Estimating area of vector polygons on spherical and ellipsoidal earth models with application in estimating regional carbon flows

Huiting Huang

Huiting Huang (2017).

Estimating Area of Vector Polygons on Spherical and Ellipsoidal Earth Models with Application in Estimating Regional Carbon Flows

Master degree thesis, 30 credits in Physical Geography and Ecosystem Analysis

Department of Physical Geography and Ecosystem Science, Lund University

Thesis nr 436

Level: Master of Science (MSc)

Course duration: January 2017 until June 2017

Estimating Area of Vector Polygons on Spherical and Ellipsoidal Earth

Models with Application in Estimating Regional Carbon Flows

Author

HUITING HUANG

Master thesis, 30 credits, in Physical Geography and Ecosystem Analysis

Supervisor Lars Harrie

Examiner Harry Lankreijer

Exam Committee Paul Miller

Department of physical Geography and Ecosystem Science

Lund University

Abstract

Estimating area of polygons on the Earth's surface is required in many fields in earth science. In the field of carbon modelling, one application of estimating polygons' area is to estimate carbon flows for regions. This thesis aims to develop a methodology to estimate area of a polygon on a spherical/ellipsoidal surface applied to the problem to estimate carbon flows in regions.

It is common that field data are stored in grid which covers the Earth's surface in earth science. Region area estimation is inevitable for computation of sum of field data or density of data in regions. Region area can be computed by summing up the whole or partial area of grid cells covered by the region. The Earth is usually modelled as a sphere or an ellipsoid. Area of the overlay polygon on spherical/ellipsoidal surface can be considered as the product of cell area and fraction (partial value) of overlay area in the grid cell. Three methodologies to estimate partial value of overlay area in a grid cell were proposed and tested: 1) using latitude-longitude plane, 2) using cylindrical area-preserving projection and 3) using the area of corresponding of spherical polygons. Cell sizes were estimated by cylindrical equal-area projection method. Tests show that area-preserving projection method is a suitable method to estimate area of a polygon on the Earth's surface for the application of regional carbon flow estimation because it trades off the quality of estimates and computational demands.

Estimation of carbon flows in regions is interesting in many research domains. Atmospheric inversion is one technique of carbon flux modelling to provide carbon flux data in grid with various resolutions. Regional carbon flows can be estimated as the sum of fluxes in grid cells overlapped by the polygonal region. In most models, flux is modelled constant everywhere in each grid cell. A case study was performed to estimate carbon flow in Sweden using the methodology developed to estimate area of polygon. The uncertainties in the estimation of carbon flow in Sweden are influenced by the estimation of geographic extent of Sweden and the flux data in grid provided by atmospheric inversions. Four groups of test were done to test the effects of different factors on the flow estimation: partial values, earth model, interpolation and inversion systems. The test result illustrates partial value, earth model and interpolation have less than 1% effect on final result. The region flow is mainly influenced by flux data modeled by different inversions.

Keywords: area of a polygon, atmospheric inversions, map projection, regional carbon flows

Acknowledgements

Foremost, I would like to express my sincere appreciation and gratitude to my supervisor Lars Harrie for his continuous support, patience and guidance. Without his help, I could have not accomplished my thesis. The door of his office was open every time when I ran into troubles or had questions about research and writing. He always replied to my emails quickly and explained clearly no matter what questions I had. Thanks to his understanding and concerns when I was sick and help to make a new plan for me when I was a little behind the schedule.

I would also like to acknowledge Ute Karstens who helped us to figure out the questions about atmospheric inversions and flux data. She never hesitated to provide useful suggestions and valuable comments whenever I need help. Thanks to her concerns when I was sick.

Great thanks to my examiner Harry Lankreijer who also provided me useful suggestions for my research.

Great thanks to all my dear friends. Without their company, encouragements and love, it would be much harder for me to finish this thesis work.

Finally, I must express my very profound gratitude to my families and to my boyfriend. Thank them for patiently listening to my complaints and encouraging me when I was feeling down. They are always there to provide me support during the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Table of Contents

1 Introduction

1.1 Background

The Earth's surface is usually approximated as a spherical surface or an ellipsoidal surface. Finding the area of polygons on a spherical/ellipsoidal surface is important in many fields of earth science (Bevis and Campereri 1987), especially for the problems requiring area of polygons on spheres/ellipsoids with high precision. Density is an important indicator for geographical analysis, e.g., density of some species in some regions, the distribution of population in a country, flux of water or greenhouse gases etc. Field data is measured on the Earth's surface at different sampling points and point data is presented on a spherical/ellipsoidal surface. To obtain the density of point data in the given region, the area of the region must be estimated. Conversely, estimating the area on the spherical/ellipsoidal surface is a problem needed to be solved if the information of density is already known but the total sum of the data is required in the given domain. Besides, areal interpolation, making maps showing accurate regions' areas and other geographical problems ask for estimation of area of a polygon on the spherical/ellipsoidal surface. There are different methods to estimate area of polygons on the Earth's surface depending on the application the method works for. Usually area estimation is done on the spherical/ellipsoidal surface directly (Bevis and Campereri 1987) or 2-dimensional plane by transforming coordinates into the plane coordinate system.

In the field of carbon modelling, one application of estimating polygons' area is to estimate carbon flows for regions. Carbon flux models have been developed to characterize, understand and estimate carbon fluxes at global and regional scales because global warming arising from enhanced greenhouse effect has caused extensive concern in scientific circles in the past decades (Schneider 1989). Carbon flux models are developed based on different approaches, modelling strategies, process representation, boundary conditions, initial conditions, and driver data (GCP - Global Carbon Project 2017). Atmospheric inversions or top-down analyses, is one type of carbon flux models using inverse approaches. Inversion techniques are used to verify "bottom-up" emission estimates which have considerable uncertainties (Bergamaschi et al. 2015).

1.2 Problem statement

In earth science, it is common that field data (e.g., water or carbon fluxes or number of species) are presented in grid which cover the Earth's surface (Figure 1.1). In order to estimate the sum or density of data in a given region, the area of the region must be estimated. The region usually covers the whole grid cells and partial cells. The area of the region is calculated as the sum of area of each grid cell covered by the region.
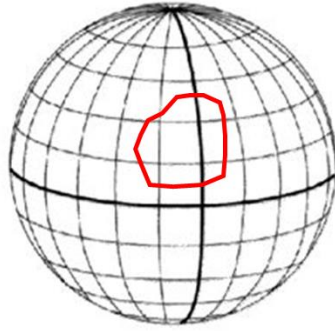
Fig 1.1 A region covers a grid with field data. Red polygon presents the extent of the region. 1 grid cell is completely covered and 8 grid cells are partially covered by this region.

Region area estimation is inevitable for regional carbon flow estimation. Estimating carbon flows for regions is a key step when performing intercomparison between inverse models on different scales (i.e., global, regional and country level) based on carbon fluxes (flow rate per unit area) in grid with various resolutions provided by inversions. Results of different inverse systems differ due to different choice of atmospheric data, transport model and prior information. So comparison between estimates of regional carbon flows from different inversions is also important to analyze model characteristics and to identify the potential model shortages (Bergamaschi et al. 2015). Many researches show results of carbon flows in regions from different inversions and perform comparisons (e.g., Bergamaschi et al. 2015; Stephens et al. 2007; Kirschke et al. 2013; Thompson et al. 2014; Peylin et al. 2013).

However, the methods to estimate the extent of regions in current research work are not clear and no quality assessment of the carbon flows in regions computed is provided. In other words, the uncertainties introduced by area estimation to the final result in the application estimating regional carbon flows is unknown.

1.3 Aim

Based on the above problem statement, the main task of this thesis work is to develop a methodology to estimate area of a polygon and to apply it to regional carbon flow estimation (Figure 1.2). In detail, this research work is divided into four steps:

- Develop a methodology to estimate area of a region (defined as a vector polygon) on a sphere/ellipsoid based on a grid.
- Estimate the uncertainty of the estimated area of the spherical/ellipsoidal region.
- Apply the methodology to estimate uncertainties in land/sea masks used in inversion systems.
- Employ the methodology to estimate carbon flow in Sweden as well as assessing the uncertainties in this estimation.
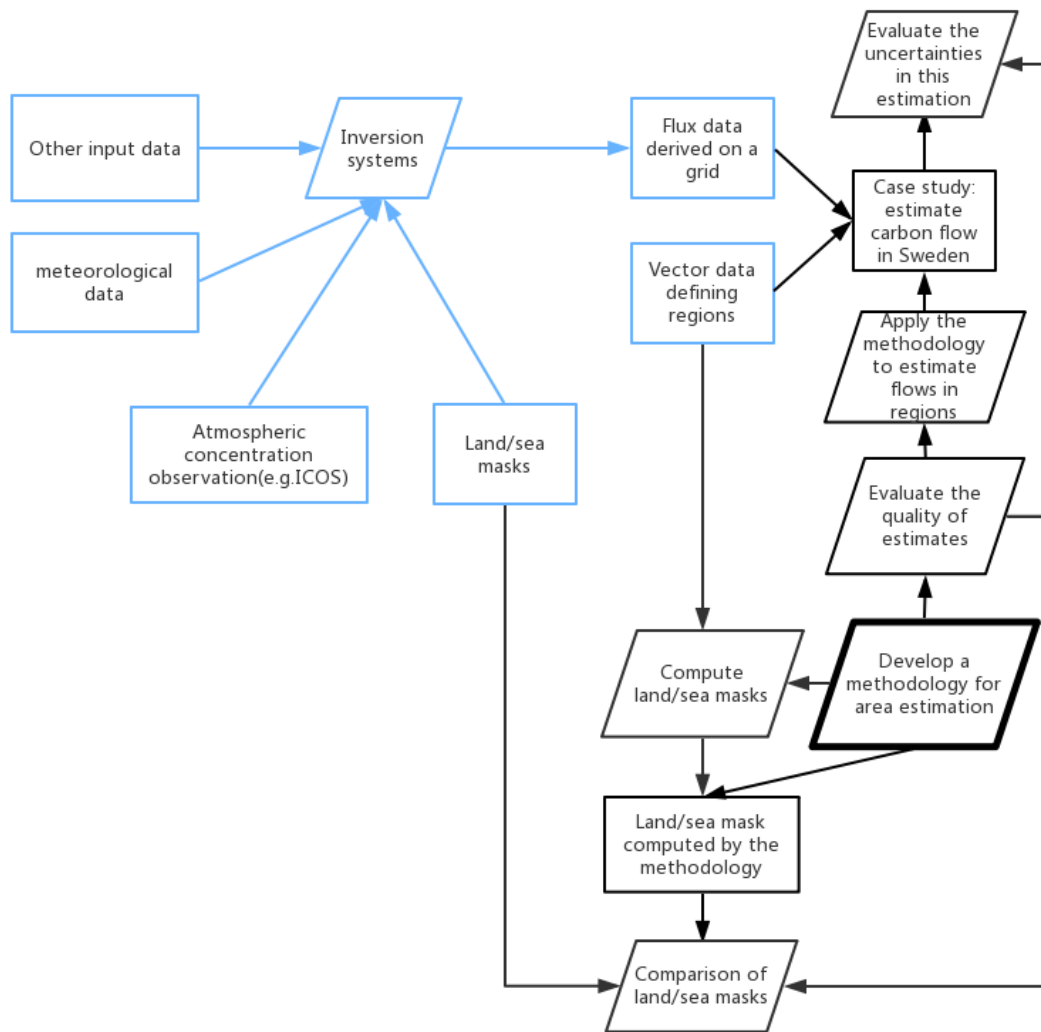
Fig 1.2 Study design. Black part is the contributor of this thesis work. The diamond in bold is the core work of this thesis project.

1.4 Project setup

This project work includes two main parts. The first part is development and evaluation of a method to estimate the area of a region defined as a vector polygon on spheres and ellipsoids and the second part is to employ the methodology developed to estimate region area to estimate carbon flows for regions and on land/sea masks.

For estimation of the area of a vector polygon, several methods were proposed, and a test was designed to evaluate the uncertainties of estimates of different methods by comparison and decide the most suitable one. The second part of this thesis work is to employ the method developed in the first part to estimate flows in Sweden based on different carbon fluxes in grid from different inverse systems. In this study, flux data from three inversion systems CTE, CAMS and Jena CarboScope were used. The evaluation of this methodology on the application of estimating regional carbon flow was done by testing the effects of different components on the estimated result, respectively. Also, the method was applied to estimate land/sea masks

based on three grids in the three inversions using vector data collected from Esri. And comparison between different results of land/sea masks were compared.

The computation was mainly performed in ArcGIS (ArcGIS 10.3.1 for desktop, ESRI, Redlands, California, United States) and by programming in Matlab (Version R2013b (8.2.0.701), August 13, 2013, MathWorks, Natick, United States) and Python programming (Python 2.7.12 64bits, Qt 4.8.7, PyQt4(API v2) 4.11.4 on Windows) in Spyder (Version 2.3.9, https://github.com/spyder-ide/spyder/releases/tag/v2.3.9), Python development environment for desktop and Jupyter Notebook (Version 4.3.1, https://github.com/jupyter/notebook/releases/tag/4.3.1), online development environment.

1.5 Disposition

This thesis starts with an introduction chapter giving the overview of this project work, including background, problem statement, aim and methodology. Chapter 2 is a theoretical part to describe background knowledge required by this thesis project, including atmospheric inversions, brief introduction of three inversion systems: CTE, CAMS and Jena CarboScope, map projections and spatial interpolation techniques. Following is practical work presented in Chapter 3, Chapter 4 and Chapter 5. Chapter 3 describes development and evaluation of the methodology to estimate vector area of on spheres and ellipsoids. In Chapter 4, the method was used to estimate land/sea masks. Application of the method to estimate region flows and test of effects of different factors on the result of carbon flow during the process of computation can be found in Chapter 5. Discussion and conclusions are following as last two chapters. Programming codes to estimate carbon flow in Sweden and the description about how to compute land/sea masks in ArcGIS are attached as Appendix.

2 Literature review

2.1 Atmospheric inversions

Greenhouse effect is known as the phenomenon that greenhouse gases in the atmosphere trap long-wave radiation from Earth's surface, giving higher global mean temperature than that without atmosphere (Mitchell 1989). The greater the concentration of greenhouse gases in the atmosphere, the more infrared energy kept (Schneider 1989). The increase of surface temperature can be modelled by feeding atmospheric greenhouse gases concentration as input data to the process functions in climate models.

Atmospheric concentration is directly linked to the emission of greenhouse gases. Observation of emission carbon dioxide (CO_2), the most important long-lived anthropogenic greenhouse gas (Bergamaschi et al. 2015), is of great importance. To analyze and predict emission of greenhouse gases at different scales, models are the important tool for scientists

(Bergamaschi et al. 2015). Atmospheric inversion is one modelling technique to estimate greenhouse gas fluxes, which has been widely employed on global and continental scale (Bergamaschi et al. 2015).

Atmospheric inversions provides optimal surface-atmosphere carbon exchange that fit to atmospheric concentration observations. Here the general principle of the inversion systems is provided, a detailed explanation would be out of the scope of this thesis work. Inversion systems consist of four different components: transport models, atmospheric measurements prior information and optimization scheme, which all vary in the different systems (Peylin et al. 2013; Ciais et al. 2010). The aim is to constrain prior surface carbon fluxes by running atmospheric transport and chemistry model in an inverse way using atmospheric concentration measurements as input data so that the prior fluxes can be adjusted to more realistic values. The prior fluxes are usually provided by a terrestrial/ocean dynamical model. Atmospheric CO_2 observations are collected from a global network consisting of more than 100 sites where CO_2 is measured either continuously or with interval using.

The link between net surface carbon exchange (x) and atmospheric concentration (y) can be expressed as the formula below (Peylin et al. 2013):

$$y = H(x) + r \hspace{3cm} (1)$$

where H is the atmospheric transport model, x is net surface carbon exchange, y is observations of atmospheric carbon concentration and r is the uncertainty of y. Figure 2.1 illustrate the principles of inverse models. In the forward way of the model, the prior surface carbon fluxes x_b is fed into the model to simulate atmospheric CO_2 concentration y_b. Inverse modelling consists of varying the input x within the range of uncertainty of prior information x_b to find the optimized surface fluxes by minimizing the difference between atmospheric observations y_{obs} and optimized CO_2 concentrations y_a within their uncertainties (Ciais et al. 2010). The optimized value of x_a is called posteriori CO_2 flux parameters, which is better match the true values. One can also be seen in Figure 2.1 that uncertainties of both optimized surface fluxes x_a and atmospheric concentrations y_a decrease.

Inverse modelling allows not only to minimize the difference between modeled and observed atmospheric concentrations but also to optimize the prior terrestrial carbon fluxes from a biospheric model (Peylin et al. 2013; Chen et al. 2015). Moreover, this strength of optimization of source/sink in inverse models can not only verify simulations from process models but also improve quantification of carbon sinks and sources in poorly observed regions (Chen et al. 2015; Engelen et al., 2002).
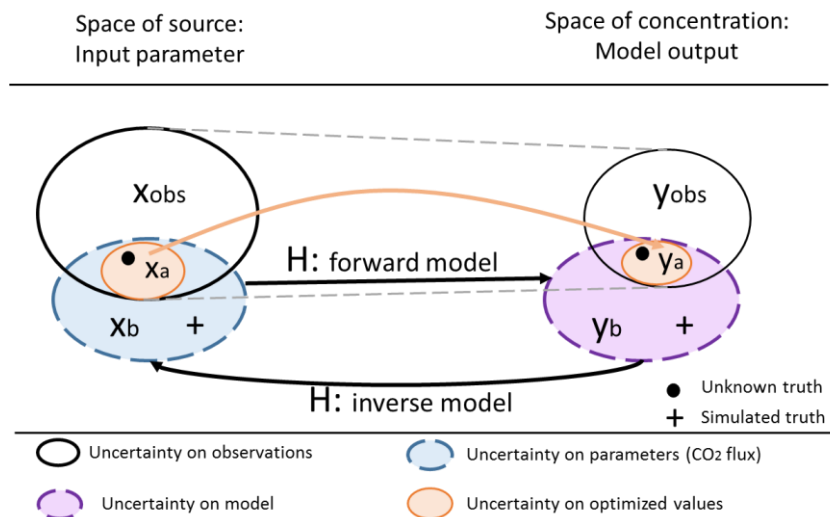
Fig 2.1 Principle of inverse models. Simplified Figure 1 by Ciais et al., 2010. Sign '●' means true values of surface fluxes, '+' describe the simulated values. x_b is the prior surface fluxes with uncertainties. y_b is the modeled atmospheric concentration corresponding to input data x_b. H is the transport model linking input parameter x with model output y. x_{obs} and y_{obs} are observational values of fluxes and concentration with their respective errors. And x_a and y_a are the optimized values of fluxes and concentration.

2.1.1 Three inversion systems

Three inversion systems were selected in this report to study and perform comparison of carbon flows in regions computed by the method developed. The same inversions used in the report: Global Carbon Budget 2016 (Le Quéré et al. 2016). They are CarbonTracker Europe (CTE), Jena CarboScope and CAMS (previously called MACC) (Le Quéré et al. 2016).

In this thesis, the focus is to analyze the contribution of model output differences to regional carbon flow estimation and compare the contribution of model output with the contribution of other components. So only basic information of three inversions is provided, but not the details of system construction.

2.1.1.1 Inversion system CTE

CarbonTracker Europe is one of systems employed within CarboEurope program using atmospheric inversion technique (Peters et al. 2010). This system is one of the first systems to use the semi-continuous measurements from European continental sites (Peters et al. 2010). The transport model used in this system is the global chemistry transport model TM5 which applies two-way nested zooming algorithm (Krol et al. 2005). TM5 inherits many concepts and parameterizations from TM3 but extends the functionality of TM3 to zoom into some special regions by refinement in both space and time. TM3 is the global chemistry transport model to solve the continuity equation for an arbitrary number of atmospheric tracers (Heimann and Körner 2003). It is driven by meteorological fields calculated by a general circulation model or

a weather forecast model (Heimann and Körner 2003). TM5 is applied to regional carbon budget study requiring high resolution and also to global study with coarser resolution (Krol et al. 2005). The grid resolution over Europe in this transport model is $1^{o} \times 1^{o}$ (Figure 2.2) (Krol et al. 2005).

Fig 2.2 Horizontal grid definition of transport model TM5. *Recreated from Krol et al. 2005 by Ute Karstens, ICOS Carbon Portal, Lund University. Copied from Ute Karstens with her permission.*

2.1.1.2 Inversion system Jena CarboScope

In the inversion system Jena CarboScope (Rödenbeck et al., 2013) the link between surface fluxes and atmospheric concentration is established by the TM3 global atmospheric tracer model. Jena CarboScope (hereinafter referred to as 'Jena') is developed by the Max Planck Institute for Biogeochemistry (MPI-BGC). The horizontal spatial distribution used in this study is 48x72 grid, approximately 4°latitude ×5° longitude, except cells (approximately 2°latitude ×5° longitude) in the first low and last row in the polar region (Figure 2.3) (Rödenbeck et al. 2013).

