Reconstruction of gap-free time series satellite observations of land surface temperature to model spectral soil thermal admittance

PhD Thesis

Hamid Reza Ghafarian Malamiri
Reconstruction of gap-free time series satellite observations of land surface temperature to model spectral soil thermal admittance

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op woensdag 6 mei 2015 om 12.30 uur

door

Hamid Reza Ghafarian Malamiri
Master of Science in Geo-information Science and Earth Observation (Geoinformatics)
University of Twente, Netherlands

geboren te Yazd, Iran.
Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. Massimo Menenti

Prof. dr. Li Jia

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter

Prof. dr. M. Menenti, Technische Universiteit Delft, promotor

Prof. dr. Li Jia, Chinese Academy of Science (China), promotor

Prof. dr. J.A. Sobrino, University of Valencia (Spain)

Dr. Ali Reza Amiri Simkooei, University of Isfahan (Iran)

Prof. dr. Wim G.M. Bastiaanssen, Technische Universiteit Delft

Prof. dr. Z. (Bob) Su, Universiteit Twente

Dr. Darren Ghent, University of Leicester (UK)

Prof. dr. H. Russchenberg, Technische Universiteit Delft, reservelid

ISBN 978-94-6186-466-6

Copyright ©2015 by Hamid Reza Ghafarian Malamiri.

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic, mechanical, including photocopying, recording or by any information storage and retrieval system, without the prior permission of the author.

Cover page: http://spaceflight.nasa.gov/
Preface

This dissertation would not have been accomplished without the support and help of many people in different ways and I would like to express my gratitude to all of them.

First, I would like to thank my promotor, Prof. dr. Massimo Menenti, for giving me the opportunity to pursue my PhD under his supervision. Thank you for your support, valuable comments and discussions throughout my research. I was impressed by your tremendous ideas over a wide range of research subjects.

I would like to thank Prof. dr. Li Jia for her scientific comments on the basic problems at the first stage of my research. Thank you for the discussions that we had on different topics during your visits to TUDelft. I wish to thank Dr. William Kustas for carefully reading my thesis and for his useful comments as a member of the evaluation committee.

I would like to gratefully acknowledge the financial support of my PhD from the Iranian Ministry of Science, Research, and Technology (MSRT).

Furthermore, I would like to thank my best friend, Ali Mousivand, for sharing his knowledge and support with some of the methods used in this thesis to retrieve the soil thermal properties and implementing the sensitivity analysis. I would like to thank Rene den Ouden for his help in data pre-processing and sharing some results. I would like to thank Ben Gorte and Marcel Kleinherenbrink for Dutch translation of the summary and propositions. I am thankful to Lidwien De Jong, Rebeca Domingo and Marjolein de Niet-de Jager for all the support during my PhD. I wish to express my appreciation to all my other friends and colleagues for making my stay in the Netherlands, a pleasant one, especially Maryam and Amin, Ali and Fatemeh, Alireza and Shahin, Hamid and Marzieh, Ali and Maryam, Enayat and Fariba and many others. I am grateful to my family and my wife’s family for all their support and encouragement.

Last but not least, my special appreciation goes to my lovely wife, Fatemeh, who has been always supporting me and sharing all the difficult times with great love and sacrifices. I owe many thanks to you and our daughter, Shaya, who brought a lot of joy and happiness to our family, for your great companionship and love.
Abstract

The soil thermal properties (soil thermal conductivity, soil heat capacity and soil diffusivity) are the main parameters in the applications that need quantitative information on soil heat transfer. Conventionally, these properties are either measured in situ or estimated by semi-empirical models using the fractions of soil constituents. The use of such methods over large and heterogeneous areas, however, is often costly, time-consuming and sometimes impractical.

