lv3=102438, Accessed on 03.04.2020 16:40.

DWD (2020b). https://www.dwd.de/DE/service/lexikon/begriffe/W/Wasseraequivalent_pdf.pdf?__blob=publicationFile&v=4, Accessed on 03.04.2020 16:50.

DWD (2020c). https://www.dwd.de/DE/forschung/klima_umwelt/klimaueberwachung/reanalysen/reanalysen_node.html, Accessed on 24.03.2020 16:51.

ECMWF (2016). https://www.ecmwf.int/en/elibrary/16648-part-iv-physical-processes, Accessed on 11.12.2019 12:06.

ECMWF (2020a). https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation, Accessed on 06.04.2020 11:58.

ECMWF (2020b). https://apps.ecmwf.int/codes/grib/param-db/, Accessed on 10.04.2020 11:38.

Haimberger, L. (2019). Oral conversation on 09.07.2019.

Haimberger, L. (2020). Oral conversation on 18.05.2020.

Hanke-Bourgeois, M. (2009), *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*, vol. 3. Vieweg+Teubner, p.317, ISBN:978-3-8348-0708-3.

Hantel, M. (2013), *Einführung Theoretische Meteorologie*, vol. 1. Springer Spektrum, ISBN:978-3-8274-3055-7, DOI:10.1007/978-3-8274-3056-4.

Hantel, M. and L. Haimberger (2016), *Grundkurs Klima*, vol. 1. Springer Spektrum, ISBN:978-3-662-48192-9, DOI:10.1007/978-3-662-48193-6.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut (2020), The ERA5 Global Reanalysis. *Quarterly Journal of Royal Meteorological Society* **146**, DOI:10.1002/qj.3803.

Holopainen, E. and C. Fortelius (1986), Accuracy of estimates of atmospheric large-scale energy flux divergence. *Monthly Weather Review* **114**, 1910–1921, DOI:10.1175/1520-0493(1986)114<1910:AOEOAL>2.0.CO;2.

IPCC (2013). http://www.climatechange2013.org/images/report/WG1AR5_ALL_FINAL.pdf, Accessed on 13.04.2020 18:37.

Kraus, H. (2008), *Grundlagen der Grenzschicht-Meteorologie: Einführung in die Physik der atmosphärischen Grenzschicht und in die Mikrometeorologie*. Springer, Berlin Heidelberg, ISBN:978-3-540-75980-5.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong (2009), Toward optimal closure of the Earth's Top-of-Atmosphere radiation budget. *Journal of Climate* **22**, 748–766, DOI:10.1175/2008JCLI2637.1.

Mayer, M. and L. Haimberger (2012), Poleward atmospheric energy transports and their variability as evaluated from ecmwf reanalyis data. *Journal of Climate* pp. 734–752, DOI:10.1175/JCLI-D-11-00202.1.

Mayer, M., L. Haimberger, J. M. Edwards, and P. Hyder (2017), Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. *Journal of Climate* **31**, 9225–9246, DOI:10.1175/JCLI-D-17-0137.1.

Mayer, M., L.Haimberger, M.Pietschnig, and A. Storto (2016), Facets of arctic energy accumulation based on observations and reanalyses 2000–2015. *Geophysical Research Letters* **43**, 10,420–10,429, DOI:10.1002/2016GL070557.

Mayer, M., S. Tietsche, L. Haimberger, T. Tsubouchi, J. Mayer, and H. Zuo (2019), An improved estimate of the coupled arctic energy budget. *Journal of Climate* **32**, 7915–7934, DOI:10.1175/JCLI-D-19-0233.1.

Shea, D. (2020). https://climatedataguide.ucar.edu/climate-data-tools-and-analysis/regridding-overview, Accessed on 27.03.2020 18:13.

Strampp, W. (2012), *Höhere Mathematik 2*. Springer, Wiesbaden, p.136, ISBN:978-3-658-09008-1, DOI:10.007/978-3-8348-2541-4.

Stull, R. (2011), Wet-bulb temperature from relative humidity and air temperature. *Journal of Applied Meteorology and Climatology* **50**, 2267–2269, DOI:10.1175/JAMC-D-11-0143.1.

Stull, R. B. (1988), *An introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers, Dordrecht,Boston,London, ISBN:978-90-277-2769-5, DOI:10.1007/978-94-009-3027-8.

von Schuckmann, K., M. D. Palmer, K. E. Trenberth, A. Cazenave, D. Chambers, N. Champollion, J. Hansen, S. A. Josey, N. Loeb, P.-P. Mathieu, B. Meyssignac, and M. Wild (2016), An imperative to monitor Earth's energy imbalance. *Nature Climate Change* **6**, 138–144, DOI:10.1038/nclimate2876.