Fig 2.3 Horizontal grid definition of transport model TM3. *Recreated from Heimann and Körner 2003 by Ute Karstens, ICOS Carbon Portal, Lund University. Copied from Ute Karstens with her permission.*

2.1.1.3 Inversion system CAMS

CAMS inversion system uses LMDZ general circulation model for the transport calculation. LMDZ is the second generation of a climate model developed at Laboratoire de Météorologie Dynamique (Hourdin et al. 2006). The horizontal resolution of LMDZ model for this study is 96 x 96 and each grid cell is around 2° latitude $\times 4^o$ longitude, except cells (approximately 1° latitude $\times 4^o$ longitude) in the first row and the last row in the polar regions (Figure 2.4) (Baek et al. 2014).
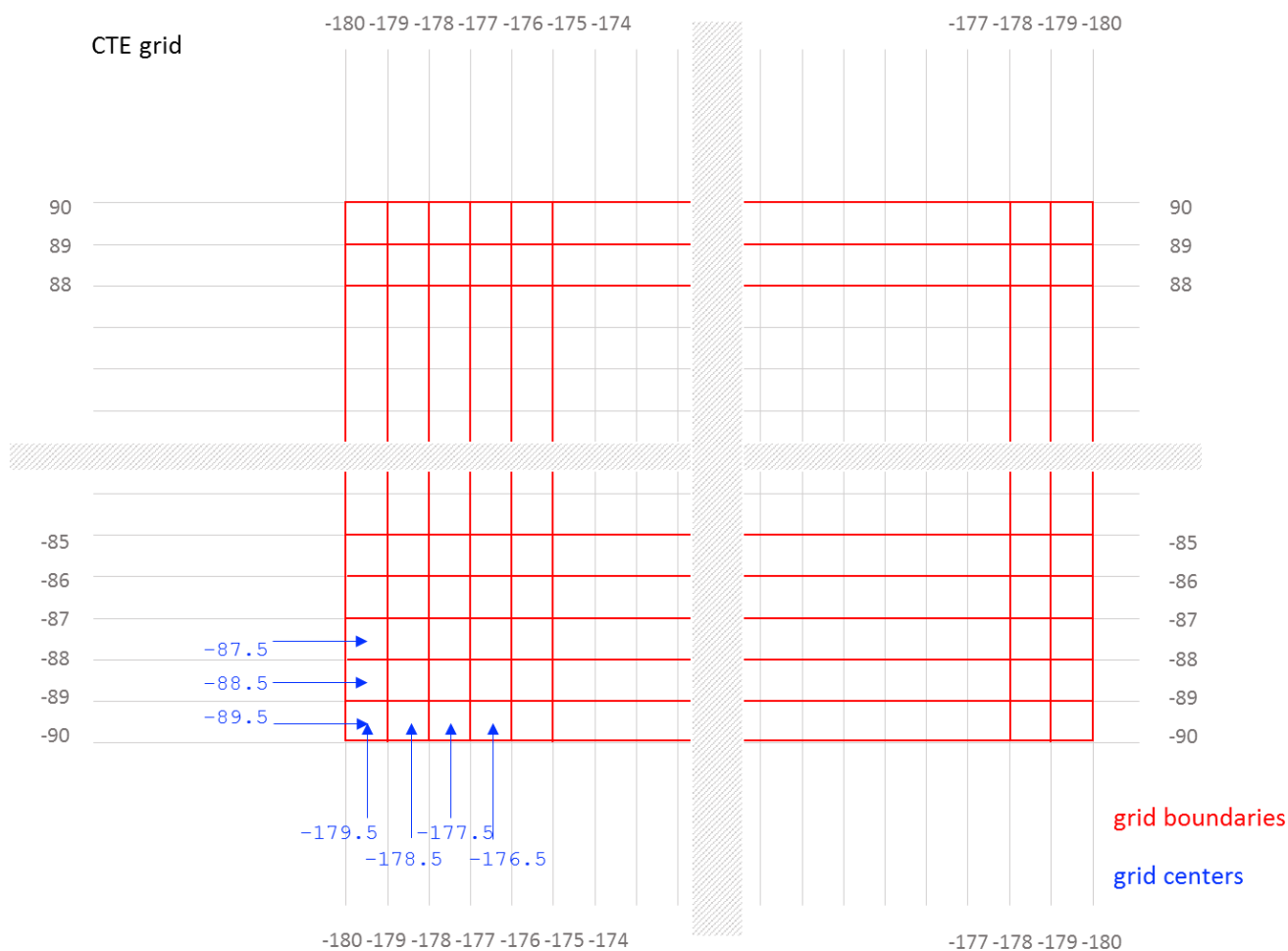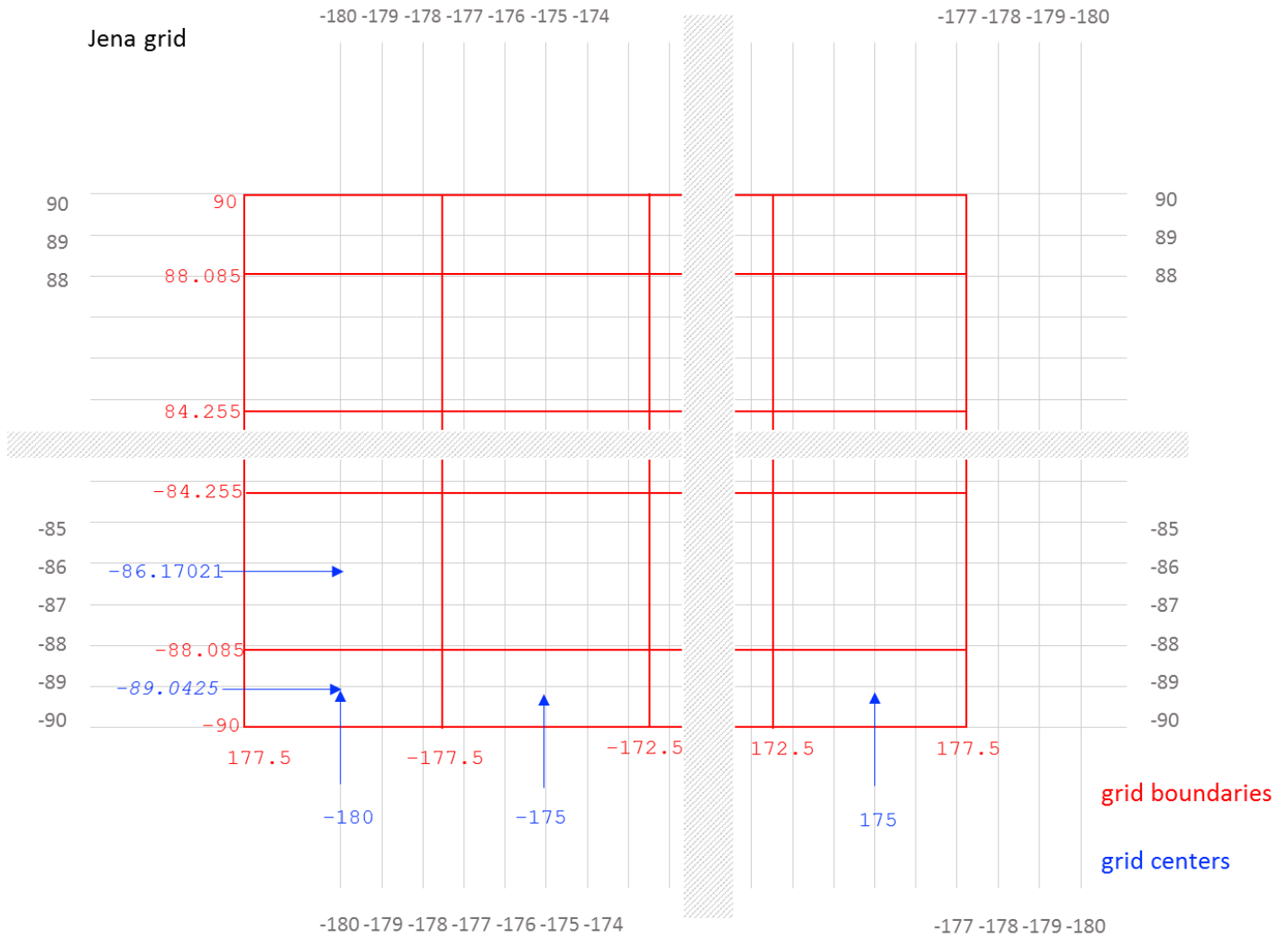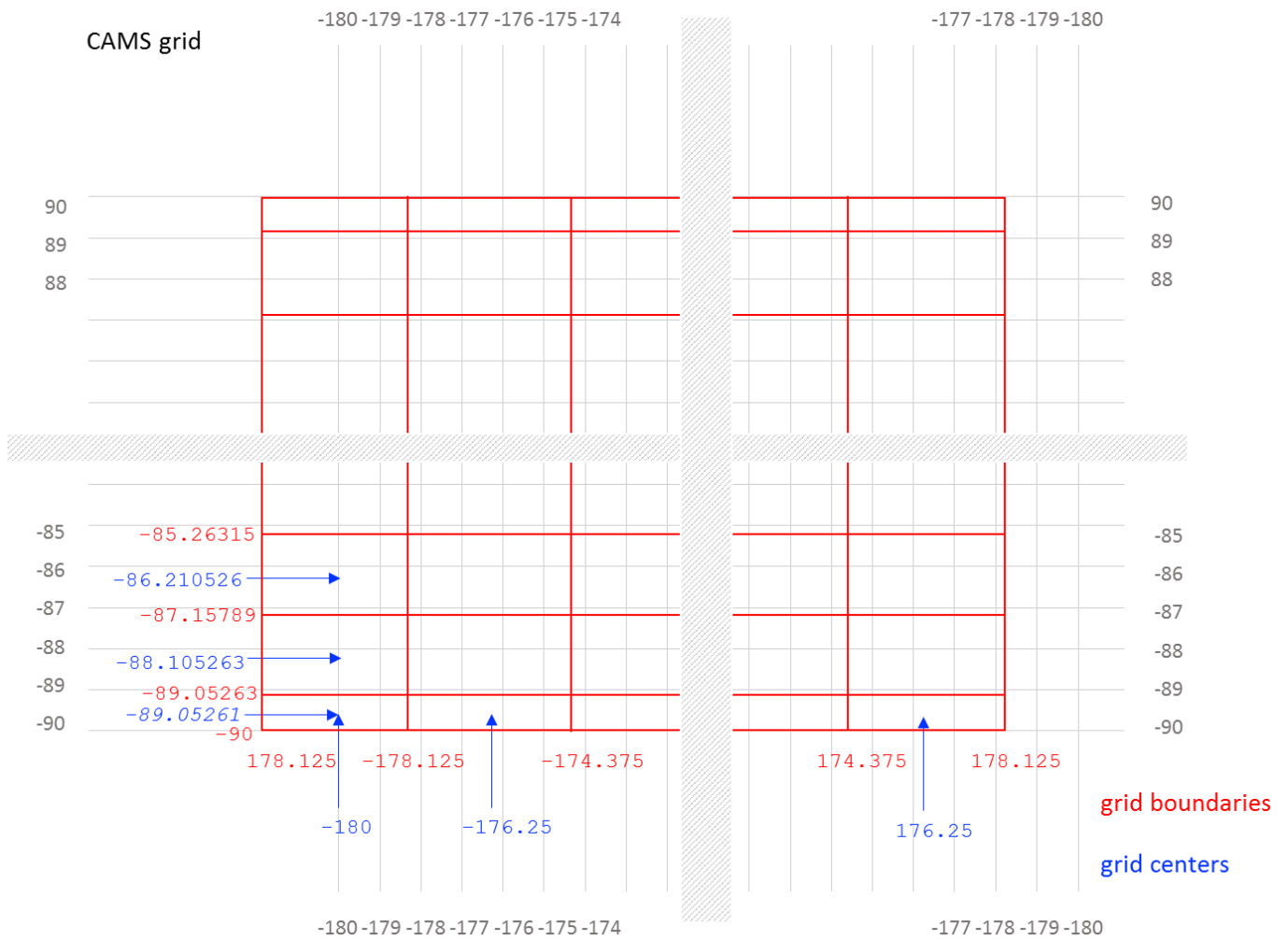
Fig 2.4 Horizontal grid definition of transport model LMDZ. *Recreated from Baek et al. 2014 by Ute Karstens, ICOS Carbon Portal, Lund University. Copied from Ute Karstens with her permission.*

Difference between transport models used in inverse systems is the driving factor behind the differences of flux estimation (Le Quéré et al. 2016). Besides transport models, various atmospheric CO_2 measurements, prior information and optimization scheme contribute to the difference between estimates. The CO_2 atmospheric observations driving the three inversions are collected from various flask and *in situ* networks (Le Quéré et al. 2016). Prior information is the estimate of surface carbon exchange, consisting of anthropogenic and natural components (Peylin et al. 2013). The natural surface fluxes in the three inversions are provided by different models: SIBCASA model for CTE, LPJ model for Jena (Le Quéré et al. 2016) and TURC model for CAMS (Chevallier et al. 2005). The prescription of anthropogenic flux varies in the three inversions as well. Since the focus of this thesis is not about inverse techniques and only model outputs were used for this study, no further detailed information about setup of these three inversion systems is given in this thesis

2.1.2 TransCom project

Atmospheric Tracer Transport Model Intercomparison Project (TransCom) was a special project of the International Geosphere-Biosphere Program (IGBP), Global Analysis, Interpretation, and Modeling (GAIM) Project. The aim of the TransCom project was to perform various types of intercomparisons between inverse models to quantify and diagnose uncertainties and biases of calculations of global carbon budget which are due to simulation errors of atmospheric transport model, uncertainties in measured CO_2 data used, and in the inverse approaches applied (Michaut 2017).

The TransCom project has evolved through several phases. The first two phases (TransCom 1 and TransCom 2) have been completed with third phase (TransCom 3) underway and fourth phase initiated (The TransCom Experiment 2017). The initial phase of TransCom performed model simulations of fossil and bio-spheric emissions of CO2 and compared them to characterize model behavior (Law et al. 1996; Gurney et al. 2003). The second phase conducted simulations of sulfur hexafluoride (SF6) emissions using eleven 3-dimension tracer models. The goal of this phase was to compare model simulations between models and observations and to diagnose different or unrealistic results models produced (Denning et al. 1999; Gurney et al. 2003). After useful insights into model behavior provided by the experiments conducted in phase one and two, the TransCom project transferred its focus to assess the sensitivity of flux estimates for annual mean, seasonal cycle and inter-annual variability to the transport model used as well as other factors in inversion process, e.g., observational data choices, priori flux uncertainties (Baker et al. 2006; Gurney et al. 2003). In this phase (TransCom 3), carbon fluxes estimation and intercomparison of those fluxes were made on 22 pre-defined emission regions (Figure 2.5) (e.g., Baker et al. 2006; Deng et al. 2007; Gurney et al. 2000; Law et al. 2003).
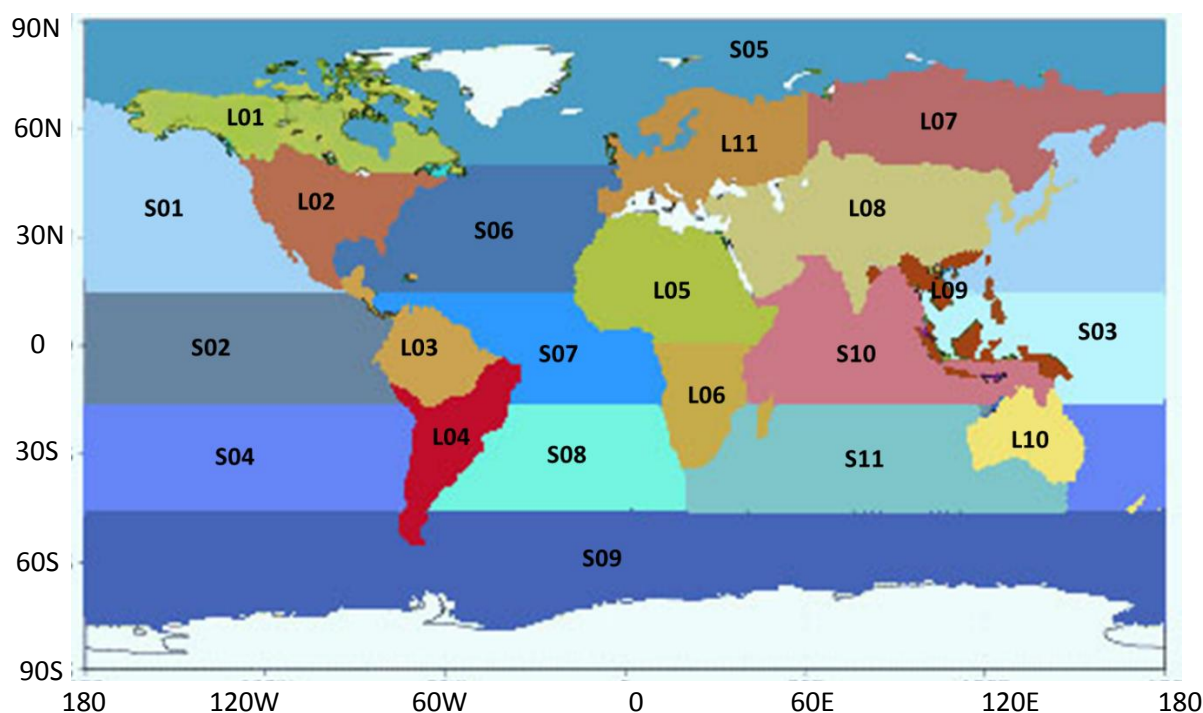
Fig 2.5 TransCom-3 divided land and ocean regions (11 each, 22 in total) of the globe. L01: Boreal North American; L02:Temperate North American; L03: Tropical American; L04: South American; L05: Tropical Africa; L06: South Africa; L07: Boreal Asia; L08: Temperate Asia; L09: Tropical Asia; L10: Australia; L11: Europe; S01: North Pacific; S02: West Pacific; S03: East Pacific; S04: South Pacific; S05: Northern Ocean; S06: North Atlantic; S07: Atlantic Tropics; S08: South Atlantic; S09: Southern Ocean; S10: Tropical Indian Ocean; S11: South Indian Ocean . The white area is not considered. Modified *from Patra et al. 2003.*

2.2 Map projection

2.2.1 Properties of map projection

A map projection is systematic representation of part or all of the Earth's surface in 3 dimensions onto a 2-dimensional plane (Snyder 1987). It is inevitable that distortion will be introduced from spherical maps into a projected plane map. Cartographers choose map projections according to the characteristics which should keep undistorted based on the purpose of the projected map. The characteristics normally considered when choosing map projections are: area, shape, scale and direction (Snyder 1987; Eldrandaly 2006). No projection can preserve all characteristics accurate and keeping one characteristic accurate is at the expense of others. Some projections are a compromise of all the characteristics. A spherical map is projected onto a plane by a developable surface which can be expanded flat without stretching or tearing (Chen et al. 1999). A map projection is constructed by wrapping the Earth by a developable surface and projecting points on the Earth's surface along rays from a projection center onto the developable surface (Eldrandaly 2006). There are three most commonly used developable surfaces: cylinders, cones and planes.

On the surface of the Earth, a graticule or a network consisting of latitude and longitude lines is used to locate the points (Figure 2.6). The graticule lines are usually called parallel circles and meridians, respectively. Parallels of graticule are formed by latitude circles around the Earth and they are parallel to the line/curve of Equator on planes. Meridians of graticule are formed by longitudes circles which intersect two poles (Snyder 1987; Kennedy and Kopp 1994). Figure 2.7 shows the projected graticule on the plane by cylindrical equal-area projection.
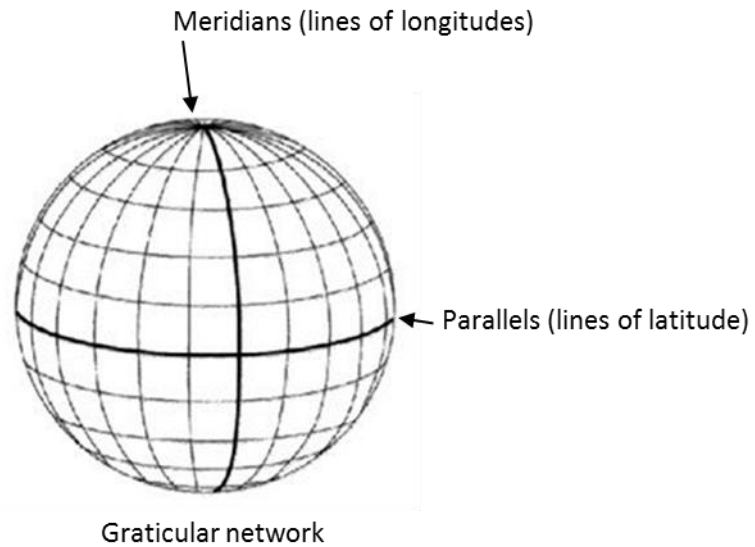
Fig 2.6 A graticule formed by the parallels and meridians.
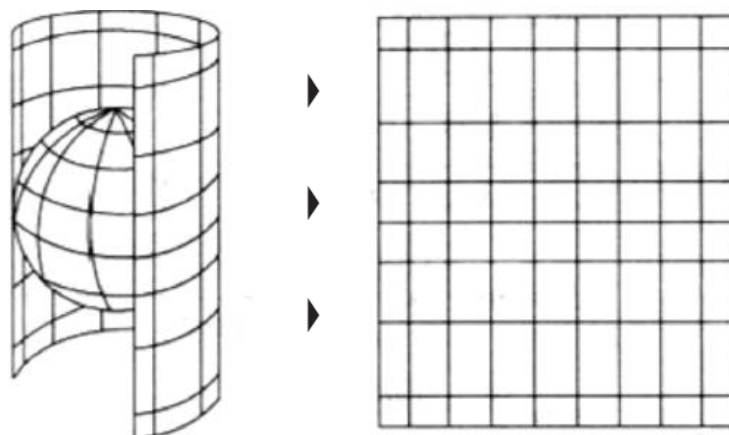
Fig 2.7 The graticule projected onto a cylindrical projection surface.