This thesis proposes and evaluates a new approach to estimate the soil thermal properties by inverse modelling of Spectral Soil Thermal Admittance (SSTA) which is determined using the time series satellite observations of Land Surface Temperature (LST) and soil heat flux \( (G_0) \) over the entire Qinghai-Tibet Plateau (QTP) from 2008 to 2010. To calculate the soil thermal admittance, the amplitudes of \( G_0 \) and LST at significant frequencies are required which needs consistent, continuous and long time series. The hourly FY-2C LST time series used in this study were often contaminated by missing data (gaps) and outliers. The HANTS algorithm and M-SSA were used to fill the gaps and remove the outliers in the LST time series. Then, the gap-filled hourly LST was used to identify the most significant periodic components over a three-year data. The amplitude of soil heat flux and LST were estimated at significant frequencies and then the soil thermal admittance at each frequency was determined over the study area. The SSTA, which is the variation of STA against frequency, contains information about the soil thermal properties of different soil layers. An inversion model was used to estimate soil thermal properties of different soil layers (assuming three-layer soil) over the Q-TP.

Keywords: Gap-filling, HANTS algorithm, (Multi) – Singular Spectrum Analysis (M – SSA), soil thermal properties, sensitivity analysis, model inversion.
Summary

The Tibetan Plateau with a heterogeneous and elevated landscape is characterized by its various land surface properties (e.g. soil water content and soil thermal properties) and has a significant impact on the Asian Monsoon system. The variability of these properties combined with their associated effects on the water and energy balance provides necessary information on water resources and the interaction between the earth’s surface and the atmosphere. Knowledge of the exchange of available energy between the land surface and the atmosphere is vital for monitoring land surface processes and for land and water management decisions. Soil heat flux as one of the components of the surface energy balance effectively combines the energy transfer processes at the surface with the soil thermal regime and determines the Land Surface Temperature (LST). The analytical relationship between LST, soil heat flux and soil thermal properties (e.g. soil thermal conductivity) leads to the definition of the apparent thermal admittance at the soil surface. The soil thermal admittance can be calculated as the ratio between the amplitudes of soil heat flux and LST at any frequency. The Spectral Soil Thermal Admittance (SSTA) was defined as a set of observations of the apparent soil thermal admittance at multiple frequencies. SSTA conveys valuable information about the soil thermal properties of different soil layers. The soil thermal properties have a great influence on partitioning of energy at the surface and determine the land surface temperature and exchange of soil heat flux across the soil profile. These properties are conventionally measured in situ or estimated using semi-empirical models, though this is resource demanding, costly and impractical for large area.

In the current study, a new approach was proposed to estimate the soil thermal properties by inverse modelling of spectral soil thermal admittance. To achieve this the following steps have to be accomplished:

The SSTA was estimated by time series satellite observations of LST and soil heat flux, provided consistent, continuous and long time series with sufficient temporal resolution (e.g. hourly). The hourly LST and daily soil heat flux time series estimated from the FY-2C geostationary satellite were used in this study covering the entire Tibetan Plateau from 2008 to 2010. The LST time series are degraded by missing data (i.e. gaps) with different number, size and distribution, outliers and noise due to clouds, aerosols and algorithm malfunctioning which cause irregular observations.
The specific results of each chapter are summarized below.

**Chapter 3** describes the use of the HANTS algorithm to identify and remove the outliers and filling the short gaps that already exist and the ones added by removing the outliers. Two versions of HANTS (i.e. new and old version) were applied to hourly time series of LST. The results suggest that the old version is more suitable than the new version due to its capability to select the significant periodic components independently in the reconstruction of the time series. Different parameter settings were evaluated by numerical experiments to choose the proper HANTS parameters. The HANTS algorithm was evaluated using both ground measurements and synthetic data in which the gaps are introduced artificially with different size, number and distribution. The results showed that the HANTS algorithm has the capability of filling short gaps, i.e. when the number of gaps is less than the half the number of observations, with final acceptable accuracy. HANTS uses temporal correlation in gap-filling. To fill long gaps, i.e. when the number of gaps is larger than half the number of observations, another methodology was used which makes use of both spatial and temporal correlation in gap-filling.

**Chapter 4** describes the (Multi)-Singular Spectrum Analysis (M-SSA) to fill the gaps and remove outliers. The SSA was applied to monthly samples of a time series of hourly LST to select the number of periodic components and window size. The window size and number of components were selected as 72 hours and seven respectively. The performance of M-SSA was evaluated using both ground measurements and synthetic data. The results showed the promising capability of M-SSA in gap-filling even when long gaps exist in the data. The M-SSA was finally applied on a monthly basis to reconstruct LST time series over the entire Tibetan Plateau from 2008 to 2010.