Wallace, J. M. and P. V. Hobbs (2006), *Atmospheric Science: An Introductory Survey*, vol. 2. Academic Press, ISBN:978-0-12-732951-2.

Warrilow, D., A. Sangster, and A. Slingo (1986), Modelling of land-surface processes and their influence on European climate. *UK Met Office Report* .

Wilks, D. S. (1995), *Statistical methods in the atmospheric sciences*. Academic Press, San Diego, p.45ff, ISBN:0-12-751965-3.

Zhuang, J. (2017a), Comparison of 5 regridding algorithms. https://xesmf.readthedocs.io/en/latest/notebooks/Compare_algorithms.html, Accessed on 11.04.2020 17:39.

Zhuang, J. (2017b), xesmf: Universal regridder for geospatial data. https://xesmf.readthedocs.io/en/latest/, Accessed on 03.02.2020 12:09.

Abbreviations and Symbols

AET	Atmospheric total energy tendency
BOT	Bottom of the ocean/land column
c_i	Specific heat capacity of ice
$\vec{c_i}$	Sea ice drift vector
c_p	Specific heat capacity of air at constant pressure
c_w	Specific heat capacity of liquid water
CERES	Clouds and the Earth's Radiant Energy System (Satellite Data)
CSF	Cold snow fall on earth's surface
C3S	Copernicus Climate Change Service
t_i	Thickness of sea ice layer
d_s	Snow thickness
dh	Differential of enthalpy
dT	Differential of temperature
E	Evaporation
ECMWF	European Center for Medium-Range Weather Forecasts
ERA5	Fifth generation of ECMWF atmospheric reanalyses
f_i	Sea ice fraction
f_l	Land fraction
f_o	Ocean fraction
F_A	Atmospheric lateral energy transport
F_B	Basal heat flux
F_I	Latent heat transport associated with sea ice
F_O	Ocean heat transport
F_S	Vertical net energy flux
$F_{S_{CALC,Riemann}}$	Calculated F_S with equation 24 by use of Riemann integration
$F_{S_{CALC,Trapez}}$	Calculated F_S with equation 24 by use of trapezoidal integration
$F_{S_{ERA}}$	F_S parameterized from ERA5
$F_{S_{INDIR}}$	Indirect calculated net energy flux by the use of the atmospheric column

g	Gravitational acceleration
G	Ground heat flux
h	Enthalpy
h_0	Specific enthalpy of water at 0 °C
HTESSEL	Improvement of TESSEL scheme
$IHCT$	Sea ice sensible heat content tendency
IFS	Integrated Forecast System
k	Atmospheric kinetic energy
L_f	Latent heat of fusion
L_v	Latent heat of vaporization
LH	Latent heat flux
$LSHCT$	Tendency of latent soil heat content
m	Moist static energy
MAE	Mean absolute error
MET	Sea ice melt tendency
M_{snow}	Energy required for snow melt
N	Immaterial sourve term
NWP	Numerical Weather Prediction Model
$OHCT$	Temporal tendency of ocean heat content
p	Atmospheric pressure
P	Total precipiation
P_{snow}	Snowfall rate
Φ	Geopotential
q_g	Specific water vapour content
q_l	Specific liquid water content
q_s	Specific ice content
ρ_i	Density of ice
ρ_s	Density of soil
ρ_{snow}	Density of snow
ρ_w	Density of water

Rad_{SFC}	Net radiation on surface
$Rad_{SFC_{solar}}$	Net solar radiation on surface
$Rad_{SFC_{thermal}}$	Net thermal radiation on surface
Rad_{TOA}	Net radiation at TOA level
$RMSE$	Root Mean squared error
RF	Rain fall on earth's surface
ΔS	Storage term
SFC	Surface (of sea, soil column,..)
SF	Snow fall on earth's surface
SH	Sensible heat flux
ST	Tendency of latent heat content of snow on earth's surface
SWE	Snow water equivalent
T_{00}	Reference temperature
T_a	Air temperature
T_i	Temperature of sea ice
T_o	Temperature of ocean
T_p	Temperature of precipitation
T_s	Temperature of soil
TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land
TOA	Top of the atmosphere
$TSHCT$	Tendency of thermal soil heat content
Θ_i	Volumetric soil ice content
Θ_w	Volumetric soil water content
4DVAR	Data Assimaliation Method
$\vec{v}$	Horizontal wind vector