Projections can be classified according to the developable surface used and the property undistorted. Following are some common used projections (Snyder 1987):

- Equal-area projections: this kind of projections preserves area, which means the area of a coin on any size on projection map is the same as the area of it on the Earth's surface. Some common area-preserving projections: Cylindrical equal-area projection (Figure 2.11), Albers Equal-Area Conic projection (Figure 2.12) and Lambert Azimuthal Equal-Area projection (Figure 2.13).

- Conformal projections: this kind of projections preserves shape (angle). However, there is still distortion in shape of large areas, though shape is essentially correct for small areas. One of the most famous conformal projections is Mecator projection. Transversal Mecator projection is used in all map service (e.g., Google Map) and all topographic mapping. Conformal projections are also suitable projection for navigation maps or weather maps (Eldrandaly 2006). Lambert Conformal Conic is one of common conformal projections.
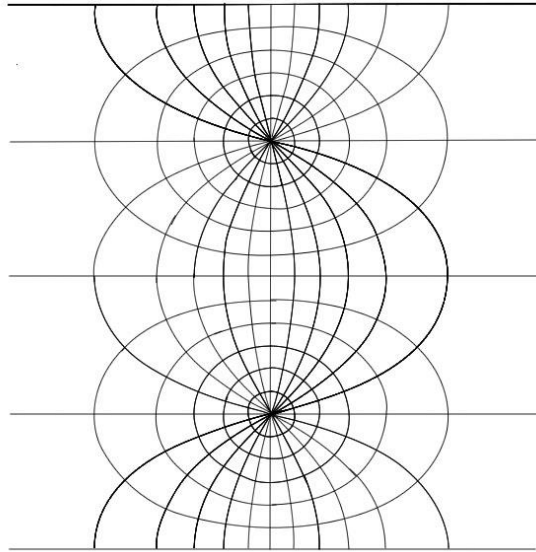
Fig 2.8 Graticule of transverse Mercator projection. This is a cylindrical conformal projection. The middle straight line is the standard (true-scale) meridian which can be any chosen one. *Modified from Snyder 1987*.

- Equidistant projections: this kind of projections presents true scale from one or two points to every other point or along every meridian. It cannot preserve distances from all the points to all other points. This projection can also be combined with different developable surfaces: Equidistant Cylindrical projection, Conic Equidistant projection and Azimuthal Equidistant projection.

20° N

60° N
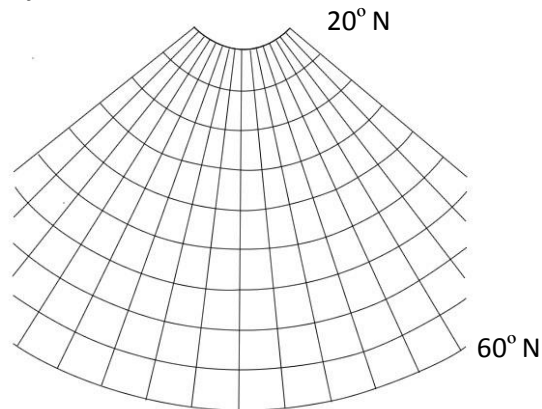
Fig 2.9 Graticule of Conic Equidistant projection with standard parallels 20° and 60° N. *Modified from Snyder 1987*.

- Azimuthal projections: on this projection, the directions or azimuths of all points on the map are presented correctly to the center. These projections can also be equal-area, conformal and equidistant. Common Azimuthal projections: Orthographic projection, Lambert Azimuthal Equal-Area projection and Azimuthal Equidistant projection.

13
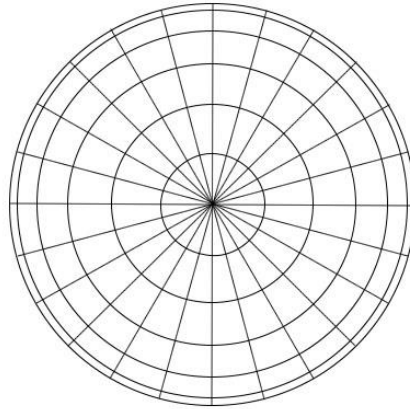
Fig 2.10 Graticule of Orthographic projection at polar aspect. *Modified from Snyder 1987*.

- Other projections: these projections do not preserve any property but reach a compromise between each property. For example, Miller cylindrical projection is neither equal-area nor conformal. It compromises between Mercator (conformal cylindrical projection) and other cylindrical projections (Snyder 1987). This projection is used for world maps and in several atlases.

2.2.2 Examples of some area-preserving map projections

In this project, one important task is to estimate grid cell size and fraction of overlay area in a cell. It is hard to compute area of a polygon on a 3-dimensional surface. To simplify computation, map projections can be used to project 3-dimensional polygons into 2-dimension plane. Equal-area projection maintains the area, which means this kind of projection could be the best choice if polygon area is estimated by projecting the Earth's surface to a plane. In the family of equal-area projections, cylindrical equal-area projection, Albers Equal-Area Conic projection, Lambert Azimuthal Equal-Area projection are some common ones.

Cylindrical equal-area projection is an orthographic projection of sphere onto cylinder (Snyder 1987) (Figure 2.11). In normal case of cylindrical equal-area projection with the Equator as standard parallel, meridians are equally spaced straight lines and perpendicular to unequally straight and horizontal parallels. Meridians and parallels are orthogonal straight lines. There are also transverse or oblique cylindrical equal-area projection depending on the location of the standard lines. There is no distortion of area anywhere on the map and the scale and shape of standard parallel in normal case, standard lines in transverse or oblique cases are undistorted. There is severe distortion of shape and scale in high-latitude area and most severe distortion appears at the pole in normal case. This projection is well suited for making maps of area near the Equator or areas predominantly extending north-south (Snyder 1987).
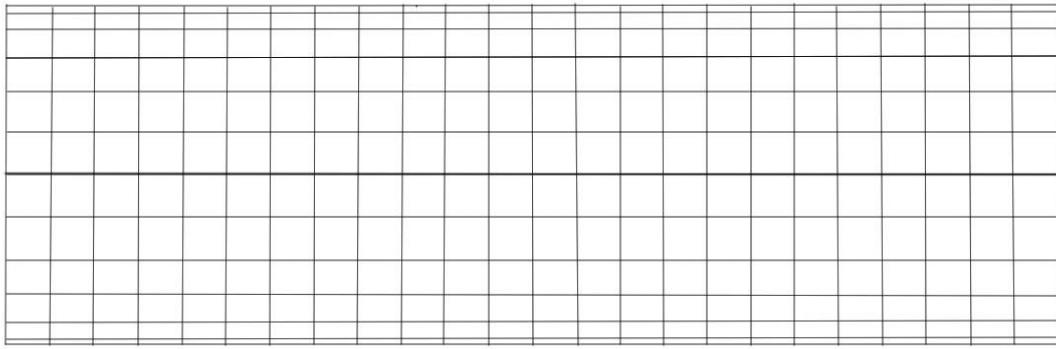
Fig 2.11 Graticule of cylindrical equal-area projection with standard parallel as the Equator. *Modified from Snyder 1987.*

Contrast to normal cylindrical equal-area projection, Alber Equal-area Conic projection is well used for equal-area maps of regions predominantly in east-west extent. For this projection, parallels are arcs of concentric circles with unequally interval and the parallels are closer and closer at the North Pole or South Pole. Meridians are equally spaced radii of the concentric circles, gathering at the north or South Pole and extending outward (Snyder 1987). The maps on this projection look like a sector (Figure 2.12). There is no distortion along standard parallels (normally only one standard parallel) and angles at which meridians and parallels intersect. However, there is still distortion for other angles.
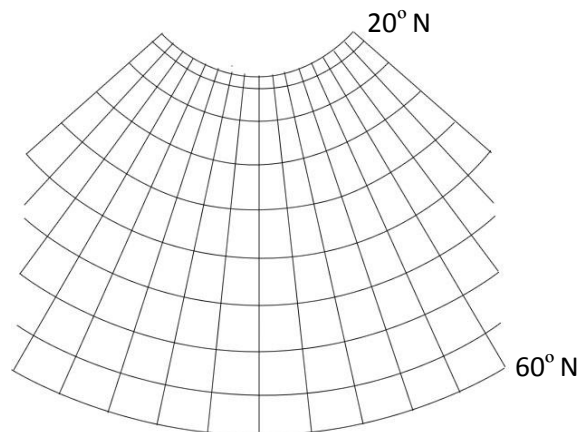
Fig 2.12 Graticule of Albers Equal-Area Conic projection, with standard parallels 20° and 60°N. *Modified from Snyder 1987.*

Lambert Azimuthal Equal-Area projection produces an equal-area map of a sphere within a circle (Tobler 1964). Normally, maps on this projection cannot show beyond one hemisphere (i.e., areas beyond the Equator cannot be shown on the polar Azimuthal Equal-Area projection). In polar aspect, all meridians are straight lines and all latitudes are concentric circles. In Equatorial aspect, the Equator and central meridian are straight (Figure 2.13(a)). All other meridians are curved and the two outmost meridians are 90th meridians east and west from the center meridian consisting a circle on the map (Figure 2.13(b)). In oblique aspect, the central meridian is straight and all other meridians and parallels are curved lines (Figure 2.13(c)). The center point is the only point without distortion and the distortion becomes more

15

and more severe as the distance from the center increases. This projection has advantage to make maps of hemisphere or continents, especially polar region (Snyder 1987).

| a | b | c |
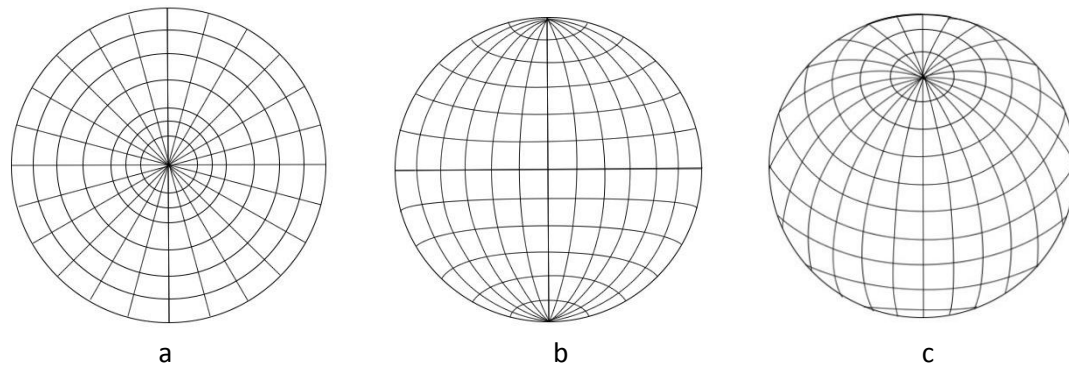
Fig 2.13 Graticule of Lambert Azimuthal Equal-Area projection: a. polar aspect; b. equatorial aspect; c. oblique aspect. *Modified from Snyder 1987.*

2.2.3 Earth models

The Earth is approximately an oblate ellipsoid rotating around its short axis (Snyder 1987). The necessity to study the Earth requires to model the Earth into a regular geometry to facilitate all the measurements and computations, i.e., locating a point on the Earth surface and computing the distance between two points. There are two main approximation of the Earth: spherical earth model and ellipsoidal earth model. A spheroid, also called an ellipsoid, is the definition of the shape and size of the Earth. Ellipsoids are defined by the length of its semi-major, semi-minor axes, flattening and the location of the center (whether it is identical to the Earth's center), a particular ellipsoid only the best fit a part of the Earth's shape (Lu et al. 2014). The Geodetic Reference System of 1980 (GRS80) is the internationally standardized model that is widely used. Although the Earth is more shaped as nearly an ellipsoid, the Earth is sometimes assumed as a sphere to make mathematical computation easier in many cases.

2.2.4 Rhumb line and geodesic line

The shortest distance between two points on a surface is a geodesic line. On a sphere, the geodesic lines are part of great circles (Weintrit 2015). A rhumb line (also called loxodrome) is a line between two points on the Earth's surface intersecting meridians at fixed angle (Rickey and Tuchinsky 1980; Snyder 1987; Alexander 2004) and usually it is longer than the geodesic lines (Snyder 1987). If the meridians and latitudes are parallels lines and orthogonal on a projected map (e.g., by a cylindrical projection), the straight line between two points are rhumb line instead of geodesic lines. This is because the angles at which the line intersects meridians or parallels are the same.

2.3 Spatial interpolation

Spatial interpolation is to predict values at other points or for other subareas or to represent

the whole surface based on a given set of spatial data either in the form of discrete points or for subareas (Lam 1983). There are two forms of spatial interpolation: point interpolation and areal interpolation (Lam 1983). Point interpolation has been studied much more extensively than areal interpolation. *Mean value*, *nearest neighbor* and *inverse distance weighting* (IDW) are some basic point interpolation methods which predict the values of unknown points according to neighborhood values. The techniques are based on the assumption that each point influences the resulting value of a point only depending on their distance to it (Mitas and Mitasova 1999). These basic methods are parts of non-geostatistical methods. *Kriging* is geostatistical interpolation technique for point interpolation (Li and Heap 2014; Lam 1983). Similar to IDW, *Kriging* weights the measured points to decide how much influence of each measured point on the interpolated point. But differently, IDW decides each observation location's weight only depending on the distance to the unmeasured location, while the *Kriging* weights are determined by not only distance but also the general form of spatial variation of the entity (Oliver and Webster 1990).

Although point interpolation attracts more concerns in scientific circle, areal interpolation is more frequently required in the field geography than others. There are two approaches applied in this problem: volume-preserving and non-volume-preserving (Lam 1983). Non-volume preserving is a point-based areal interpolation. This approach transforms areal interpolation into point interpolation by overlapping a grid on source zone that needed to be interpolated and assign a control point to each source grid cell. Then value on each control point is assigned by a point interpolation technique. Finally, estimated values on control points within the target zone are averaged to estimate the value in the target zone. For volume-preserving approach, there are two methods: *Overlay* and *pycnophylactic* for volum-preserving areal interpolation. The *Overlay* method computes the value in target zone from area-weights which is determined by the fractional overlay area of original and target zones. This method is suitable for the homogeneous surface. However, the values in most surfaces are not constant everywhere in reality. The *pycnophylactic* method makes improvements to this problem. This method takes adjacent source zones into account by assigning mean density to each source zone superimposed by interpolation zone and make the density on the source zones as smooth as possible (Lam 1983).

The interpolation in my study was done in a grid on a sphere surface. According to the definition of three inverse models to be analyzed, cell value refers to mean flux in a whole cell and flux is constant at every point in a cell. That requires areal interpolation on a sphere surface instead of areal interpolation on planes and *Overlay* areal interpolation could be the suitable method to use.

Conservative remapping is a technique that can be used to remap state variables or fluxes (e.g., water and heat) from one component grid to another and maintain fluxes between different model components (e.g., atmosphere-land and ocean-ice models) which are coupled into a climate system (Jones 1999). Different models usually have independent grid definition. Conservative remapping computes interpolation weights of each source grid cell overlapped by a target grid cell and integrates the flux in the source grid cells according to its respective

weights into the target grid cell. It is classified into first- and second-order accurate remapping (Jones 1999). If the modeled flux data is constant in a cell, the conservative remapping is first-order area-weighted remapping, which corresponds to *Overlay* areal interpolation. Figure 2.14 shows an example of overlay between the source grid and destination grid. The flux in the triangular destination cell k is overlapped by 6 quadrilateral grid cells n (n = 1, 2...6). Interpolation weights are based on the fraction of overlay area of the quadrilateral cells and the triangular cell. And the flux in the new box is the integral of the fluxes in each overlapped quadrilateral cell multiplying with its respective interpolation weights (Jones 1999).

If a cell value only represents a single point in a grid cell, e.g., cell center point, the gradient of the flux in a cell (i.e., the first-order derivatives of the flux) is not zero and the remapping is thus second-order accurate (Jones 1999). The second-order weights are an area-weighted distance from the center point of original cell. For example, the center point in a cell represents the temperature or precipitation in the gridded data of temperature or precipitation for a region. In this study, the flux data is homogeneous in each grid cell, so only first-order remapping is employed to interpolating of carbon fluxes between different grids. Second-order remapping is out of scope of this thesis.
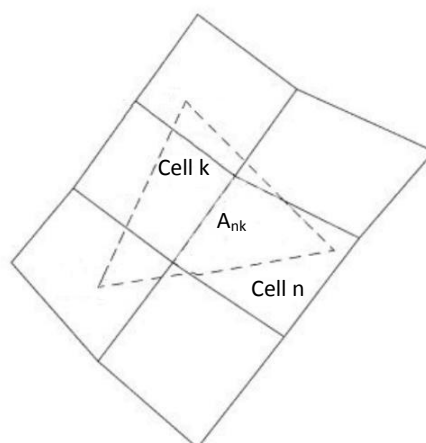
Fig 2.14 A simplified figure from Jones 1999. This presents the example of a dashed-outline triangular destination grid cell k overlapping a quadrilateral source grid cells n.

3 Estimating polygon areas on spheres and ellipsoids

3.1 Background

As described in section 1.2, a vector polygon overlapping a grid usually covers the whole grid cells and partial cells. The area of the vector polygon is the sum of area of each grid cell covered by the polygon. So the key point for estimating area of a polygon is to estimate overlapped area of each grid cell, which can be calculated by the following equation:

$$\text{Overlay area} = \text{cell size} \times \text{partial value} \qquad (2)$$

where partial value is the fraction of overlay area in the whole cell. If the cell is completely covered by a region, partial value equals 1. Estimating overlapped area in one grid cell resolves into two estimations: cell area and fraction of overlay area.

3.2 Methods to estimate partial values

According to equation 2, the problem of estimating vector areas boils down to the problem of computing cell areas and fraction of overlay area in one grid cell. Fraction of overlay area can be computed as overlay area divided by the whole cell area. Three methods are proposed here: 1) using a latitude-longitude plane, 2) using an equal-area projection and 3) using the area of the corresponding spherical polygons. Given that the computation based on an ellipsoidal earth model is quite advanced and complicated, here only the methods are discussed to estimate partial values based on the spherical earth model. The estimation process is the same on ellipsoidal earth model. The only difference is that the equations for ellipsoidal earth model are more complex than spherical earth model.

3.2.1 Using a latitude-longitude plane

To compute the area of a polygon on a sphere, a simple way is to convert spherical coordinate system to a latitude-longitude plane (hereinafter referred to as planar method) as shown in equation 3 below.

$$x = \lambda$$
$$y = \varphi \qquad\qquad\qquad (3)$$

where λ and φ are latitude and longitude defined in the selected geographic coordinate system, while x and y are the coordinates in the local coordinate system.

Then the polygon area can be computed in a local coordinate system as shown in Figure 3.1. In this latitude-longitude local coordinate system, a $1°\text{x}1°$ grid cell is transformed to a 1x1 size square. And its cell area is simply 1. So the partial value is the overlay area. The area of overlay polygon can be estimated by:

$$area\_polygon = \frac{1}{2}\sum_{i=1}^{n}(x_i y_{i+1} - x_{i+1} y_i) \qquad\qquad (4)$$

where n is the number of vertices of the polygon (Worboys and Duckham 2004). The vertices should be ordered in clockwise order and that the first and last point must have multiple identities.

Although this method is easy to implement, there are obvious shortcomings. When spherical coordinate system is converted into latitude-longitude plane, all grid cells become the same squares with size 1. However, the scale of grid cells increase towards the pole in east-west direction, while it is constant in north-south direction. Hence, there is area distortion at high-latitude area. Additionally, the polygon lines in the local coordinate system are not a part of circles but rhumb lines, which change the shape of the polygon to some degree and then also introduce errors. But the errors due to rhumb lines are normally
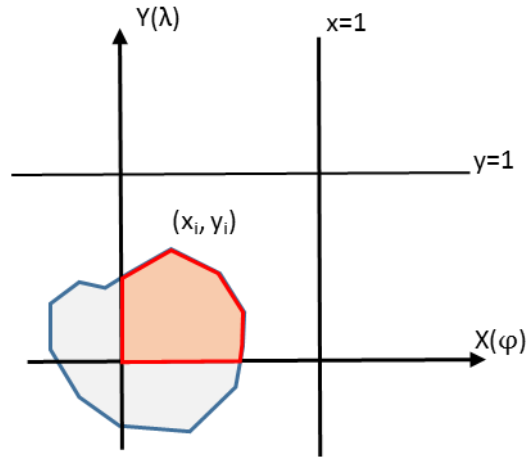
smaller compared to area distortion.

Fig 3.1 The blue polygon is the region for which the carbon flow should be estimated and red polygon is the overlay area in a cell. φ and λ present latitude and longitude, which are transformed into x and y in the local coordinate system. (x_i, y_i) is the coordinate of polygon's break point.