**Chapter 5** describes the estimation of the amplitudes of LST and soil heat flux at significant frequencies to determine the spectral soil thermal admittance. The significant periodic components were identified on three-year hourly LST data using the FFT. The amplitudes of LST at those frequencies were estimated by Fourier analysis and least square method. The amplitudes of soil heat flux were estimated using a procedure that requires only the daily average and one instantaneous value of soil heat flux each day. The soil thermal admittance at each frequency was calculated as the ratio between estimated amplitudes of soil heat flux and LST at the same frequency over the Tibetan Plateau. The final products were maps of the soil thermal admittance at each frequency.
The spectral soil thermal admittance was defined and determined for each pixel in the study area as the set of values of soil thermal admittance as a function of frequency.

**Chapter 6** describes a new method to estimate the soil thermal properties of different soil layers (assuming three-layer soil in this study) by inverse modelling of SSTA estimated from satellite observations. The estimated SSTA using satellite observations was inverted against the SSTA simulated by a physical forward model. A sensitivity analysis was conducted to find the influential parameters to which the forward model is sensitive. The capability of the forward model to simulate the SSTA was evaluated using the ground measurements of SSTA as reference against the SSTA calculated with the forward model and soil thermal properties measured or estimated by applying semi-empirical methods to soil textural data. The accuracy of retrieved soil thermal properties was validated by using synthetic soil thermal properties estimated by the semi-empirical model assuming three-layered soil with various fraction of soil constituents and soil water content. These synthetic soil thermal properties were used to generate synthetic SSTA data, and then inversion model was used to retrieve the soil thermal properties using the synthetic SSTA as observations. Finally, the soil thermal properties of a three-layered soil were estimated pixel by pixel over the study area, and the corresponding map for each soil thermal property in each soil layer was created. The retrieved soil thermal properties were compared with soil thermal properties estimated globally using Pedotransfer Functions (PTFs) established as described by Dai et al. (2013) and Shangguan et al. (2013). The results showed that the retrieved soil thermal conductivity falls within the range of the reference soil thermal conductivity for dry and saturated conditions.

Finally, **Chapter 7** presents the overall conclusions of this study and an outlook for future work by proposing recommendations.

The estimation of soil thermal properties in different soil layers by inverse modelling of spectral soil thermal admittance is the central topic of this study. The results presented in this thesis can be used as input parameters in applications dealing with quantitative information on soil heat transfer which, in turn, depends on the soil thermal properties.
Samenvatting

Het Tibetaans Plateau met zijn heterogene en hooggelegen landschap wordt gekarakteriseerd door verschillende eigenschappen van het oppervlak (zoals bodemvochtigheid en de thermische eigenschappen van de bodem) en heeft een significante invloed op het Aziatisch Moessonsysteem. De variabiliteit van deze eigenschappen gecombineerd met hun effecten op de water- en energiebalans levert informatie over watervoorraden en over de interactie tussen het aardoppervlak en de atmosfeer. Kennis over de uitwisseling van de beschikbare energie tussen het aardoppervlak en de atmosfeer is essentieel bij het waarnemen van oppervlakteprocessen en het beheer van land en water. De warmteflux van de grond als een van de componenten van de oppervlakte-energiebalans combineert effectief de energieoverdracht processen aan het oppervlak met het thermische regime van de grond en bepaalt de landoppervlaktetemperatuur (LST). De analytische relatie tussen de LST, de warmteflux en thermische eigenschappen van de grond (voornamelijk de thermische geleidbaarheid) leidt tot de definitie van de schijnbare thermische admittantie van het grondoppervlak. De thermische admittantie kan berekend worden als de verhouding tussen de amplitudes van de warmteflux en de LST bij elke frequentie. De spectrale thermische admittantie van de grond (SSTA) is gedefinieerd als een verzameling waarnemingen van de schijnbare thermische admittantie van het grondoppervlak bij verschillende frequenties. SSTA geeft waardevolle informatie over de thermische eigenschappen van verschillende bodemlagen. De thermische eigenschappen hebben grote invloed op de verdeling van energie aan het oppervlak en bepalen de LST en de uitwisseling van warmtefluxen over het bodemprofiel. Deze eigenschappen worden conventioneel in-situ gemeten of geschat met gebruik van semi-empirische modellen, ook al is dit rekenkundig intensief, duur en niet praktisch voor een groot oppervlak. In deze studie wordt een nieuwe manier voorgesteld om de thermische eigenschappen van de bodem te schatten door middel van invers modelleren van de SSTA. De volgende stappen zijn nodig om dit mogelijk te maken:

De SSTA kan geschat worden uit tijdreeksen van satellietwaarnemingen van de LST en van de warmteflux van de grond, mits de tijdreeksen consistent, continu, lang genoeg zijn en voldoende temporele resolutie hebben (in dit geval respectievelijk eens per uur en eens per dag). De urrmetingen van de LST en de dagelijkse warmteflux werden in

De specifieke resultaten per hoofdstuk worden hieronder opgesomd.

**Hoofdstuk 3** beschrijft het gebruik van het HANTS algoritme om uitschieters te identificeren en te verwijderen en om de gaten in de data te vullen. Twee versies van het HANTS algoritme (een oude en een nieuwe versie) worden toegepast op de uurwaarnemingen van de LST. De resultaten geven de indruk dat de oude versie van het algoritme geschikter is dan de nieuwe versie, doordat het deze het vermogen heeft om de significante periodieke componenten onafhankelijk te selecteren in de reconstructie van de tijdreeksen. Verschillende parameter-instellingen zijn geëvalueerd met numerieke experimenten om de juiste HANTS parameters te selecteren. Het HANTS algoritme is geëvalueerd door gebruik te maken van in-situ metingen en synthetische data, waarin kunstmatige gaten zijn geïntroduceerd van verschillende groottes, aantallen en verdelingen. De resultaten laten zien dat het HANTS algoritme in staat is kleine gaten met een aanvaardbare nauwkeurigheid te vullen, mits het aantal gaten kleiner is dan de helft van het aantal waarnemingen. HANTS gebruikt temporele correlatie bij het vullen van de gaten. Om grote gaten te vullen, dus wanneer het aantal gaten meer de helft van de observatie omvat, wordt een andere methode gebruikt, die gebruik maakt van zowel ruimtelijke als temporele correlatie.

**Hoofdstuk 4** beschrijft de (Multi-)singuliere spectrale analyse (M-SSA) om gaten te vullen en uitschieters te verwijderen. SSA wordt toegepast op maandelijkse sets van de uurwaarnemingen van de LST om het aantal periodieke componenten en de grootte van het venster te bepalen. Voor de venstergrootte is 72 uur gekozen en het voor het aantal componenten de waarde zeven. De prestatie van de M-SSA is geëvalueerd aan de hand van in-situ metingen en synthetische data. De resultaten laten het veelbelovende vermogen van M-SSA voor het vullen van data zien, zelfs wanneer er grote gaten in de data zitten. De M-SSA is uiteindelijk maandelijks toegepast om tijdreeksen van LST te reconstrueren over het gehele Tibetaans Plateau van 2008 tot 2010.

**Hoofdstuk 5** beschrijft het schatten van de amplitudes van de LST en de warmteflux van de grond op significante frequenties om de SSTA te bepalen. De significante periodische componenten zijn geïdentificeerd uit de drie jaar lange uurwaarnemingen
van LST met behulp van het FFT-algoritme. De amplitudes van de LST op deze frequenties zijn geschat met behulp van Fourier-analyse en de kleinste-kwadratenmethode. De amplitudes van de warmteflux zijn geschat met een methode die alleen het dagelijkse gemiddelde en de momentane waarde van de warmteflux nodig heeft per dag. De thermische admittantie van de grond bij elke frequentie is berekend als de verhouding tussen de geschatte amplitudes van de warmteflux en de LST bij die frequentie op het Tibetaans Plateau. De uiteindelijke producten zijn kaarten van de thermische admittantie van de bodem bij elke frequentie. De SSTA is gedefinieerd en bepaald voor elke pixel in het onderzoeksgebied als de verzameling waardes van de thermische admittantie als een functie van frequentie.