3.2.2 Using an equal-area projection

Projecting 3-dimensional sphere onto a 2-dimensional plane makes computation on the spherical surface much easier. Map projection can be applied to this problem. Since the task is to estimate the polygon area on a spherical surface, the most important thing to consider when choosing map projection is to preserve area on the 2-dimensional map. Given that it is easiest to have a projection that map the graticule (grid) onto an orthogonal grid, cylindrical equal-area projection is the most appropriate choice. The formulas for the normal case of this projection are:

$$x = R \ (\lambda - \lambda_0) \ cos\varphi_s$$
$$y = R \ sin\varphi/cos\varphi_s \tag{5}$$

Where λ and φ are latitude and longitude defined in the selected geographic coordinate system, while x and y are projected coordinates. φ_s is the standard parallel and λ_0 is the central meridian. If the standard parallel is the Equator and the central meridian is made as 0 which coincides with Greenwich meridian, the formulas can be simplified as:

$$x = R \ \lambda$$
$$y = R \ sin\varphi \tag{6}$$

When all the coordinates of polygon's break points are transformed into the local coordinate system, the area of cell and overlay region are estimated by Equation 4. According to the definition of area preserving projection, there is no area distortion anywhere on the projected map. The partial value can be estimated by dividing overlay area by cell area.

This method (hereinafter referred to as 'projection method') also has the same problem as the planar method that the polygon boundaries are also rhumb lines but not great circle arcs.

3.2.3 Using the area of the corresponding spherical polygons

An approach using spherical coordinates to compute area of a spherical polygon is proposed in the article by Bevis and Campereri 1987. The area spherical polygon can be computed by the formula:

$$A_P = R^2 [\sum \alpha_i - (n-2)\pi]$$ (7)

where α is the interior angle of the polygon, n is the number of polygon's vertices and R is the radius of the sphere. If the polygon is a triangle with three vertices, then the formula can be simplified as:

$$A_T = R^2 [A + B + C - \pi]$$ (8)

where A, B and C are the interior angles of the triangle. The angles A, B and C are computed according to rules of spherical trigonometry (Snyder 1987).

The area computed by this approach (hereinafter referred to as 'spherical method') is theoretically most correct. The partial value can be estimated by dividing the area by cell area. However, this computation is complex since it requires computations of the spherical angles. But it is still practicable to compute the area of a spherical triangle ($n = 3$) by this method.

3.3 Evaluation of method to estimate partial values: triangle test

The three methods to compute partial values was evaluated in a test with triangles based on a spherical earth model. The uncertainties of results computed by three methodologies proposed above can reflect the uncertainties introduced to the results computed on an ellipsoidal earth model by the three methodologies. This test was implemented in Matlab by the author and Lars Harrie, ICOS Carbon Portal, Lund University.

In this triangle test, the partial value of a spherical triangle within a 1° x1^o cell was estimated by three methods. The results estimated using the formula to compute a spherical triangle's area (the spherical method) are considered as true values. The quality of results computed by other methods was assessed by compared to true values. For projection method and spherical method, cell size should be known to calculate partial value. Cylindrical equal-area projection was applied to estimate cell areas (discussed in section 3.4) in this test. As discussed above, both planar method and projection method have a shortcoming that polygon break lines are rhumb lines but not great circle arcs. Therefore, the relationship between length of polygon break lines and partial value is interesting to study.

The length of the triangle short edges was measured in degree on lat/lon plane. The triangle was designed to share its two short edge with the cell boundary and it was presented as a right-angled isosceles triangle on the lat/lon plane (planar method) (Figure 3.2). The length of the triangle short edges was changed between 0 and 1 by step 0.5 degree and the location of the cell where the triangle was inside was moved from the Equator to high latitudes (at different latitude bands 0-1, 20-21, 40-41, 60-61, 80-81). The difference of partial value of triangles between methods is shown in following figures (Figure 3.3).
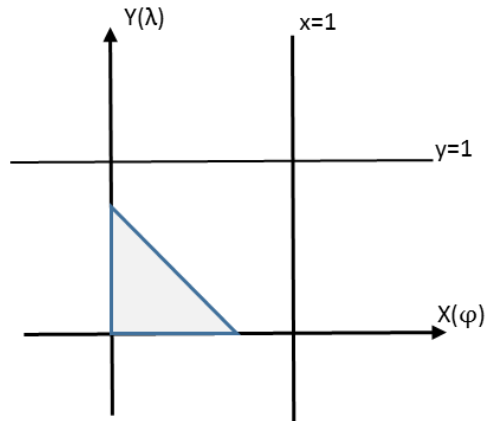
Fig 3.2 Triangle test design. A right-angled isosceles triangle sharing one short edge with the 1x1 cell boundary on a lat/lon plane.

The difference of the triangle's partial value between planar method and the spherical method is larger than that between projection method and spherical method. And the difference between projection method and spherical method is much close to 0 when the break line (triangle's long edge) is smaller than 0.5 degree.

CTE inversion system has the finest resolution of flux data while other inversion systems usually have coarser ones. Therefore the test to estimate partial values of triangle in a bigger cell grid, e.g., $10^{\circ} \times 10^{\circ}$ grid cell, was supplemented. In a $10^{\circ} \times 10^{\circ}$ grid cell, the length of short edges were changed from 0 to 10 by step 0.1 degree. According to results shown in Figure 3.4, differences of partial values computed by projection method and planar method to true values in each subplots are larger than the differences in Figure 3.3 at the corresponding latitude locations. This means that more uncertainty of partial values will be introduced if grid cell size is large.

Fig 3.3 Difference of partial values of the triangle on different locations in a $1^{\circ} \times 1^{\circ}$ grid cell. Red line presents the difference between planar method and spherical method. Blue line presents the difference between projection method and spherical method.
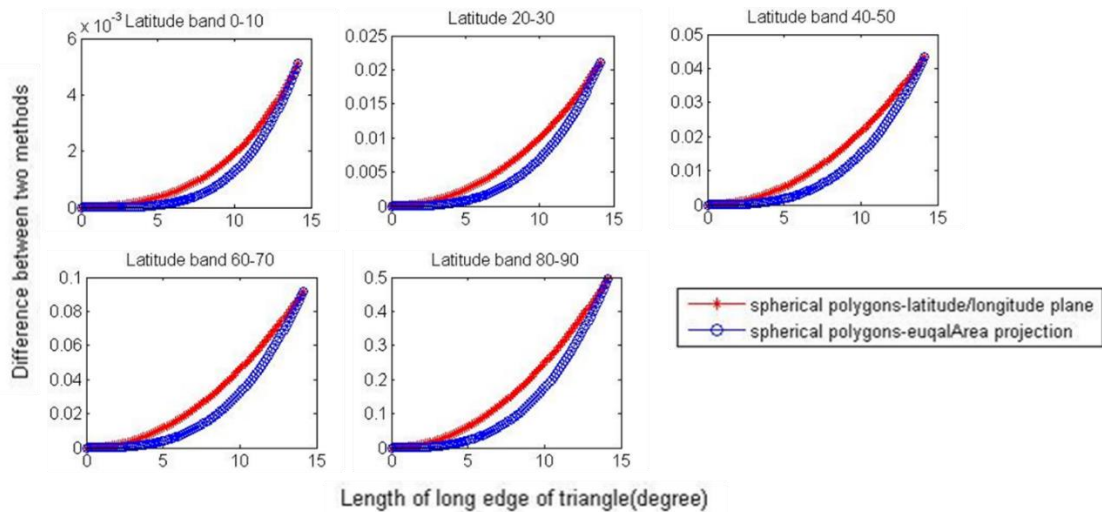
Fig 3.4 Difference of partial values of the triangle on different locations in a 10° x 10° grid cell. Red line presents the difference between planar method and spherical method. Blue line presents the difference between projection method and spherical method.

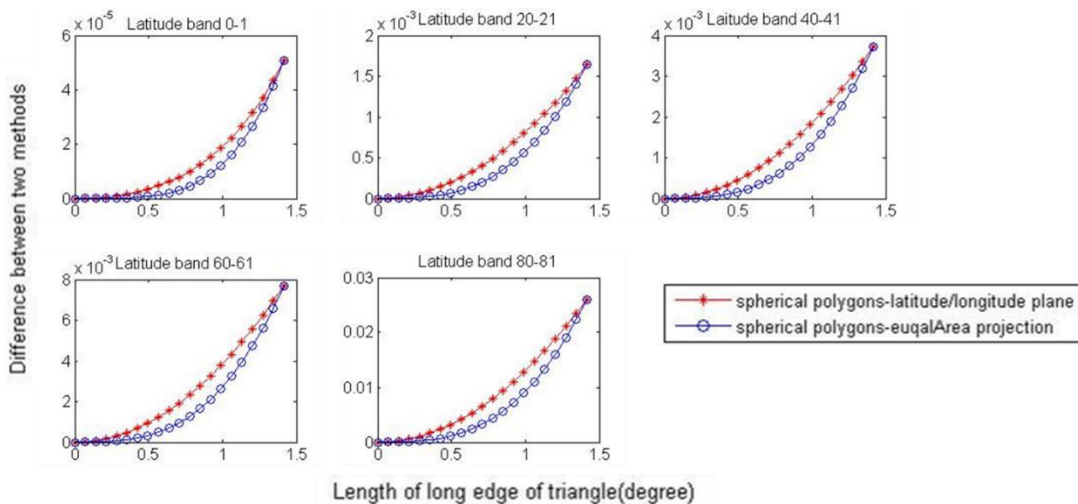## 3.4 Methods to estimate cell areas

Equal-area projection is used to estimate cell areas as discussed in section 3.2. Error will be introduced by this method to estimate partial value due to the shortcoming that break lines of the overlay polygon are rhumb lines on the plane but great circle arcs on the sphere/ellipsoid. However, cylindrical equal-area projection provides true values of cell sizes because cell boundaries are also rhumb lines on the sphere/ellipsoid (Figure 3.5). Polygon lines on the plane are always straight rhumb lines. Cell boundaries consist of latitudes and longitudes which are both rhumb lines and great circle arcs on the sphere/ellipsoid. The projected cell boundaries are also rhumb lines on the plane so that projected cell grids are presented in the right shape. But polygon lines within the grid cell cannot be always rhumb lines on the sphere/ellipsoid.
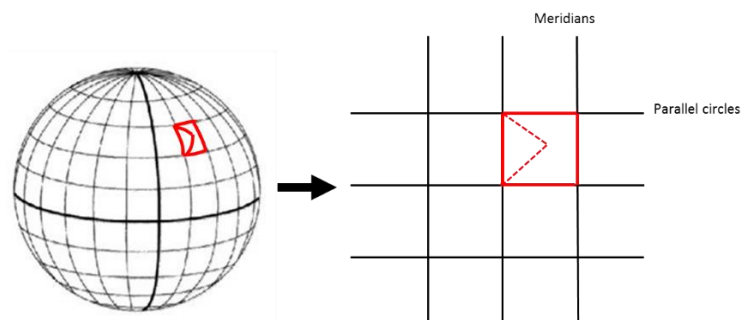
Fig 3.5 One grid cell projected to plane by equal-area projection. Solid red line are cell boundaries and dashed ones are boundaries of overlay polygon.

Even though equal-area projection provides true values, the choice of different earth models will influence the results of cell area estimation. The earth model used for a certain application depends on different purposes. This part only gives estimates of cell areas on a spherical earth

23

model and on an ellipsoidal earth model and comparison between two estimations was performed below. Given that two hemispheres are considered symmetrical, only the results on the Northern hemisphere is given. The computation was done in 1°x1^{o} grid in ArcGIS.

3.4.1 Spherical earth model

The spherical earth model chosen here is an earth-centered sphere with radius 6,371,000 meters. 1°X1^{o} grid with the size of 180x360 were created in cylindrical equal-area projection coordinate system based on the spherical earth model in ArcGIS and the cell areas were computed automatically when the grid was created. Figure 3.6 presents the cell size of a grid cell on each latitude band in the Northern Hemisphere. The size of grid cells on the same latitude band are identical.
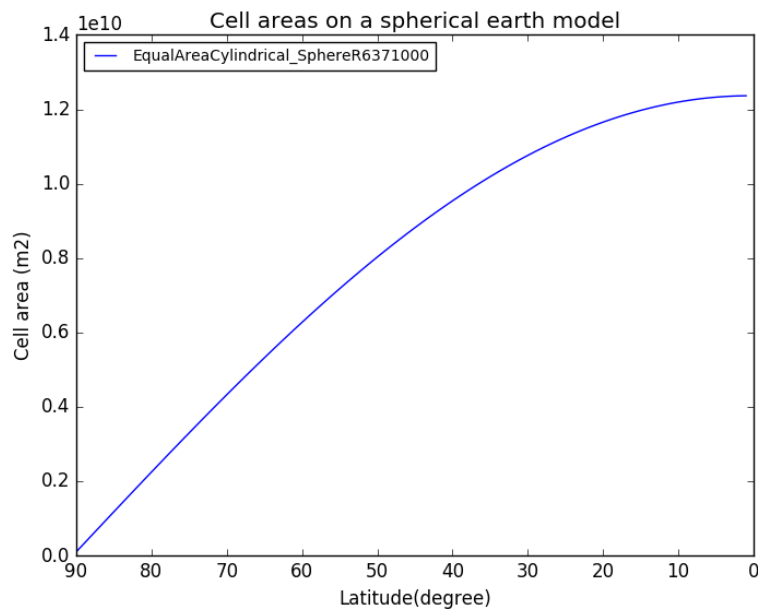
Fig 3.6 Cell areas estimated on a spherical earth model at each latitude band in the Northern Hemisphere.

3.4.2 Ellipsoidal earth model

Ellipsoids can be defined by these parameters semi-major axis and semi-minor axis. Based on these parameters, the flattening and eccentricity can also be derived (van Sickle 2004) (Figure 3.7).
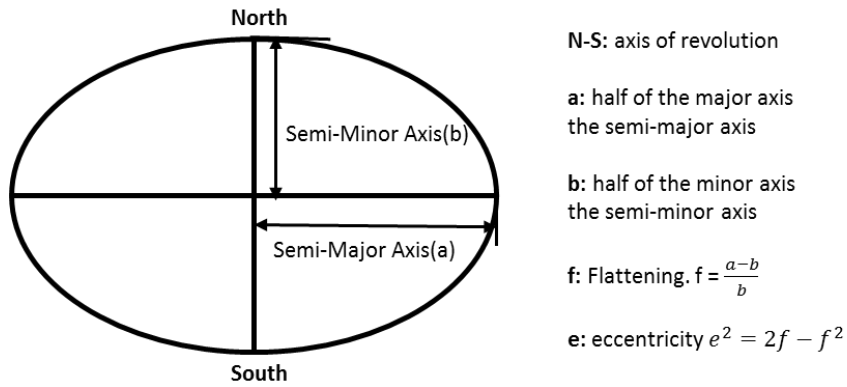
Fig 3.7 Parameters of biaxial ellipsoid

In this study, the ellipsoid GRS80 is chosen as the ellipsoidal earth model. The parameters of GRS80 ellipsoid are (Moritz 1980, presented in van Sickle 2004):

Semi-major (a)	6378137.0 meters
Semi-minor (b)	6356752.31424 meters
1/Flattening (f)	298.257223563
Eccentricity (e^2)	0.00669438002290

The $1^{\circ} x1^{\circ}$ grid was transformed to cylindrical equal-area projection coordinate system based on WGS84 ellipsoid in ArcGIS. The WGS84 and GRS80 ellipsoids only differs on millimeter-level so that they can be treated as the same model. Figure 3.8 presents the results of cell sizes on WGS84 ellipsoid.
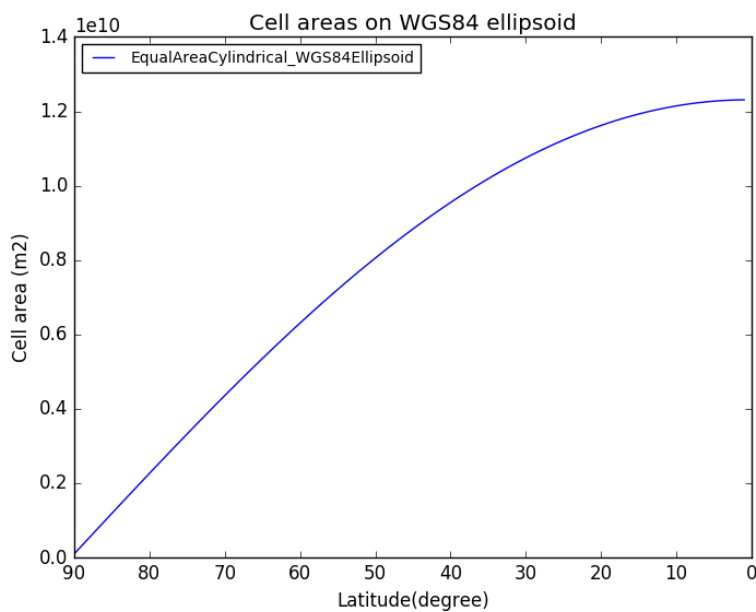
Fig 3.8 Cell areas estimated on WGS84 ellipsoid at each latitude band in the Northern Hemisphere

25

3.5 Evaluation of methods to estimate cell areas

In this part, the comparison between two estimates was performed (Figure 3.9). The relative difference between ellipsoidal results and spherical results is less than 1% for the globe. Highest value appears on the pole and it is decreasing as the latitude is close to the equator. There is no difference between two results around 35° latitude. Cell areas on ellipsoidal model are a little smaller (less than 0.4%) than spherical model around tropical and subtropical regions (latitude < 35°), while estimates on ellipsoidal model are larger than spherical model when the location is more north (Latitude > 35°).

Relative difference in cell areas between ellipsoidal equal-area projection and spherical equal-area projection
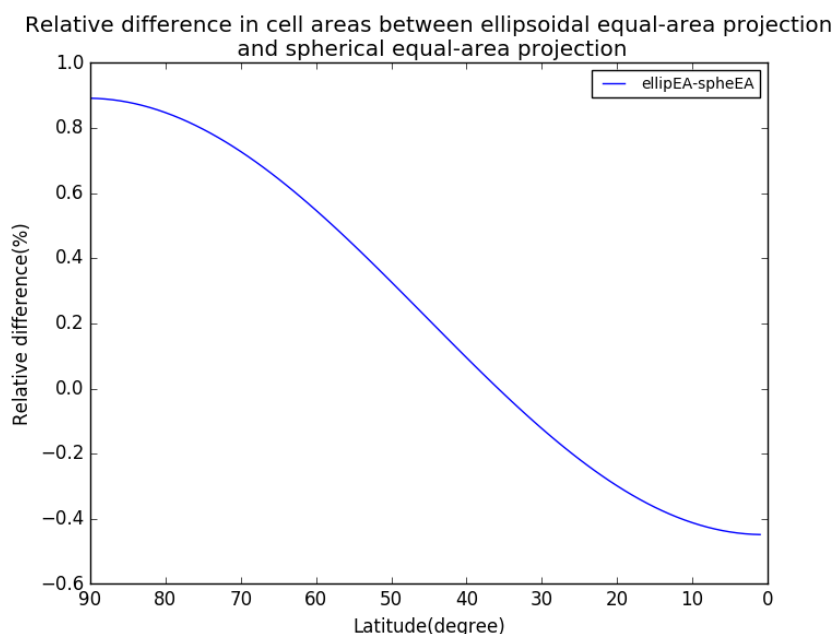
Fig 3.9 Relative difference in cell sizes between the WGS84 ellipsoid and a spherical earth model

4 Land/sea mask test

4.1 Background

Land/sea mask provides the extent of land and ocean all over the world. It is defined in grid with number from 0 to 1 to respect the fraction of land/sea in one grid cell. In other words, land/sea mask gives the partial value of land/sea area in each grid cell. Different inverse systems have different land/sea masks as input data. The definition of land/sea mask affects the models estimation of carbon flows in regions. A case study was performed here to apply the projection method to estimate partial value for in three inversions: CTE, CAMS and Jena. Comparison between the results obtained in this study and inversions' input data of land/sea mask was made to see the differences.

The definition of land/sea mask in CTE model is based on TransCom regions. However, only 0

and 1 but no values in between are used in CTE land/sea mask. The cell value is decided by the dominant region in the cell. But partial values, which are fraction of land or ocean in a grid cell, are used in the land/sea masks in CAMS and Jena inversions.

4.2 Method

Two land/sea masks were computed in $1°x1°$ grids in the cylindrical equal-area projection coordinate system on two earth models, a sphere with radius 6371000 meters and the GRS80 ellipsoid. The area of overlay region between vector data defining land of globe and each grid cell was computed by 'intersect' tool in 'overlay analysis' toolset in ArcGIS. Cell areas were already computed using $1°x1°$ grid created in section 3.4. Preliminary computation results obtained in ArcGIS were exported to Python to create land/sea masks. Partial value in each cell were computed by dividing overlay area in that cell by the cell area (see Appendix).

4.2.1 Vector data used for computing the land/sea mask

The vector data used in this land/sea mask test is a world map collected from Esri (Figure 4.1). This world map is in the geographic coordinate system based on WGS84 ellipsoid, which can be treated as the same model as GRS80.
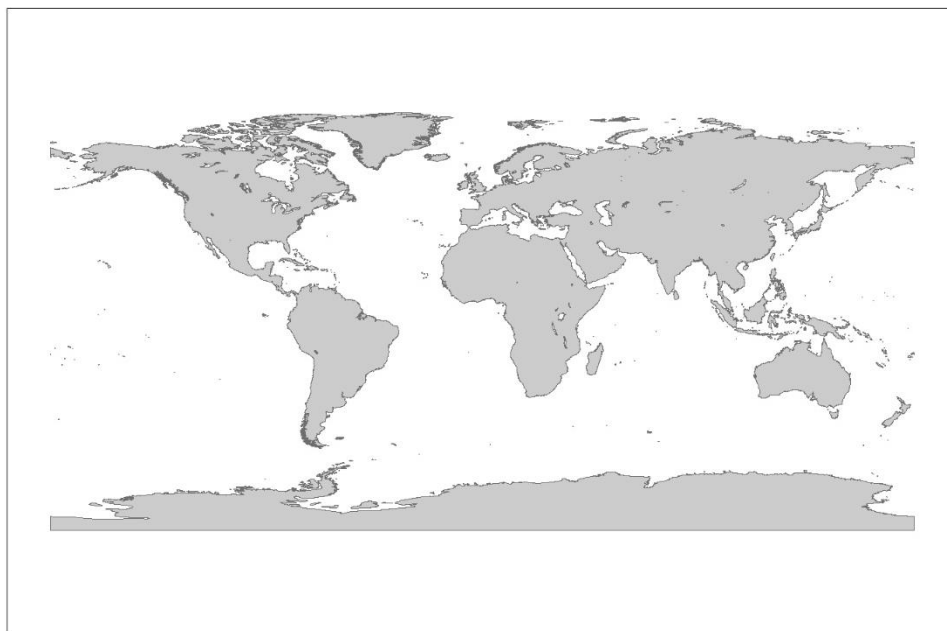
Fig 4.1 World map defined by vector data from Esri

The triangle test showed that equal-area projection method provides that almost true partial values when the length of edges of a polygon are pretty small (Figure3.3 and Figure 3.4). For the vector data used in this study, the mean distances between break points of the polygons from the Equator to high latitudes are smaller than 0.06 degree (Table 4.1). According to the triangle test, the computing process of equal-area projection method propagates pretty small

uncertainties to results using this vector data. The uncertainties were estimated as relative difference of partial values between projection method and true values, which are computed by spherical method:

$$\text{Uncertainty of partial values} = \frac{values\ by\ projection\ method - true\ values}{true\ values} \qquad (9)$$

Where values by projection method and true values were obtained by roughly finding the corresponding partial values in the triangle test when the length of triangle's long edge is equal to the mean distances.

Uncertainty of partial value estimation is less than 1% when mean distance of polygons' break lines are smaller than 0.06 degree.

Table 4.1. Mean distances of polygons' edges at different latitude bands from the Equator to high latitudes

Latitude band (degree)	Mean distance (degree)	Uncertainty of partial value
0-1	0.030	0.55%
20-21	0.035	0.6%
40-41	0.043	0.38%
60-61	0.045	0.5%
80-81	0.059	1%

Besides, the different definition of the extents and boundaries of regions also influence the partial value estimation. And more uncertainties are introduced due to problems of positional accuracy and omitting of some small islands or lakes.

4.3 Comparison of land/sea masks

Comparison of land/sea masks between the masks computed by the method developed and the masks used in inverse systems was made utilizing a Python script created by Lars Harrie, ICOS Carbon Portal, Lund University. Two land/sea masks computed by the projection method on different earth models were made on $1^{o}x1^{o}$ grid. Since the land/sea mask used in CTE model is also in $1^{o}x1^{o}$ grid, land/sea masks provided by CAMS and Jena were interpolated from their original grid onto a $1^{o}x1^{o}$ grid to enable direct comparison.

4.3.1 Comparison of land/sea masks between different earth models

In land/sea masks created on two earth models, partial values of land for grid cells completely covered by ocean or completely covered by land are the same, 0 and 1 respectively. The differences on two earth models only appear in cells partially covered by land. Figure 4.2

shows mean relative difference between partial values of land at each latitude band on two earth models, which is less than 0.02%, mostly less than 0.01%.

Mean relative difference in partial values of land between WGS84 ellipsoid and the spherical earth model on each latitude band
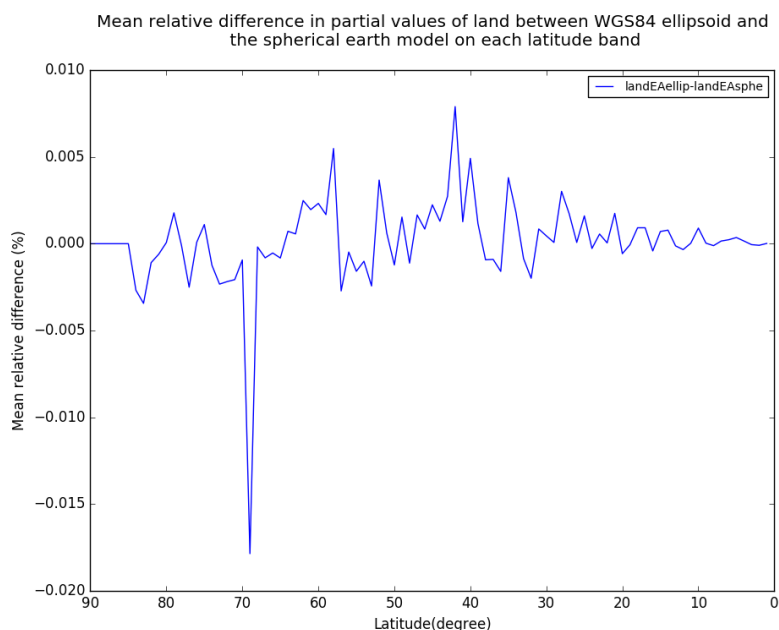
Fig 4.2 Difference in partial values of land between WGS84 ellipsoid and the spherical earth model in the Northern Hemisphere.

4.3.2 Comparison of land/sea masks used in CTE, CAMS and Jena

The land/sea masks computed in this study was compared to the ones the three model computes as well. Given that the difference of land/sea masks on two earth models are tiny, only land/sea mask on the spherical earth model was used for comparison.

Land/sea mask created by CTE model is based on TransCom regions. The green part in Figure 4.3 is the undefined region in TransCom regions, which is shown as 'sea' on the CTE land/sea mask. However, the difference between two land/sea masks are not apparent over other regions.

landSeaMask_Spherical-CTE

Fig 4.3 Comparison between CTE land/sea mask and the one computed on the spherical earth model. Color scale presents the difference of fraction of land between two land masks.

The largest difference appears in Antarctic between CAMS land/sea mask and the spherical one (Figure 4.4), which results from the different definition of Antarctic in two geographic data used for the two estimates. The purple red part is defined as sea in CAMS system while it is a part of land of Antarctic in the vector data collected from Esri. Unlike the CTE system, the land/sea mask of CAMS shows difference from the land/sea mask on the spherical model mainly along the coastal. CAMS estimates lower land fraction but higher ocean fraction along the coastal than the mask computed on the spherical earth model. But for Caspian Sea and other small lakes, higher fraction of land around borders of lakes shown in the land/sea mask by spherical earth model than CAMS land/sea mask. This could because, for one thing, the definition of small lake extent in CAMS is smaller than the lake extent in collected Esri vector data and, for another thing, $1^\circ x 1^\circ$ resolution is too coarse to provide precise estimation of land/sea fraction for the small areas.

landSeaMask_Spherical-CTE

Fig 4.4 Comparison between CAMS land/sea mask and the one computed on the spherical earth model. Color scale presents the difference of fraction of land between two land masks.

Similar to the comparison result between CAMS and the mask computed on the spherical earth model, Jena provides lower land fraction and higher ocean fraction in coastal regions (Figure 4.5) and relative larger difference appears in small lakes. But this phenomenon is more obvious for Jena system and appears almost along the whole coastal lines, but only in some coastal areas in CAMS system. Light colors appearing in small part of Antarctic illustrate Jena system has similar geographical definition to the vector data used in this study.
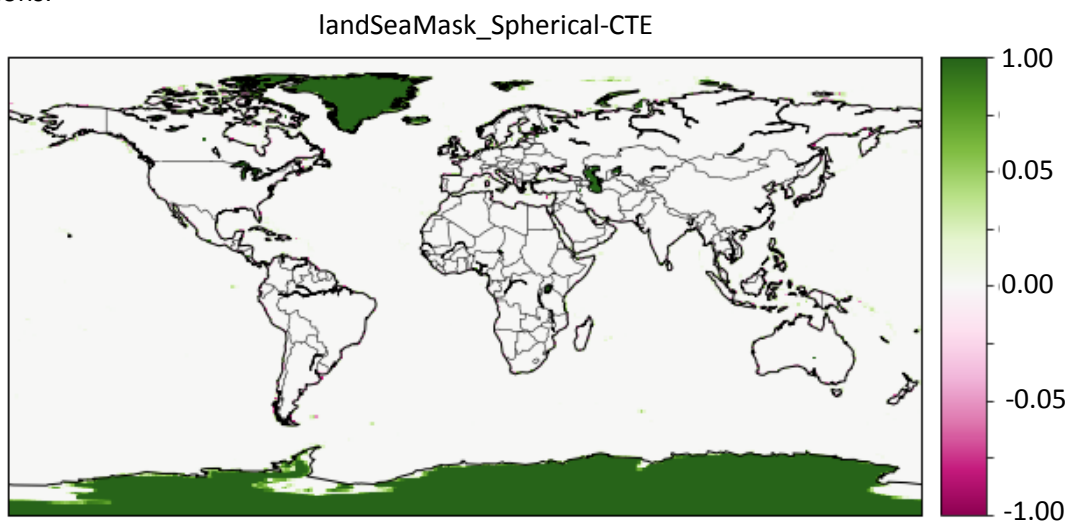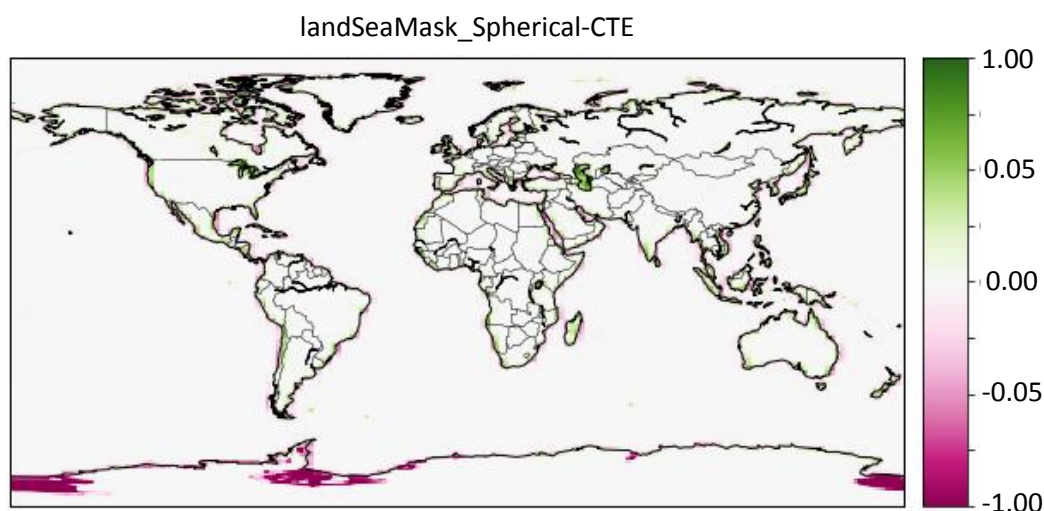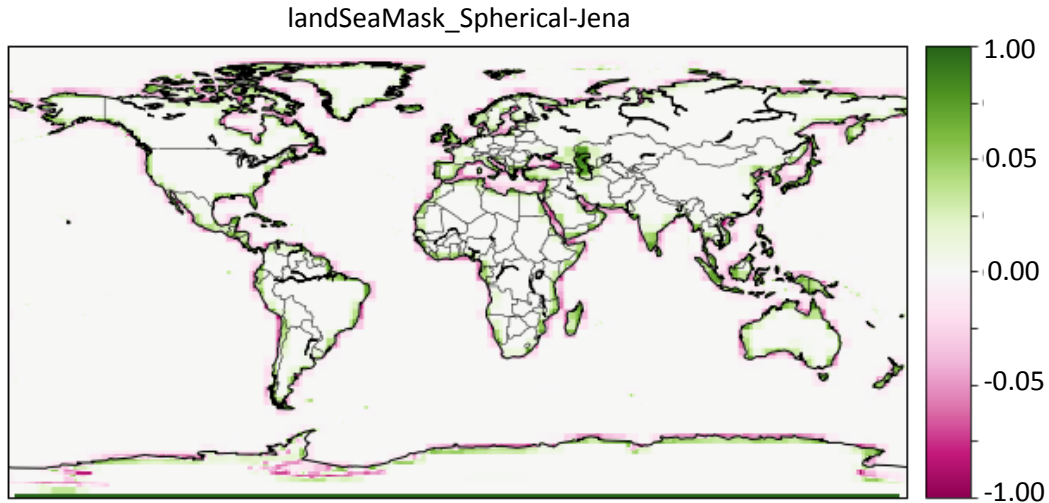
Fig 4.5 Comparison between Jena land/sea mask and the one computed on the spherical earth model. Color scale presents the difference of fraction of land between two land masks.

5 Case study: estimation of carbon flow in Sweden

5.1 Background

Methodology to estimate area of a polygon based on a grid was applied to estimate carbon flow in Sweden in this chapter. Swedish carbon flow was estimated by cylindrical equal-area projection method on different earth models using different inversion flux data. Based on the estimation of overlay area in grid cells, Sweden carbon flow can be computed as sum of fluxes in grid cells which are overlapped by polygon defining Sweden. In this thesis, flux data is assumed homogeneous in each grid cell. So the equation to estimate Sweden carbon flow is:

$$Sweden\ carbon\ flow = \sum_1^n flux_i \times overlay\ area_i \tag{9}$$

where n is number of grid cells overlapped by Sweden, $flux_i$ is the mean flux in grid cell i and $Overlay\ area_i$ is the area of overlay region of Sweden in grid cell i. According to equation 2, equation 9 can be transformed as:

$$Sweden\ carbon\ flow = \sum_1^n flux_i \times partial\ value_i \times cell\ area_i \tag{10}$$

Which can also be written as:

$$Sweden\ carbon\ flow = flux\ data\ \times Sweden\ mask \times cell\ areas \tag{11}$$

where flux data is in various grids depending on which inversion system is used. Sweden mask presents partial value of Sweden area in each grid cell, which has the same size as flux data.

According to equations above, different components, partial value, earth model and inversion system could affect the estimation result of regional carbon flow. To investigate uncertainties propagated by these components, comparison of Sweden carbon flow computed by different combination of the components was made.

5.2 Data and data processing

5.2.1 Flux data

Carbon flux data were collected from results of three inversion systems: CTE, CAMS and Jena CarboScope. bio-spheric flux data from three inversions is monthly but in different units: mol/m^2/second for CTE, kgC/m^2/month for CAMS and PgC/grid cell/year for Jena. The time period of flux data also vary. CTE flux data is available during 2001-2014, while longer time period for CAMS and Jena data: 1979-2015 for CAMS and 1980-2016 for Jena (Table 5.1).

When comparing Sweden carbon flow computed by different flux data in different resolutions, common unit and resolution of flux data is required to facilitate a more direct comparison (Peylin et al. 2013). The units of three flux data were converted to a common one: PgC/m^2/year and flux data of CAMS and Jena were resampled to CTE grid (1°x 1°) by first-order conservative remapping. Interpolation three flux data onto the same grid helps to get rid of the influence of different resolution on the final result and the effects of main differences (i.e., transport model, prior information, atmospheric measurements and optimization scheme) between inversions can be shown directly in the comparison of final result of the estimation.

Table 5.1 Flux data of three inversions

Inversion system	Grid	Unit of monthly flux data	Time period
CTE	180x360	mol/m2/s	2001-2014
CAMS	48x72	kgC/m^2/month	1979-2015
Jena	96x96	PgC/grid cell/year	1980-2016

5.2.2 Sweden masks

There are 7 Sweden masks created for the following test. Six of them were computed in three grids (CTE, CAMS and Jena) and on two earth models by cylindrical equal-area projection method. Another one Sweden mask was created by planar method.

Partial values of Sweden area in each grid cell were calculated in ArcGIS and masks were created by python programming. Vector data defining Sweden extent was extracted from the global vector data used in section 4.2.1. 'Sweden' polygon was intersected by three inversion grids on two earth models to compute partial values in grid cells by tool 'Intersect'. This tool calculated overlay area of 'Sweden' polygon within grid cells and partial values were calculated by dividing overlay areas by cell areas. Another Sweden mask was created by planar method in CTE grid on GRS80 ellipsoidal earth model. There is no difference between spherical and ellipsoidal earth model based on which to compute land/sea mask or region mask by planar method. There are 7 Sweden masks in total:
- CTE/spherical earth model

- CTE /ellipsoidal earth model
- CAMS/spherical earth model
- CAMS /ellipsoidal earth model
- Jena/spherical earth model
- Jena /ellipsoidal earth model
- CTE/planar method.

5.2.3 Cell areas

Six cell areas were also calculated in three grids and on two earth models. CTE, CAMS and Jena grid were created by 'create fishnet' tool in cylindrical equal area projection coordinate system based on the spherical earth model and GRS80 ellipsoidal earth model in ArcGIS. Every grid cell was stored as a polygon and area of each one could be found in the 'attribute table'. Cell areas stored in grids with different sizes were created by Python programming. The six cell areas are:

- CTE/spherical earth model
- CTE /ellipsoidal earth model
- CAMS/spherical earth model
- CAMS /ellipsoidal earth model
- Jena/spherical earth model
- Jena /ellipsoidal earth model.

5.3 Method

5.3.1 Test design

The uncertainties of carbon flow estimation can come from these components: partial value, earth model, and inversion system. But interpolation also introduces uncertainties to carbon flow estimation. To evaluate the effects of this methodology on the application of estimating regional carbon flow, the effects of four components were investigated: partial value, earth model, inversion system and interpolation. The tests were performed using Python scripts in Jupyter notebook (see Appendix).

4 groups of test were performed (Table 5.2). In each group, different results of Sweden carbon flow in January 2001 were estimated and compared. The component to be tested was changed within a couple of options and other components were kept constant. Effects of different factors are presented by relative difference between various results of Sweden carbon flow in each group.

The effects of partial value and earth model do not differ at different time. But the effects of regridding and inversion systems may differ as time changes because flux data is changing over time. In the test of effect of interpolation, time series of original and resampled flux data were

plotted to see how steadily the difference between two time series changes over time, which can indicate how the effects differ over time. And in the test of effect of inverse system, estimation of Sweden carbon flow in June 2001 was complemented in order to compare different results at different time.

Table 5.2 Test design to investigate effects of different components on Sweden carbon flow.

Test	Flux data		Sweden mask	Cell areas (earth model)
Test effect of partial value	CTE		Planar method in CTE grid	Equal-area projection on a spherical model in CTE grid
	CTE		Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
	CTE		Equal-area projection on GRS80 ellipsoidal model in CTE grid	Equal-area projection on a spherical model in CTE grid
Test effect of earth model	CTE		Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
	CTE		Equal-area projection on a spherical model in CTE grid	Equal-area projection on GRS80 ellipsoidal model in CTE grid
Test effect of interpolation	CAMS	CAMS in original grid	Equal-area projection on a spherical model in CAMS grid	Equal-area projection on a spherical model in CAMS grid
		CAMS in interpolated grid	Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
	Jena	Jena in original grid	Equal-area projection on a spherical model in Jena grid	Equal-area projection on a spherical model in Jena grid
		Jena in interpolated grid	Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
Test effect of inversion system	CTE		Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
	CAMS in interpolated grid		Equal-area projection on a spherical model in CTE grid	Equal-area projection on a spherical model in CTE grid
	Jena in interpolated grid		Equal-area projection on a spherical model	Equal-area projection on a spherical model

		in CTE grid	in CTE grid

5.4 Results

The following tables and figures show results of 4 groups of test.

5.4.1 Effect of partial value

Table 5.3 Effect of partial value. Relative difference to CTE/planar mask/spherical cell size is the difference between other two results and CTE/planar mask/spherical cell size divided by CTE/planar mask/spherical cell size. The relative difference to CTE/spherical EA partial value/spherical cell size and the relative difference to CTE/ellipsoidal EA partial value/spherical cell size were calculated in the same manner.

Test	Sweden carbon flow(PgC/yr)	Relative difference(%)		
		To CTE/planar mask/spherical cell size	To CTE/spherical EA partial value/spherical cell size	To CTE/ellipsoidal EA partial value/spherical cell size
CTE/planar mask/spherical cell size	0.1459512		-0.0052	-0.0053
CTE/spherical EA partial value/spherical cell size	0.14595875	0.0052		-6.8184E-05
CTE/ellipsoidal EA partial value/spherical cell size	0.14595885	0.0053	6.8184E-05	

5.4.2 Effect of earth model

Table 5.4 Effect of earth model. The relative difference is the difference between CTE/spherical EA partial value/spherical cell size and CTE/spherical EA partial value/ellipsoidal cell size divided by CTE/spherical EA partial value/spherical cell size.

	Sweden carbon flow(PgC/yr)	Relative difference(%)
CTE/spherical EA partial value/spherical cell size	0.14595875035	
CTE/spherical EA partial value/ellipsoidal cell size	0.146814324746	0.586175473

5.4.3 Effect of interpolation

5.4.3.1 CAMS

Table 5.5 Effect of CAMS interpolation. The relative difference is the difference between CAMS original/spherical EA partial value/spherical cell size and CAMS regridded/spherical EA partial value/spherical cell size divided by CAMS original/spherical EA partial value/spherical cell size.