_Hoofdstuk 6_ beschrijft een nieuwe methode om de thermische eigenschappen van de grond te bepalen voor verschillende lagen (in deze studie door het verdelen van de grond in drie lagen) door inverse modelleren van SSTA schattingen van satellietwaarnemingen. De geschatte SSTA vanuit satellietwaarnemingen is geïnverteerd tegen de gesimuleerde SSTA van een fysisch model. Een gevoeligheidsanalyse is verricht om te bepalen voor welke parameters het model het meest gevoelig is. Het vermogen van het model om SSTA te simuleren is geëvalueerd met behulp van in-situ metingen van SSTA als referentie tegen de SSTA berekend met een model en de thermische eigenschappen van de grond (gemeten of geschat met een semi-empirische methode uit textuurmatige data). De nauwkeurigheid van de verkregen thermische bodemeigenschappen is gevalideerd door middel van synthetische thermische eigenschappen van de bodem, geschat door een semi-empirisch model waar een drielaagse bodem wordt aangenomen met verschillende bodembestanddelen en verschillende fracties bodemvochtigheid. Deze synthetische thermische eigenschappen zijn gebruikt om synthetische SSTA data te genereren en invers modelleren met synthetische SSTA data is gebruikt om de thermische eigenschappen van de grond te verkrijgen. Uiteindelijk zijn de drielaagse thermische eigenschappen van de grond pixel voor pixel geschat over het gehele studiegebied en is een corresponderen kaart van elke thermische eigenschap in elke laag gecreëerd. De verkregen thermische eigenschappen zijn vergeleken met de thermische eigenschappen die globaal geschat zijn met Pedotransfer functies (PTFs) op de manier van Dai et al. (2013) en Shangguan et al. (2013). De resultaten laten zien dat de verkregen thermische geleidbaarheid van de grond correct is binnen het bereik van de referentie thermische geleidbaarheid voor droge en verzadigde omstandigheden.
Uiteindelijk presenteert **hoofdstuk 7** de belangrijkste conclusies van deze studie en geeft het aanbevelingen voor verder onderzoek.

Het schatten van de thermische eigenschappen van verschillende grondlagen door middel van het omgekeerd modelleren van SSTA is het centrale onderwerp. De resultaten gepresenteerd in dit proefschrift kunnen gebruikt worden als invoerparameters voor toepassingen die kwantitatieve informatie behandelen op het gebied van bodem warmteverplaatsing, welke afhangt van de thermische eigenschappen van de grond.
Contents

1. General introduction ........................................................................................................ 1
   1.1. Introduction ........................................................................................................ 1
1.2. General problems ......................................................................................................... 3
   1.2.1. Gaps and outliers in hourly LST time series data ........................................... 4
   1.2.2. Estimation of soil thermal admittance at significant frequencies .......... 5
   1.2.3. Estimation of soil thermal properties via inverse modelling of spectral soil
          thermal admittance ......................................................................................... 6
1.3. General objective ........................................................................................................ 7
1.4. Overview of approach ............................................................................................... 8
1.5. Research questions .................................................................................................... 9
1.6. Outline of the thesis .................................................................................................. 9

2. Description of the study area and data sets .................................................................... 11
   2.1. Introduction on the CEOP-AEGIS project ......................................................... 11
   2.2. Study area ........................................................................................................... 12
   2.3. Satellite data ......................................................................................................... 12
      2.3.1. Retrieval of emissivity and LST using FY-2C observations .................... 14
      2.3.2. Retrieval algorithm of LST and LSE ....................................................... 14
      2.3.3. Retrieval algorithm of soil heat flux ....................................................... 15
   2.4. Ground station data ............................................................................................. 15
      2.4.1. Nagqu (BJ) station ................................................................................... 16
   2.5. Exploration of the LST time series data ............................................................. 16
      2.5.1. Mean LST over study area during November 2008 and 2008 year .......... 16
      2.5.2. Spatial extent of gaps ............................................................................... 18
      2.5.3. Temporal extent of gaps ........................................................................... 19
      2.5.4. Maximum continuous gap size in LST time series ................................. 19
   2.6. Conclusion ............................................................................................................ 20