	Sweden carbon flow(PgC/yr)	Relative difference(%)
CAMS original/spherical EA partial value/spherical cell size	0.0846	
		-0.9817
CAMS regridded/spherical EA partial value/spherical cell size	0.0838	
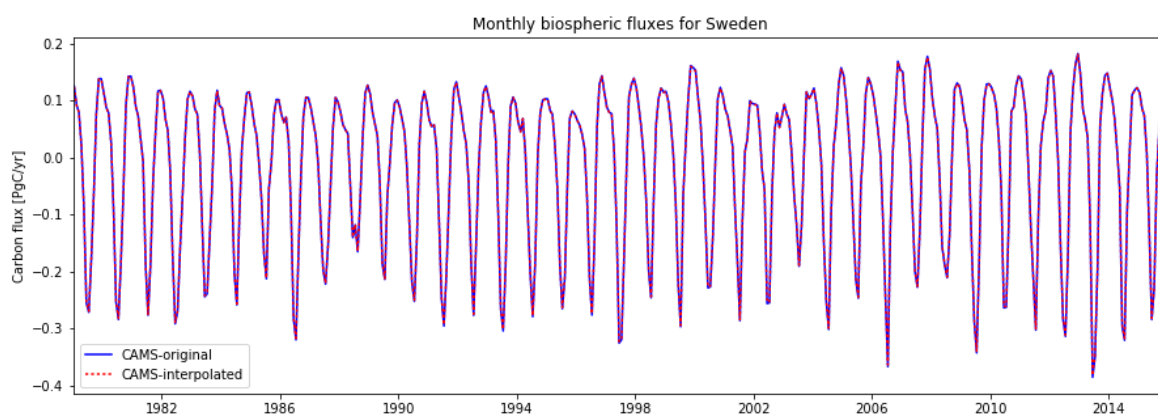
Figure 5.1 Time series 2001-2014 of CAMS original and interpolated Sweden carbon flow

5.3.4.2 Jena CarboScope

Table 5.6 Effect of Jena interpolation. The relative difference is the difference between Jena original/spherical EA partial value/spherical cell size and Jena regridded/spherical EA partial value/spherical cell size divided by Jena original/spherical EA partial value/spherical cell size.

	Sweden carbon flow(PgC/yr)	Relative difference(%)
Jena original/spherical EA partial value/spherical cell size	0.0814	
		-1.1119
Jena regridded/spherical EA partial value/spherical cell size	0.0805	
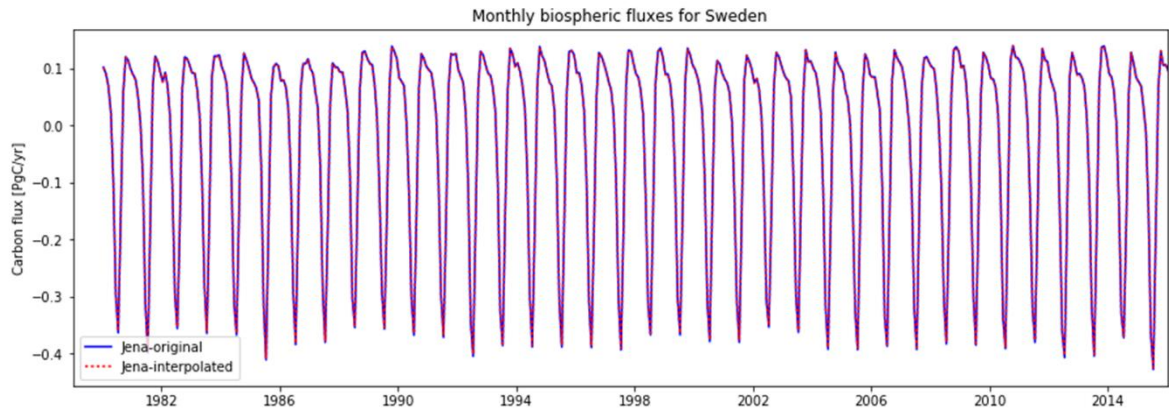
Figure 5.2 Time series 2001-2014 of Jena original and interpolated Sweden carbon flow

5.4.4 Effect of inversion system

Table 5.7 Effect of inversion system. a) Sweden carbon flow in January 2001. The relative difference to CTE result is the difference between CAMS result and CTE result divided by CTE result. b) Sweden carbon flow in June 2001. The relative difference to CTE result is the difference between Jena result and CTE result divided by CTE result

a)

Sweden carbon flow(PgC/yr)		Relative difference to CTE result(%)
CTE/spherical EA partial value/spherical cell size	0.1460	
CAMS regridded/spherical EA partial value/spherical cell size	0.0838	-42.5936
Jena regridded/spherical EA partial value/spherical cell size	0.0805	-44.8751

b)

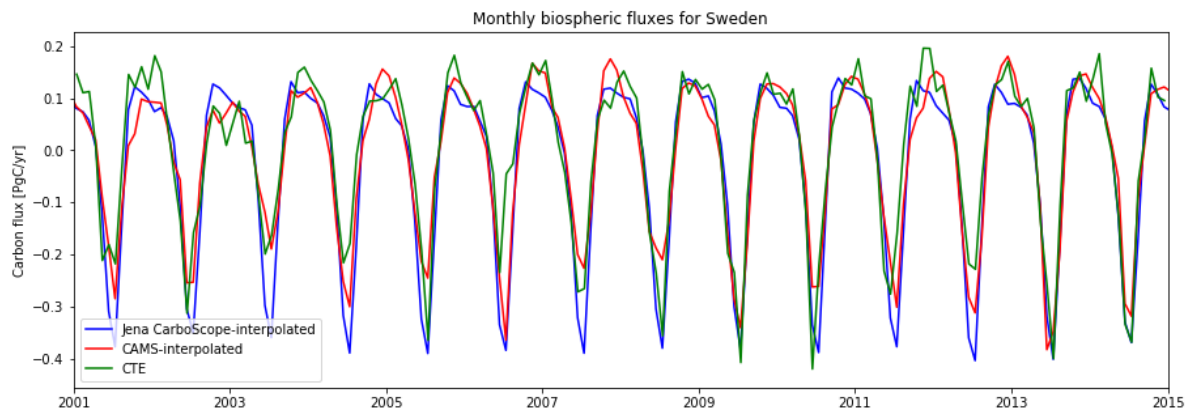| Sweden carbon flow(PgC/yr)		Relative difference to CTE result(%)
CTE/spherical EA partial value/spherical cell size	-0.1809	
CAMS regridded/spherical EA partial value/spherical cell size	-0.2274	25.6972
Jena regridded/spherical EA partial value/spherical cell size	-0.3043	68.1954

Figure 5.3 Time series 2001-2014 of Sweden carbon flow estimated by three inversion flux data.

5.4.5 Analyses of results

Partial value has very small influence on the final result. Less than 0.01% relative difference between Sweden flows computed by different Sweden masks. Sweden flows are almost the same computed by partial value on the spherical earth model and GRS80 ellipsoidal earth model. Even if partial value is estimated by planar method, there is only around 0.005% relative difference to other two results in the first test.

Influence of earth model used to estimate cell areas are larger than partial value but still less than 1%. In the case of Sweden, estimation of carbon flow on GRS80 earth model is 0.6% higher than the spherical earth model.

Less than 1% uncertainty was introduced by first-order conservative remapping method in test of effect of interpolation. And according to time series for both CAMS and Jena, Original flux data and interpolated data match quite well all the time. The difference between two time series for each inversion is pretty small and almost the same all the time.

The final result is highly affected by the choice of inversion system. The relative difference varies from 20% to 70%. CAMS and Jena inversion underestimated around 50% carbon flow compared to CTE inversion in winter (January 2001), while CAMS overestimate around 25% carbon flow and Jena overestimate more than 50% than CTE in summer (June 2001). This means the uncertainties of regional carbon flow vary over time, which is also reflected by the time series of Sweden carbon flow estimated by three flux data. The difference between every two results is changing all the time, usually the difference is smaller in winter compared to summer. The final results differ due to the variance of flux data involved in the estimation. The difference of modeled flux data is contributed by different factors comprising an inversion system: transport model, natural flux, atmospheric measurements and optimization scheme.

To conclude, the uncertainties from area estimation introduced by the projection method to the final result is less than 1%, the most uncertainties come from the flux data modeled by

different inversions. Compared to the uncertainties from flux data, the uncertainties in polygon area estimation by the projection method is too small to worry about.

6 Discussion

6.1 Methodology to estimate partial value and cell area

Three methods to estimate area of polygons were proposed in this thesis. The lat/lon planar method is the easiest one to compute a spherical polygon's area, but obvious shortcomings can be seen. The spherical method theoretically gives the most correct results, but it is not easy to implement because of its big complexity. The projection method is the most appropriate method amongst the three methods because it trades off between the quality of estimates and computational demands.

Area of a polygon consists of two components: partial value and cell area. This problem can boil down to estimate these two components. Uncertainties will be introduced to partial values but not to cell areas by the projection method. Because boundaries of polygon within a grid cell are great circle arcs on a sphere/ellipsoid surface but they are straight lines on a plane which will make shape of polygon on the plane inconsistent with the shape on the Earth's surface. However, for a grid cell, boundaries on the projected plane overlap its original boundaries on the Earth's surface (Figure 6.1).

For the polygon within a grid cell, errors of area will be smaller if break lines of the polygon are shorter. In figure 6.1, arc AB will be projected to red lines on the plane if break line AB is divided into two segments AA', A'B. Then errors of the triangle area will be reduced. On the spherical/ellipsoidal surface, shorter or smaller great circle arcs are closer to straight lines.
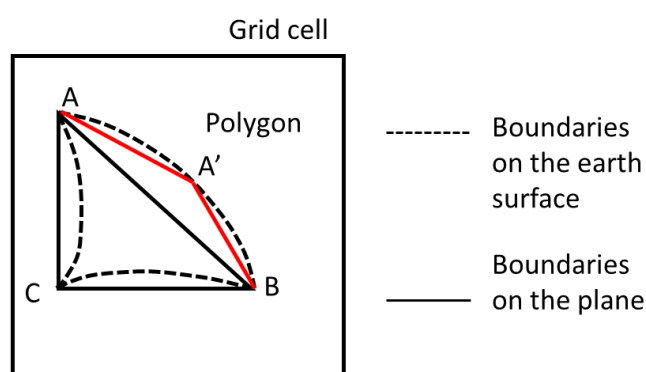

Fig 6.1 An example showing a polygon in a grid cell. Polygon ABC is a triangle on a sphere/ellipsoid surface. Dashed lines are boundaries on a sphere/ellipsoid surface, which are great circle arcs. Solid lines are boundaries of polygon and the grid cell on a plane. Dashed lines and solid lines of the grid cell boundaries overlap each other.

6.2 Application of methodology: regional carbon flow estimation

6.2.1 Problem of region flow estimation

In this thesis, inversion provides land flux data and ocean flux data separately. And flux data is mean flux for the whole grid cell, which assumes flux is constant at every point in a grid cell. For a grid cell including both land and ocean part (Figure 6.2 (a)), the land flux is considered homogenous in the whole cell including ocean part for which actually there is no land flux. bio-spheric carbon flow in a grid cell like that should be estimated by the equation $\text{flow} = \text{flux} \times \text{cell area}$ but not the equation $\text{flow} = \text{flux} \times \text{partial value} \times \text{cell area}$. Additionally, interpolation from coarse grid to fine grid automatically assign values of land flux to fine grid cells containing ocean (Figure 6.2). This problem will bring uncertainties to interpolated land flux data, conversely for ocean flux data. The flow in a grid cell including both land and sea is underestimated due to this problem.

If the grid cell where one part of land is to the border of land from other region (Figure 6.2 (b)), however, there is no problem for region flow estimation by this equation, $\text{flow} = \text{flux} \times \text{partial value} \times \text{cell area}$.

The underestimation of flow in the cells where land border to ocean leads to underestimation of carbon flow in the whole region. For further study to get more precise estimation of regional carbon flows, estimation of flow in a grid cell should be performed in two different ways depending on the grid cell. The flow should be estimated by the equation $\text{flow} = \text{flux} \times \text{cell area}$ for the grid cells including both land and ocean, but $\text{flow} = \text{flux} \times \text{partial value} \times \text{cell area}$ for grid cells only containing land.
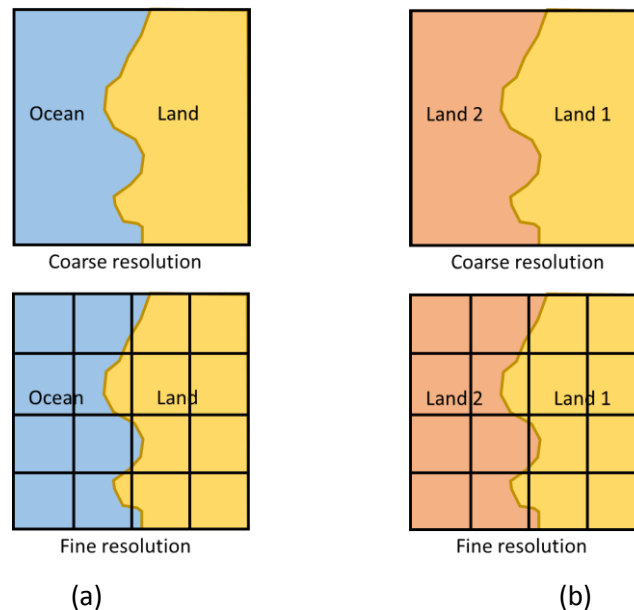

Fig 6.2 Examples of grid cells with coarse and fine resolution. (a) A grid cell including both land and ocean. Land border is to the ocean. (b) A grid cell including land of different regions. Land 1 border is to land 2.

6.2.2 Application of area estimation in other domains

Besides the application of regional carbon flow estimation, the equal-area projection method to estimate area of polygon can be applied to other domains, such as precipitation estimation, quantity of species estimation and water flux estimation in a given region. Estimating region area is only one key step of the whole procedure of estimation. For all the problems to estimate sum of certain geographic data in a region, the estimation result is mainly affected by the quality of field data stored in grids. Estimating region area by the projection method could only propagate small uncertainties (less than 1%) to the final result.

Equal-area projection method is a method which can be employed easily and can provide relative high quality estimation of area for the application to estimate regional carbon flow. There are other methods providing more accurate estimation but more difficult to be applied and less efficient, e.g., the spherical method proposed in this thesis. So if accuracy of estimates is the most important purpose, there could be better method than projection method. And the thesis demonstrates equal-area projection method is good enough only for the application of regional carbon flow estimation. The most suitable method to estimate area of a polygon could vary depending on applications or domains.

7 Conclusions

This thesis aims to develop a method to estimate area of a polygon on spherical and ellipsoidal earth models and employs this method to estimate regional carbon flow. Region area consists of two components: cell area and partial value. The problem to estimate region area is transformed to estimate cell area and partial value. Three methodologies to estimate partial value have been proposed and tested: 1) using latitude-longitude plane, 2) using cylindrical area-preserving projection and 3) using the area of corresponding of spherical polygons. Spherical method provides true values, but it is often too complicated to be applied into practice. Triangle test shows that partial values computed by planar method and projection method are pretty close to true values when length of the triangle's edge is less than 0.5 degree and the results of projection method are closer to true values compared the results of planar method. Hence, projection method is the most suitable method amongst three methods proposed to estimate partial value because it compromises between precision and computational demands. Area-preserving projection is also a good choice for cell area estimation because it gives true values. But the result of cell area estimation is influenced by different earth models which depend on how the inversion system is constructed. The relative difference between cell areas computed on two earth models is less than 1%, while land/sea masks computed on spherical and ellipsoidal earth model are almost the same (less than 0.01%). A disadvantage of equal-area projection method to estimate partial values is that great circle arcs become rhumb lines when boundaries of polygons are projected from the spherical/ellipsoidal surface to a plane. This shortcoming brings errors to estimation of area of

polygons. Shorter length of break lines of polygons can reduce errors to estimations.

Area-preserving projection method was employed to compute land/sea masks of vector data of world map and estimate Sweden carbon flow. Uncertainties propagated to regional carbon flow by different effects were investigated by four groups of test. Three inversion systems, CTE, CAMS and Jena CarboScope were chosen for case study. By comparing land/sea masks estimated by the projection method and the masks used for the three inversion systems currently, it is found that the differences mostly appear along coastlines. The tests were performed to evaluate the effects of partial values, earth model, inversion system and interpolation on Sweden carbon flow estimation. Results indicate that partial values, earth model and interpolation affect region flow estimation slightly (less than 1%). Inversion system is the main factor to influence the result of regional carbon flow. The relative difference between Sweden carbon flows estimated using three inversion data differ over time.

A limitation of the estimation of regional carbon flow is that regional carbon flow is underestimated in grid cells containing both ocean and land. For further study, estimation of carbon flow in a grid cell could be improved by treating grid cells where border of region is to sea and grid cells where region border is to another region in different ways in the case that flux data is homogeneous in each grid cell.

References

Alexander, J. 2004. Loxodromes: a rhumb way to go. *Mathematics Magazine* 77: 349-356. JSTOR. doi:10.2307/3219199.

Bevis, M., and G. Cambareri. 1987. Computing the area of a spherical polygon of arbitrary shape. *Mathematical Geology* 19: 335-346. Springer Nature. doi:10.1007/bf00897843.

Barkley, M., P. Monks, U. Frieß, R. Mittermeier, H. Fast, S. Körner, and M. Heimann. 2006. Comparisons between SCIAMACHY atmospheric CO_2 retrieved using (FSI) WFM-DOAS to ground based FTIR data and the TM3 chemistry transport model. *Atmospheric Chemistry and Physics* 6: 4483-4498. Copernicus GmbH. doi:10.5194/acp-6-4483-2006.

Baker, D., R. Law, K. Gurney, P. Rayner, P. Peylin, A. Denning, P. Bousquet, and L. Bruhwiler et al. 2006. TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988-2003. *Global Biogeochemical Cycles* 20. Wiley-Blackwell. doi:10.1029/2004gb002439.

Baek, S., O. Boucher, K. Dassas, J. Escribano, L. Fita, F. Forget, A. Gainusa-Bogdan and L. Guez, et al. 2014. LMDZ5: a documentation. L' équipe de modélisation du Laboratoire de Météorologie Dynamique (LMD).

Bergamaschi, P., M. Corazza, U. Karstens, M. Athanassiadou, R. Thompson, I. Pison, A. Manning, and P. Bousquet et al. 2015. Top-down estimates of European CH_4 and N_2O emissions based on four different inverse models. *Atmospheric Chemistry and Physics* 15: 715-736. Copernicus GmbH. doi:10.5194/acp-15-715-2015.

Chen, H., I. Lee, S. Leopoldseder, H. Pottmann, T. Randrup, and J. Wallner. 1999. On surface approximation using developable surfaces. *Graphical Models and Image Processing* 61:

110-124. Elsevier BV. doi:10.1006/gmip.1999.0487.

Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F. Bréon, A. Chédin, and P. Ciais. 2005. Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *Journal of Geophysical Research* 110. Wiley-Blackwell. doi:10.1029/2005jd006390.

Ciais, P., P. Rayner, F. Chevallier, P. Bousquet, M. Logan, P. Peylin, and M. Ramonet. 2010. Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives. *Greenhouse Gas Inventories*: 69-92. Springer Netherlands. doi:10.1007/978-94-007-1670-4_6.

Chen, J., J. Fung, G. Mo, F. Deng, and T. West. 2015. Atmospheric inversion of surface carbon flux with consideration of the spatial distribution of US crop production and consumption. *Biogeosciences* 12: 323-343. Copernicus GmbH. doi:10.5194/bg-12-323-2015.

Denning, A., M. Holzer, K. Gurney, M. Heimann, R. Law, P. Rayner, I. Fung, and S. Fan et al. 1999. Three-dimensional transport and concentration of SF6. A model intercomparison study (TransCom 2). *Tellus B* 51: 266-297. Informa UK Limited. doi:10.1034/j.1600-0889.1999.00012.x.

Deng, F., J. Chen, M. Ishizawa, C. Yuen, G. Mo, K. Higuchi, D. Chan, and S. Maksyutov. 2007. Global monthly CO_2 flux inversion with a focus over North America. *Tellus B* 59: 179-190. Informa UK Limited. doi:10.1111/j.1600-0889.2006.00235.x.

Engelen, R. 2002. On error estimation in atmospheric CO_2 inversions. *Journal of Geophysical Research* 107. Wiley-Blackwell. doi:10.1029/2002jd002195.

Eldrandaly, K. 2006. A COM-based expert system for selecting the suitable map projection in ArcGIS. *Expert Systems with Applications* 31: 94-100. Elsevier BV. doi:10.1016/j.eswa.2005.09.008.

GCP - Global Carbon Project. 2017. *Globalcarbonproject.org*. http://www.globalcarbonproject.org/reccap/protocol.htm. Accessed January 21.

Gurney, K., R. Law, P. Rayner, and A.S. Denning. 2000. Transcom 3 experimental protocol. Department of atmospheric Science. Colorado State University, USA, paper, NO.707.

Gurney, K., R. Law, A. Denning, P. Rayner, D. Baker, P. Bousquet, L. Bruhwiler, and Y. Chen et al. 2003. TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. *Tellus B* 55: 555-579. Informa UK Limited. doi:10.1034/j.1600-0889.2003.00049.x.

Heimann, M. and S. Körner. 2003. The global atmospheric tracer model tm3, Tech. Rep. 5 Rep., Max Planck Institute for Bio- geochemistry (MPI-BGC), Jena, Germany, pp. 131.

Hourdin, F., I. Musat, S. Bony, P. Braconnot, F. Codron, J. Dufresne, L. Fairhead, and M. Filiberti et al. 2006. The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics* 27: 787-813. Springer Nature. doi:10.1007/s00382-006-0158-0.

Jones, P. 1999. First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Monthly Weather Review* 127: 2204-2210. American Meteorological Society. doi:10.1175/1520-0493(1999)127<2204:fasocr>2.0.co;2.

Kennedy, M. and S. Kopp. 1994. Understanding map projections. GIS by ESRI, Environmental System Research Institute, Redlands, USA.

Krol, M., S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters,

and F. Dentener et al. 2005. The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chemistry and Physics* 5: 417-432. Copernicus GmbH. doi: 10.5194/acp-5-417-2005.

Kirschke, S., P. Bousquet, P. Ciais, M. Saunois, J. Canadell, E. Dlugokencky, P. Bergamaschi, and D. Bergmann et al. 2013. Three decades of global methane sources and sinks. *Nature Geoscience* 6: 813-823. Springer Nature. doi: 10.1038/ngeo1955.

Lam, N. 1983. Spatial Interpolation Methods: A Review. *The American Cartographer* 10: 129-150. Informa UK Limited. doi:10.1559/152304083783914958.

Li, J., and A. Heap. 2014. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software* 53: 173-189. Elsevier BV. doi:10.1016/j.envsoft.2013.12.008.

Law, R., P. Rayner, A. Denning, D. Erickson, I. Fung, M. Heimann, S. Piper, and M. Ramonet et al. 1996. Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions. *Global Biogeochemical Cycles* 10: 783-796. Wiley-Blackwell. doi:10.1029/96gb01892.

Law, R., Y. Chen, K. Gurney, and T. 3 Modellers. 2003. TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices. *Tellus B* 55: 580-595. Informa UK Limited. doi:10.1034/j.1600-0889.2003.00053.x.

Lu, Z., S. Qiao, and Y. Qu. 2014. *Geodesy*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Le Quéré, C., R. Andrew, J. Canadell, S. Sitch, J. Korsbakken, G. Peters, A. Manning, and T. Boden et al. 2016. Global carbon budget 2016. *Earth System Science Data* 8: 605-649. Copernicus GmbH. doi:10.5194/essd-8-605-2016.

Moritz, H. 1980. Geodetic reference system 1980. *Bulletin Géodésique* 54: 395-405. Springer Nature. doi:10.1007/bf02521480.

Mitchell, J. 1989. The "Greenhouse" effect and climate change. *Reviews of Geophysics* 27: 115. Wiley-Blackwell. doi:10.1029/rg027i001p00115.

Michaut, C. 2017. Atmospheric Tracer Transport Model Intercomparison Project (TransCom). *Wcrp-climate.org*. https://www.wcrp-climate.org/modelling-wgcm-mip-catalogue/modelling-wgcm-mips-2/241-modelling-wgcm-catalogue-transcom. Accessed March 24.

Mitas, L. and H. Mitasova. 1999. Spatial interpolation. In: Longley, P., Good-child, M., Maguire, D., Rhind, D., editors, *Geographical Information Systems: Principles, Techniques, Management and Applications*, vol. 1. Wiley, London, pp. 481 – 492

Oliver, M. and R. Webster. 1990. Kriging: a method of interpolation for geographical information systems. *International journal of geographical information systems* 4: 313-332. Informa UK Limited. doi:10.1080/02693799008941549.

Patra, P., S. Maksyutov, and T. Modelers. 2003. Sensitivity of optimal extension of CO2 observation networks to model transport. *Tellus B* 55: 498-511. Informa UK Limited. doi:10.1034/j.1600-0889.2003.00056.x.

Peters, W., M. Krol, G. Van Der Werf, S. Houweling, C. Jones, J. Hughes, K. Schaefer, and K. Masarie et al. 2010. Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations. *Global Change Biology* 16: 1317-1337. Wiley-Blackwell. doi:10.1111/j.1365-2486.2009.02078.x.

Peylin, P., R. Law, K. Gurney, F. Chevallier, A. Jacobson, T. Maki, Y. Niwa, and P. Patra et al. 2013.

Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions. *Biogeosciences* 10: 6699-6720. Copernicus GmbH. doi:10.5194/bg-10-6699-2013.

Rickey, V. F., and P. M. Tuchinsky. 1980. An application of geography to mathematics: History of the integral of the secant. *Mathematics Magazine* 53: 162-166. JSTOR. doi:10.2307/2690106

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann. 2003. CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics* 3: 1919-1964. Copernicus GmbH. doi:10.5194/acp-3-1919-2003.

Snyder, J. 1987. *Map projections - a working manual*. U.S. Geological Survey Professional Paper 1395. Supersedes USGS Bulletin 1532. Washington, D.C.: U.S. Government Printing Office.

Schneider, S. 1989. The greenhouse effect: science and policy. *Science* 243: 771-781. American Association for the Advancement of Science (AAAS). doi:10.1126/science.243.4892.771.

Stephens, B., K. Gurney, P. Tans, C. Sweeney, W. Peters, L. Bruhwiler, P. Ciais, and M. Ramonet et al. 2007. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2. *Science* 316: 1732-1735. American Association for the Advancement of Science (AAAS). doi:10.1126/science.1137004.

Tobler, W. 1964. Some new equal area map projections. *Survey Review* 17: 240-243. Maney Publishing. doi:10.1179/003962664792000934.

The TransCom Experiment. 2017. *Transcom.project.asu.edu*. http://transcom.project.asu.edu/transcom.php. Accessed March 24.

Thompson, R., K. Ishijima, E. Saikawa, M. Corazza, U. Karstens, P. Patra, P. Bergamaschi, and F. Chevallier et al. 2014. TransCom N_2O model inter-comparison – Part 2: Atmospheric inversion estimates of N_2O emissions. *Atmospheric Chemistry and Physics* 14: 6177-6194. Copernicus GmbH. doi:10.5194/acp-14-6177-2014.

van Sickle, J. 2010. *Basic GIS coordinates*. CRC Press, Florida, USA.

Weintrit, A. 2015. The need to develop new geodesic-based computational algorithms for marine navigation electronic devices and systems. *International Association of Institutes of Navigation World Congress (IAIN)*. IEEE. doi:10.1109/iain.2015.7352263.

Worboys, M., and M. Duckham. 2004. *GIS: a computing perspective*. 2nd ed. CRC Press, Florida, USA.

Affiliation

Ute Karstens: ICOS Carbon Portal, Lund University
Lars Harrie: ICOS Carbon Portal, Lund University

Appendix

1. Codes for Sweden carbon flow estimation

Import modules

```
#Importing modules
%pylab inline
import warnings
import netCDF4 as cdf
import numpy as np
import datetime as dt
import os
import pandas as pd
import PySCRIP as scrip
from PySCRIP.config import PySCRIPConfig
from ipywidgets import widgets
from mpl_toolkits.basemap import Basemap
from matplotlib import pylab as plt
import pyshtools
import csv


#import mpld3
#mpld3.enable_notebook()
warnings.filterwarnings('ignore',category=DeprecationWarning)
```

Import modeled carbon flux in grid from three inversions

```
#Import modeled data from three inversions:CTE,CAMS and Jena

#select versions, by defining filenames
filename_CT='/opt/data/modeledData/CTE2015_monthly.nc'

filename_CAMS='/opt/data/modeledData/CAMS_v15r2_allmonths.nc'
filename_CAMS_regridded='/opt/data/modeledData/CAMS_v15r2_allmonths.CTEgrid_conservative.nc'

filename_Jena='/opt/data/modeledData/Jena_s81_v3.8_monthly.nc'
filename_Jena_regridded='/opt/data/modeledData/Jena_s81_v3.8_monthly.fluxes_perm2.CTEgrid_conservative.nc'
filename_Jena_prior='/opt/data/modeledData/Jena_prior_v3.8_monthly.nc'

#unit conversion factors
fac1=12./1.e12 #for CAMS (data in kgC/m2/month to PgC/yr/m2)
fac2=86400.*365.*12./1.e15 #for CTE (data in mol/m2/s to PgC/yr/m2)
fac3=(1.e15)/365 #PgC/yr -> gC/day
#For Jena: PgC/grid cell/yr -> PgC/m2/yr

# Read from the NetCDF file - the data are stored South to North
f_CTE = cdf.Dataset(filename_CT)
f_CAMS = cdf.Dataset(filename_CAMS)
f_CAMS_regridded = cdf.Dataset(filename_CAMS_regridded)
f_Jena = cdf.Dataset(filename_Jena)
f_Jena_regridded = cdf.Dataset(filename_Jena_regridded)
#f_Jena_prior = cdf.Dataset(filename_Jena_prior)
```

```python
# area[0,:,:] contains (partial) land area of grid cells
# area[1,:,:] contains (partial) ocean area of grid cells
# area[2,:,:] contains total area of grid cells
# using the area of the whole grid cell, i.e. area[2,:,:]

cell_area_Jena = f_Jena.variables['area'][2,:,:]
for i in range(72):
    cell_area_Jena[0][i] = cell_area_Jena[0][0]/72
    cell_area_Jena[47][i] = cell_area_Jena[47][0]/72
#print land_area_Jena

#just retrieve modeled data for time=2001/1
bio_poste_CTE = f_CTE.variables['bio_flux_opt'][0,:,:]*fac2
bio_poste_CAMS = f_CAMS.variables['flux_apos_bio'][264,:,:]*fac1
bio_poste_CAMS_regridded_f = f_CAMS_regridded.variables['flux_apos_bio'][264,:,:]*fac1

bio_poste_Jena=f_Jena.variables['co2flux_land'][252,:,:]/cell_area_Jena #unit:PgC/yr/per cell area
bio_poste_Jena_regridded_f = f_Jena_regridded.variables['co2flux_land'][252,:,:]#unit:PgC/yr/m2


print 'CTE'
print bio_poste_CTE.shape
#print bio_poste_CTE.sum()
print 'CAMS'
print bio_poste_CAMS.shape
#print bio_poste_CAMS.sum()
#print bio_poste_CAMS_regridded.shape
print 'Jena'
print bio_poste_Jena.shape
#print bio_poste_Jena.sum()
```

Define CTE, CAMS and Jena grid

```python
import PySCRIP as scrip
from PySCRIP.config import PySCRIPConfig
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap
import numpy as np

#Define a CTE 1x1 grid(180x360) by creating midpoints of cells
lonCTE = linspace(-179.5, 179.5, 360)
latCTE = linspace(-89.5, 89.5, 180)
lonVectorCTE, latVectorCTE = meshgrid(lonCTE, latCTE)

#Define CAMS grid (96x96) by creating midpoint of cells
dLatCAM=180.0/95.0
latStartCAM = -90.0
dLonCAM = 360.0/96.0
lonStartCAM = -180.0
latCAMS=np.zeros(96)
lonCAMS=np.zeros(96)
# Fix longitude vector
for j in range (0,96):
    lonCAMS[j] =  lonStartCAM + j * dLonCAM
latCAMS[0] = -90+dLatCAM/4
latCAMS[95] = 90-dLatCAM/4
for i in range(1, 95):
    latCAMS[i] = -90+dLatCAM*i
lonVectorCAMS, latVectorCAMS = np.meshgrid(lonCAMS, latCAMS)
```

```
#Define Jena grid (48x72) by creating midpoint of cells
latJena=np.zeros(48)
lonJena=np.zeros(72)
dlatJena = 180.0 / 47.0
latJena[0]=-89.0425
latJena[47]=89.0425
for i in range(0, 46):
    latJena[i+1] = -86.17021 + dlatJena*i
dlonJena = 360.0 / 72.0
for j in range (0,72):
    lonJena[j]=-180 + dlonJena*j
lonVectorJena, latVectorJena = np.meshgrid(lonJena, latJena)
```

Interpolate CAMS flux data from CAMS grid to CTE (1°X1°) grid

```
# Interpolate CAMS flux data from its original grid to a common grid CTE (1X1) grid to facilitate an direct comparison
import PySCRIP as scrip
from PySCRIP.config import PySCRIPConfig
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap
import numpy as np

a = PySCRIPConfig("/opt/data/interpolation/pyscrip_ICOS_configuration.yaml")
#a = PySCRIPConfig("/opt/data/interpolation/pyscrip_test_configuration.yaml")

# Retrieve the file that contains the mapping (from the YAML-file)
simulation = 'ICOS'
maptype = 'conservative'
g1  = 'CAM'
g2  = 'CTE'
theMap = a.getmap(simulation, maptype, g1, g2)
print(theMap.fname)

# Perform the mapping transformation
# Here first order conservative remapping is conducted (indicated by src_grad1=None)
bio_poste_CAMS_regridded = scrip.remap(bio_poste_CAMS, theMap.fname, 'scrip', src_grad1=None)

print bio_poste_CAMS_regridded.shape
#print bio_poste_CAMS_regridded.sum()
```

Interpolate Jena flux data from Jena grid to CTE (1°X1°) grid

```
# Interpolate CAMS flux data from its original grid to a common grid CTE (1X1) grid to facilitate an direct comparison
import PySCRIP as scrip
from PySCRIP.config import PySCRIPConfig
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap
import numpy as np

a = PySCRIPConfig("/opt/data/interpolation/pyscrip_ICOS_configuration.yaml")
#a = PySCRIPConfig("/opt/data/interpolation/pyscrip_test_configuration.yaml")

# Retrieve the file that contains the mapping (from the YAML-file)
simulation = 'ICOS'
maptype = 'conservative'
g1  = 'Jena'
g2  = 'CTE'
theMap = a.getmap(simulation, maptype, g1, g2)
print(theMap.fname)

# Perform the mapping transformation
# Here first order conservative remapping is conducted (indicated by src_grad1=None)
bio_poste_Jena_regridded = scrip.remap(bio_poste_Jena, theMap.fname, 'scrip', src_grad1=None)

print bio_poste_Jena_regridded.shape
#print bio_poste_Jena_regridded.sum()
```

Import cell areas computed on the spherical earth model and on GRS80 ellipsoidal earth model

```python
# Import data of cell areas :
# - CTE cell size on the spherical earth model
# - CTE cell size on the ellipsoidal earth model
# - CAMS cell size on the spherical earth model
# - CAMS cell size on the ellipsoidal earth model
# - Jena cell size on the spherical earth model
# - Jena cell size on the ellipsoidal earth model

#df_cell_size_sphe = pd.read_csv('', delimiter=',')
#df_cell_size_ellip = pd.read_csv('', delimiter=',')

df_cell_size_sphe = pd.read_excel('/opt/data/geodata/cell_size_sphere.xlsx',sheetname='Sheet1')
df_cell_size_ellip = pd.read_excel('/opt/data/geodata/cell_size_ellipsoid.xlsx',sheetname='Sheet1')

cell_size_CTE_sphe = np.zeros((180,360))
cell_size_CTE_ellip = np.zeros((180,360))
cell_size_CAMS_sphe = np.zeros((96,96))
cell_size_CAMS_ellip = np.zeros((96,96))
cell_size_Jena_sphe = np.zeros((48,72))
cell_size_Jena_ellip = np.zeros((48,72))
```

```python
k = 0
for i in df_cell_size_sphe['CTE'][0:180]:
    for j in range(360):
        cell_size_CTE_sphe[k][j] = i
    k = k+1

k = 0
for i in df_cell_size_ellip['CTE'][0:180]:
    for j in range(360):
        cell_size_CTE_ellip[k][j] = i
    k = k+1

k = 0
for i in df_cell_size_sphe['CAMS'][0:96]:
    for j in range(96):
        cell_size_CAMS_sphe[k][j] = i
    k = k+1

k = 0
for i in df_cell_size_ellip['CAMS'][0:96]:
    for j in range(96):
        cell_size_CAMS_ellip[k][j] = i
    k = k+1

k = 0
for i in df_cell_size_sphe['Jena'][0:48]:
    for j in range(72):
        cell_size_Jena_sphe[k][j] = i
    k = k+1
```

```python
k = 0
for i in df_cell_size_ellip['Jena'][0:48]:
    for j in range(72):
        cell_size_Jena_ellip[k][j] = i
    k = k+1

print cell_size_CTE_sphe.shape
#print cell_size_CTE_sphe
print cell_size_CTE_ellip.shape
print cell_size_CAMS_sphe.shape
print cell_size_CAMS_ellip.shape
#print cell_size_Jena_sphe
print cell_size_Jena_ellip.shape
```

Import Sweden masks

```
#Sweden masks are stored from North to South, here we should flip them to make them consistent with the flux data
# Part 1:Import Sweden masks computed in CTE(1x1) grid by planar method,spherical equal-area projection and ellipsoidal equal-area projection

file_Swe_mask_1x1_planar = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_planar.csv',"r"), delimiter=',')
file_Swe_mask_1x1_sphe_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_EA_sphe.csv',"r"), delimiter=',')
file_Swe_mask_1x1_ellip_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_EA_ellip.csv',"r"), delimiter=',')


#Sweden mask computed by planar method
Swe_mask_1x1_planar=np.zeros((180,360))
i=0;
for row in file_Swe_mask_1x1_planar:
    i=i+1
    Swe_mask_1x1_planar[180-i,:] = row[0:360]



#Sweden mask computed in CTE grid by spherical equal-area projection
Swe_mask_1x1_sphe_EA=np.zeros((180,360))
i=0;
for row in file_Swe_mask_1x1_sphe_EA:
    i=i+1
    Swe_mask_1x1_sphe_EA[180-i,:] = row[0:360]
#print Swe_mask_1x1_sphe_EA.sum()

#Sweden mask computed by ellipsoidal equal-area projection
Swe_mask_1x1_ellip_EA=np.zeros((180,360))
i=0;
for row in file_Swe_mask_1x1_ellip_EA:
    i=i+1
    Swe_mask_1x1_ellip_EA[180-i,:] = row[0:360]
```

```
# Part 2:Import Sweden masks computed in CAM grid by spherical equal-area projection and ellipsoidal equal-area projection
file_Swe_mask_CAMS_sphe_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_CAMS_sphe_EA.csv',"r"), delimiter=',')
file_Swe_mask_CAMS_ellip_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_CAMS_ellip_EA.csv',"r"), delimiter=',')


#Sweden mask in CAMS grid
Swe_mask_CAMS_sphe_EA=np.zeros((96,96))
i=0;
for row in file_Swe_mask_CAMS_sphe_EA:
    i=i+1
    Swe_mask_CAMS_sphe_EA[96-i,:] = row[0:96]

Swe_mask_CAMS_ellip_EA=np.zeros((96,96))
i=0;
for row in file_Swe_mask_CAMS_ellip_EA:
    i=i+1
    Swe_mask_CAMS_ellip_EA[96-i,:] = row[0:96]
```

```
# Part 3:Import Sweden masks computed in Jena grid by spherical equal-area projection and ellipsoidal equal-area projection
file_Swe_mask_Jena_sphe_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_Jena_sphe_EA.csv',"r"), delimiter=',')
file_Swe_mask_Jena_ellip_EA = csv.reader(open('/opt/data/geodata/countries/Sweden_landSeaMask_Jena_ellip_EA.csv',"r"), delimiter=',')


#Sweden mask in Jena grid
Swe_mask_Jena_sphe_EA=np.zeros((48,72))
i=0;
for row in file_Swe_mask_Jena_sphe_EA:
    i=i+1
    Swe_mask_Jena_sphe_EA[48-i,:] = row[0:72]

Swe_mask_Jena_ellip_EA=np.zeros((48,72))
i=0;
for row in file_Swe_mask_Jena_ellip_EA:
    i=i+1
    Swe_mask_Jena_ellip_EA[48-i,:] = row[0:72]
```

Comparison of Sweden flow computed by different methods

```
# Compare Sweden carbon flows computed by different methods
# test the effect of the partial values,earth model(cell size),interpolation and inversion separatly


#Test-partial value: using CTE modeled data
Swe_CTE_planar = bio_poste_CTE*Swe_mask_1x1_planar*cell_size_CTE_sphe
Swe_CTE_sphe_EA = bio_poste_CTE*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
Swe_CTE_ellip_EA_partial_value = bio_poste_CTE*Swe_mask_1x1_ellip_EA*cell_size_CTE_sphe

#Test-earth model:using CTE modeled data
Swe_CTE_sphe_EA = bio_poste_CTE*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
Swe_CTE_ellip_EA_cell_size = bio_poste_CTE*Swe_mask_1x1_sphe_EA*cell_size_CTE_ellip

#Test-CAMS interpolation
Swe_CAM_ori = bio_poste_CAMS*Swe_mask_CAMS_sphe_EA*cell_size_CAMS_sphe
Swe_CAM_regridded = bio_poste_CAMS_regridded*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe

#Test-Jena interpolation
Swe_Jena_ori = bio_poste_Jena*Swe_mask_Jena_sphe_EA*cell_size_Jena_sphe
Swe_Jena_regridded = bio_poste_Jena_regridded*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe

#Test-inversion
Swe_CTE_sphe_EA = bio_poste_CTE*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
Swe_CAM_regridded = bio_poste_CAMS_regridded*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
Swe_Jena_regridded = bio_poste_Jena_regridded*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
```

```
print Swe_CTE_planar.sum()
print Swe_CTE_sphe_EA.sum()
print Swe_CTE_ellip_EA_partial_value.sum()
print Swe_CTE_ellip_EA_cell_size.sum()
print Swe_CAM_ori.sum()
print Swe_CAM_regridded.sum()
print Swe_Jena_ori.sum()
print Swe_Jena_regridded.sum()

Swe_CAMS_regridded_f = bio_poste_CAMS_regridded_f*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
print '\nRegridded CAMS flux data read from library files:'
print Swe_CAMS_regridded_f.sum()
Swe_Jena_regridded_f = bio_poste_Jena_regridded_f*Swe_mask_1x1_sphe_EA*cell_size_CTE_sphe
print '\nRegridded Jena flux data read from library files:'
print Swe_Jena_regridded_f.sum()
```

Plot time series of Sweden carbon flows

```
# Import flux data from three inverisons: CTE, CAMS and Jena

#Retrieve flux data for all time steps
bio_poste_CTE = f_CTE.variables['bio_flux_opt'][:,:,:]*fac2
bio_poste_CAMS = f_CAMS.variables['flux_apos_bio'][:,:,:]*fac1
bio_poste_Jena=f_Jena.variables['co2flux_land'][:,:,:]/cell_area_Jena
bio_poste_CAMS_regridded_f = f_CAMS_regridded.variables['flux_apos_bio'][:,:,:]*fac1
bio_poste_Jena_regridded_f = f_Jena_regridded.variables['co2flux_land'][:,:,:]



print 'CTE'
print bio_poste_CTE.shape
print 'CAMS'
print bio_poste_CAMS.shape
print 'Jena'
print bio_poste_Jena.shape
print 'CAMS_regridded'
print bio_poste_CAMS_regridded_f.shape
print 'Jena_regridded'
print bio_poste_Jena_regridded_f.shape
```

```python
# This is a tool to compute regional fluxes for all time steps

import PySCRIP as scrip
from PySCRIP.config import PySCRIPConfig
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap
import numpy as np

# Parameters setting
# EModel:Earth model, define which earch model to use
# inver: Inversion,  which inversion to use:CTE,CAMS,Jena
# res: Resolution,  which grid to use
# count: Country

def estimate_carbon_flows_3(count,modeledData,cellSize,count_mask):

    if modeledData == 'CTE':
        flux = bio_poste_CTE
    elif modeledData == 'CAMS-original':
        flux = bio_poste_CAMS
    elif modeledData == 'CAMS-interpolated':
        flux = bio_poste_CAMS_regridded_f
    elif modeledData == 'Jena-original':
        flux = bio_poste_Jena
    elif modeledData == 'Jena-interpolated':
        flux = bio_poste_Jena_regridded_f


    if cellSize == 'CTE-sphere':
        cellArea = cell_size_CTE_sphe
    elif cellSize == 'CTE-ellipsoid':
        cellArea = cell_size_CTE_ellip
    elif cellSize == 'CAMS-sphere':
        cellArea = cell_size_CAMS_sphe
    elif cellSize == 'CAMS-ellipsoid':
        cellArea = cell_size_CAMS_ellip
    elif cellSize == 'Jena-sphere':
        cellArea = cell_size_Jena_sphe
    elif cellSize == 'Jena-ellipsoid':
        cellArea = cell_size_Jena_ellip

    if count == 'Sweden':
        if count_mask == 'CTE-sphere':
            mask = Swe_mask_1x1_sphe_EA
        elif count_mask == 'CAMS-sphere':
            mask = Swe_mask_CAMS_sphe_EA
        elif count_mask == 'Jena-sphere':
            mask = Swe_mask_Jena_sphe_EA
        elif count_mask == 'CTE-ellipsoid':
            mask = Swe_mask_1x1_ellip_EA
        elif count_mask == 'CAMS-ellipsoid':
            mask = Swe_mask_CAMS_ellip_EA
        elif count_mask == 'Jena-ellipsoid':
            mask = Swe_mask_Jena_ellip_EA
        elif count_mask == 'CTE-planar':
            mask = Swe_mask_1x1_planar


    #print 'flux', flux
    #print 'flux.shape', flux.shape
    #print 'mask.shape', mask.shape
    #print 'cellArea.shape', cellArea.shape

    flow = flux*mask*cellArea
    #sum here only over latitude and longitude indices => regional flux for all time steps
    #print 'Carbon flow of '+count+': '
    #print str(flow.sum(axis=(1,2)))+' PgC/yr'
    #print 'flow.shape', flow.shape

    return flow.sum(axis=(1,2))
```

```python
# example to plot time series of regional fluxes (for Sweden)
# read time information from netcdf files
# cf. Ingrid's notebook GCP-inversions-comparison_v4.0
# for more detailed comparisons of the monthly values the dates need to be synchonized, e.g. by using a pandas dataframe

#Carboscope Jena
dates_Jena=f_Jena.variables['mtime'][:]
dates_Jena=[dt.datetime(2000,1,1,0,0,0) + dt.timedelta(seconds=s.item()) for s in dates_Jena]
#print 'dates_Jena'
#print dates_Jena

#CAMS
dates_CAMS=f_CAMS.variables['time'][:]
dates_CAMS=[dt.datetime(2000,1,1,0,0,0) + dt.timedelta(days=s.item()) for s in dates_CAMS]
#print 'dates_CAMS'
#print dates_CAMS

#CTE
dates_CTE=f_CTE.variables['date'][:]
dates_CTE=[dt.datetime(2000,1,1,0,0,0) + dt.timedelta(days=s.item()) for s in dates_CTE]
#print 'dates_CTE'
#print dates_CTE
```

```python
#example for regional flux
Country = 'Sweden'

Modeled_flux = 'Jena-original'
Cell_size = 'Jena-sphere'
Country_mask = 'Jena-sphere'
flux_Jena_SWE=estimate_carbon_flows_3(Country,Modeled_flux,Cell_size,Country_mask)
print flux_Jena_SWE.shape

Modeled_flux = 'CTE'
Cell_size = 'CTE-sphere'
Country_mask = 'CTE-sphere'
flux_CTE_SWE=estimate_carbon_flows_3(Country,Modeled_flux,Cell_size,Country_mask)
print flux_CTE_SWE.shape

Modeled_flux = 'CAMS-original'
Cell_size = 'CAMS-sphere'
Country_mask = 'CAMS-sphere'
flux_CAMS_SWE=estimate_carbon_flows_3(Country,Modeled_flux,Cell_size,Country_mask)
print flux_CAMS_SWE.shape

Modeled_flux = 'CAMS-interpolated'
Cell_size = 'CTE-sphere'
Country_mask = 'CTE-sphere'
flux_CAMS_interpolated_SWE=estimate_carbon_flows_3(Country,Modeled_flux,Cell_size,Country_mask)
print flux_CAMS_interpolated_SWE.shape
```

```python
Modeled_flux = 'Jena-interpolated'
Cell_size = 'CTE-sphere'
Country_mask = 'CTE-sphere'
flux_Jena_interpolated_SWE=estimate_carbon_flows_3(Country,Modeled_flux,Cell_size,Country_mask)
print flux_Jena_interpolated_SWE.shape

print flux_CTE_SWE[5]
print flux_CAMS_interpolated_SWE[257]
print flux_Jena_interpolated_SWE[269]
```

```python
fig=plt.figure(figsize=(15,5))
plt.plot(dates_Jena,flux_Jena_SWE,label='Jena CarboScope',color='blue')
plt.plot(dates_CAMS,flux_CAMS_SWE,label='CAMS',color='red')
plt.plot(dates_CTE,flux_CTE_SWE,label='CTE',color='green')
plt.title('Biospheric fluxes for '+Country)
plt.ylabel('Carbon flux [PgC/yr]')
plt.legend(loc='best')
plt.show
```

```python
fig=plt.figure(figsize=(15,5))
plt.plot(dates_Jena,flux_Jena_interpolated_SWE,label='Jena CarboScope-interpolated',color='blue')
plt.plot(dates_CAMS,flux_CAMS_interpolated_SWE,label='CAMS-interpolated',color='red')
plt.plot(dates_CTE,flux_CTE_SWE,label='CTE',color='green')
plt.xlim(dt.datetime(2001,1,1),dt.datetime(2015,1,1))
plt.title('Monthly biospheric fluxes for '+Country)
plt.ylabel('Carbon flux [PgC/yr]')
plt.legend(loc='best')
plt.show
```

```
fig=plt.figure(figsize=(15,5))
plt.plot(dates_CAMS,flux_CAMS_SWE,label='CAMS-original',color='blue')
plt.plot(dates_CAMS,flux_CAMS_interpolated_SWE,label='CAMS-interpolated',color='red',linestyle= ':')
plt.xlim(dt.datetime(1979,1,1),dt.datetime(2016,1,1))
plt.title('Monthly biospheric fluxes for '+Country)
plt.ylabel('Carbon flux [PgC/yr]')
plt.legend(loc='best')
plt.show
```

```
fig=plt.figure(figsize=(15,5))
plt.plot(dates_Jena,flux_Jena_SWE,label='Jena-original',color='blue')
plt.plot(dates_Jena,flux_Jena_interpolated_SWE,label='Jena-interpolated',color='red',linestyle= ':')
plt.xlim(dt.datetime(1979,1,1),dt.datetime(2016,1,1))
plt.title('Monthly biospheric fluxes for '+Country)
plt.ylabel('Carbon flux [PgC/yr]')
plt.legend(loc='best')
plt.show
```

2. Manual to compute overlay area/partial value in ArcGIS

1. Create grids
   1) Choose "Create Fishnet" tool to create a grid in the coordinate system GCS_WGS_1984.
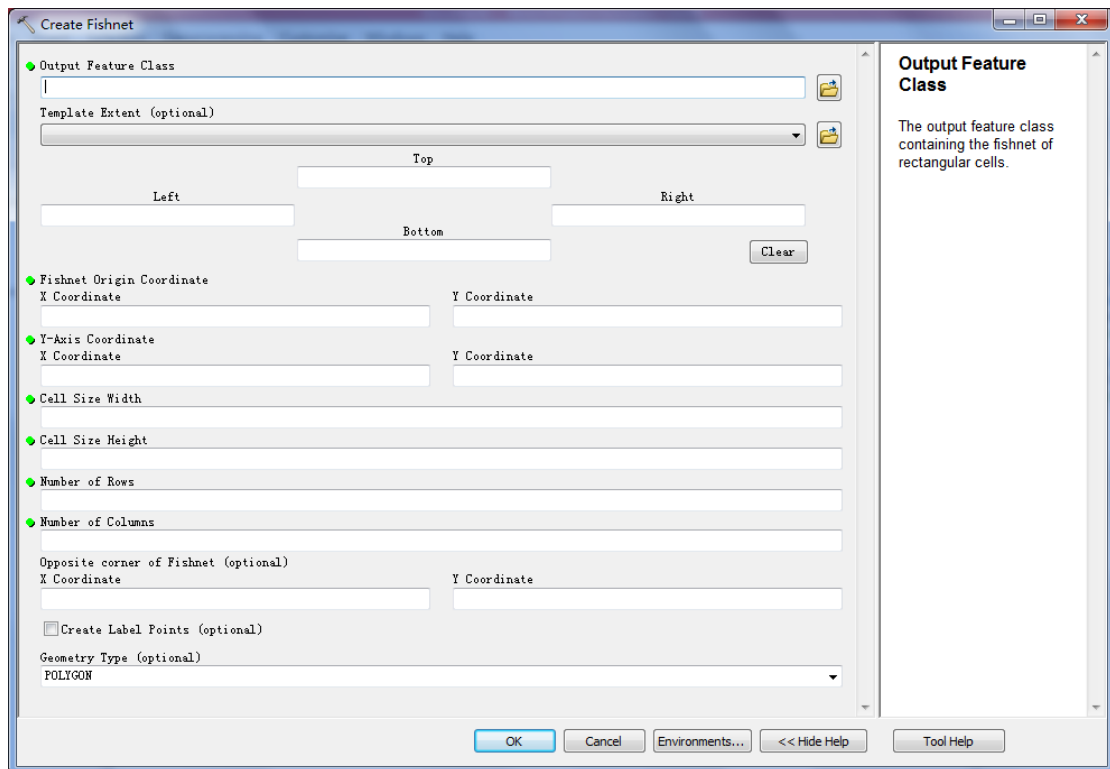
   Set parameters:
   a) template extent: top 90º, bottom -90°, left -180º, right 180°;
   b) "Fishnet origin coordinate" and "Y-axis coordinate" will be set automatically;
   c) Cell size and number can be set according to the grid definition;
   d) Create label points: usually not tick;
   e) Geometry Type: choose "polygon", which makes it possible to intersect grid cells with region area to compute the overlay area in a grid cell.

   2) Create a new map and set the "projected coordinate system" in "Data Frame

Properties" as "Cylindrical_Equal_Area (sphere)". Add the grid created into this map in the cylindrical equal area projection on a spherical earth model

or

Create a new map and set the "projected coordinate system" in "Data Frame Properties" as "Cylindrical_Equal_Area (world)". Add the grid created into this map in the cylindrical equal area projection on an ellipsoidal earth model.

2. Add vector data of regions as a layer into the map including the grid.
3. Compute overlay area (partial value)
   1) Choose "Intersect" tool
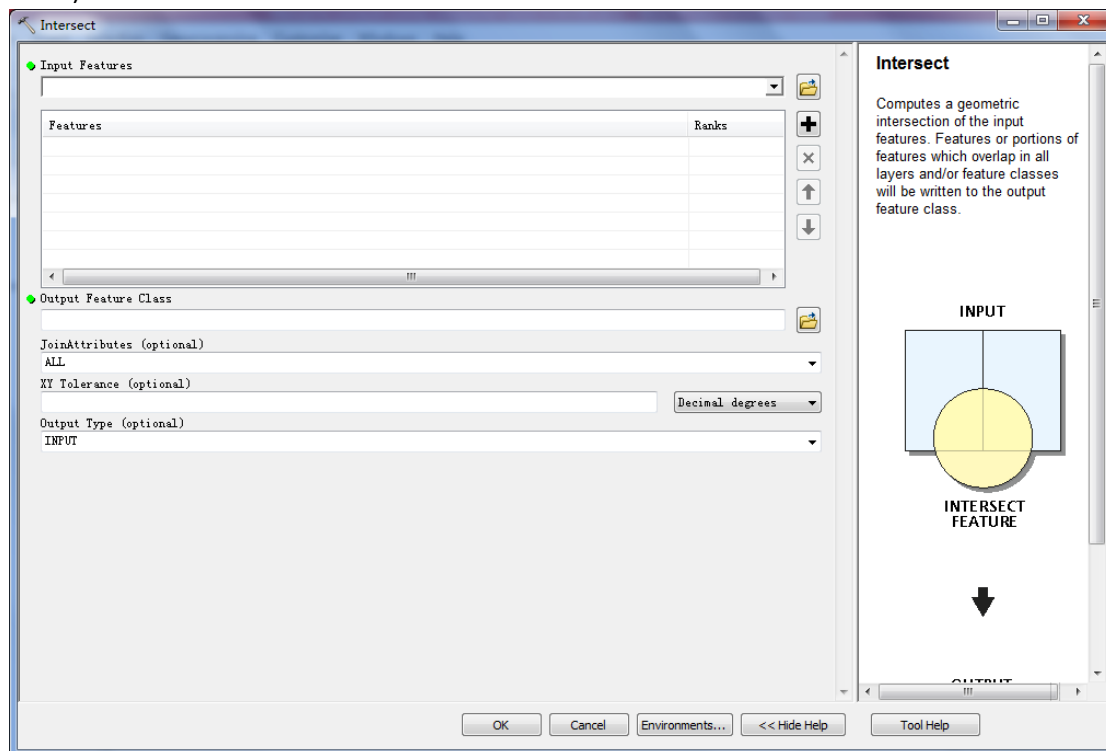
Set parameters:
a) Scroll down "Input Features" and choose the layer of vector data and the layer of the grid created;
b) Choose a location in "Output Feature Class" to store the output feature.

   2) Open "attribute table" of the output feature. The area of overlay area in each grid cell can be found in the table, otherwise "Add field", right click the new column name and choose "calculate geometry" to calculate the area of the overlay area in each grid cell. The attribute table keeps all the fields of both layers intersected. So the area of the grid cells can also be found in the table. A partial value can be calculated by dividing overlay area by the area of the grid cell. However, the "attribute table" only includes the overlay polygons with area larger than 0 in a grid cell. If the grid cell does not cover any part of regions, there is no record for that grid cell in the table.

   3) Export the attribute table to Python, to create land/sea masks. Create a zero array in

the size of the grid create in step 1 using command np.zeros(). Change 0 to the number of overlay area/partial value in the attribute table at the right location in the array.

400   Sofia Sjögren (2016) Effective methods for prediction and visualization of contaminated soil volumes in 3D with GIS

401   Jayan Wijesingha (2016) Geometric quality assessment of multi-rotor unmanned aerial vehicle-borne remote sensing products for precision agriculture

402   Jenny Ahlstrand (2016) Effects of altered precipitation regimes on bryophyte carbon dynamics in a Peruvian tropical montane cloud forest

403   Peter Markus (2016) Design and development of a prototype mobile geographical information system for real-time collection and storage of traffic accident data

404   Christos Bountzouklis (2016) Monitoring of Santorini (Greece) volcano during post-unrest period (2014-2016) with interferometric time series of Sentinel-1A

405   Gea Hallen (2016) Porous asphalt as a method for reducing urban storm water runoff in Lund, Sweden

406   Marcus Rudolf (2016) Spatiotemporal reconstructions of black carbon, organic matter and heavy metals in coastal records of south-west Sweden

407   Sophie Rudbäck (2016) The spatial growth pattern and directional properties of *Dryas octopetala* on Spitsbergen, Svalbard

408   Julia Schütt (2017) Assessment of forcing mechanisms on net community production and dissolved inorganic carbon dynamics in the Southern Ocean using glider data

409   Abdalla Eltayeb A. Mohamed (2016) Mapping tree canopy cover in the semi-arid Sahel using satellite remote sensing and Google Earth imagery

410   Ying Zhou (2016) The link between secondary organic aerosol and monoterpenes at a boreal forest site

411   Matthew Corney (2016) Preparation and analysis of crowdsourced GPS bicycling data: a study of Skåne, Sweden

412   Louise Hannon Bradshaw (2017) Sweden, forests & wind storms: Developing a model to predict storm damage to forests in Kronoberg county

413   Joel D. White (2017) Shifts within the carbon cycle in response to the absence of keystone herbivore *Ovibos moschatus* in a high arctic mire

414   Kristofer Karlsson (2017) Greenhouse gas flux at a temperate peatland: a comparison of the eddy covariance method and the flux-gradient method

415   Md. Monirul Islam (2017) Tracing mangrove forest dynamics of Bangladesh using historical Landsat data

416   Bos Brendan Bos (2017) The effects of tropical cyclones on the carbon cycle

417   Martynas Cerniauskas (2017) Estimating wildfire-attributed boreal forest burn in Central and Eastern Siberia during summer of 2016

418   Caroline Hall (2017)The mass balance and equilibrium line altitude trends of glaciers in northern Sweden

419   Clara Kjällman (2017) Changing landscapes: Wetlands in the Swedish municipality Helsingborg 1820-2016

420   Raluca Munteanu (2017) The effects of changing temperature and precipitation rates on free-living soil Nematoda in Norway.

421   Neija Maegaard Elvekjær (2017) Assessing Land degradation in global drylands and possible linkages to socio-economic inequality

422   Petra Oberhollenzer, (2017) Reforestation of Alpine Grasslands in South Tyrol: Assessing spatial changes based on LANDSAT data 1986-2016

423   Femke, Pijcke (2017) Change of water surface area in northern Sweden

424   Alexandra Pongracz (2017) Modelling global Gross Primary Production using the correlation between key leaf traits

425   Marie Skogseid (2017) Climate Change in Kenya - A review of literature and evaluation of temperature and precipitation data

426   Ida Pettersson (2017) Ekologisk kompensation och habitatbanker i kommunalt planarbete

427   Denice Adlerklint (2017) Climate Change Adaptation Strategies for Urban Stormwater Management – A comparative study of municipalities in Scania

428   Johanna Andersson (2017) Using geographically weighted regression (GWR) to explore spatial variations in the relationship between public transport accessibility and car use : a case study in Lund and Malmö, Sweden

429   Elisabeth Farrington (2017) Investigating the spatial patterns and climate dependency of Tick-Borne Encephalitis in Sweden

430   David Mårtensson (2017) Modeling habitats for vascular plants using climate factors and scenarios - Decreasing presence probability for red listed plants in Scania

431   Maja Jensen (2017) Hydrology and surface water chemistry in a small forested catchment : which factors influence surface water acidity?

432   Iris Behrens (2017) Watershed delineation for runoff estimations to culverts in the Swedish road network : a comparison between two GIS based hydrological modelling methods and a manually delineated watershed

433   Jenny Hansson (2017) Identifying large-scale land acquisitions and their agro-ecological consequences : a remote sensing based study in Ghana

434   Linn Gardell (2017) Skyddande, bevarande och skapande av urbana ekosystemtjänster i svenska kommuner

435   Johanna Andersson (2017) Utvärdering av modellerad solinstrålning i södra Sverige med Points Solar Radiation i ArcGIS

436   Huiting Huang(2017) Estimating area of vector polygons on spherical and ellipsoidal earth models with application in estimating regional carbon flows