3. Filling short gaps with the HANTS algorithm .............................................................. 23
   3.1. Introduction ........................................................................................................... 23
   3.2. HANTS algorithm ............................................................................................... 26
      3.2.1. The evolution of HANTS over time ......................................................... 28
      3.2.2. Definition of significant frequencies to model time series with HANTS. 29
3.3. Fourier series analysis of LST time series data ............................................. 31
  3.3.1. Fast Fourier Transform (FFT) and power spectrum analysis ................. 31
  3.3.2. Goodness of fit criterion .............................................................. 34
3.4. Results and discussions ............................................................................. 35
  3.4.1. Analysis of a three days LST time series with FFT and HANTS .......... 35
  3.4.2. Identification of most relevant periodic components in one month LST data 37
  3.4.3. HANTS results applied on one month LST time series ..................... 38
  3.4.4. HANTS performance and validation with ground measurements ....... 41
  3.4.5. HANTS performance over the entire study area ............................... 44
  3.4.6. Evaluation of the HANTS gap-filling using synthetic data ............... 46
3.5. Conclusion ................................................................................................. 47
4. Filling long gaps with the singular spectrum analysis .................................... 49
  4.1. Introduction ............................................................................................. 49
  4.2. The singular spectrum ........................................................................... 51
  4.3. The SSA algorithm ............................................................................... 52
  4.4. Different types of noise in time series analysis ....................................... 55
    4.4.1. White noise ................................................................................... 56
    4.4.2. Autocorrelated noise (red noise) ................................................. 57
  4.5. Signal detection ...................................................................................... 57
  4.6. Monte Carlo SSA .................................................................................. 58
    4.6.1. Eigenspectrum-shape Monte Carlo test ....................................... 59
    4.6.2. Data-based and null-hypothesis-based Monte Carlo test ............. 59
  4.7. Trend analysis and stationarity ............................................................... 61
  4.8. Results and discussions ......................................................................... 62
    4.8.1. Reconstruction of LST time series using SSA ............................... 63
    4.8.2. Window size and most relevant periodic components .................. 63
    4.8.3. Finding the most significant periodic components ....................... 66
    4.8.4. Visual interpretation of singular value spectrum ......................... 66
    4.8.5. Monte Carlo SSA .......................................................................... 67
    4.8.6. Positive and negative outliers removal ....................................... 72
    4.8.7. Validation of SSA using ground measurements of LST ............... 76
    4.8.8. Evaluation of M-SSA performance in gap-filling and noise reduction ... 77
  4.9. Conclusion ............................................................................................. 82
5. Estimation of the Spectral Soil Thermal Admittance (SSTA) ........................................ 83
   5.1. Introduction ............................................................................................................. 83
   5.2. Theory and method ............................................................................................... 86
   5.3. Results and discussions ........................................................................................ 89
      5.3.1. Finding the most significant periodic components in LST time series .......... 90
      5.3.2. Estimating the amplitude and phase of LST data at significant frequencies 90
      5.3.3. Estimating the amplitude of soil heat flux ....................................................... 93
      5.3.4. Validation of amplitudes estimated using incomplete soil heat flux data .... 99
      5.3.5. Soil thermal admittance .................................................................................. 100
   5.4. Conclusion ............................................................................................................. 102

6. Retrieval of soil thermal properties by inverse modeling of spectral soil thermal
   admittance ..................................................................................................................... 103
   6.1. Introduction ............................................................................................................. 103
   6.2. Soil thermal properties .......................................................................................... 106
      6.2.1. Soil thermal conductivity ................................................................................. 106
      6.2.2. Soil volumetric heat capacity ......................................................................... 108
      6.2.3. The physical model of soil thermal admittance assuming a three-layer soil 109
   6.3. Methodology .......................................................................................................... 110
      6.3.1. Smoothing spectral soil thermal admittance ................................................... 110
      6.3.2. Sensitivity analysis ......................................................................................... 113
      6.3.3. Parameter retrieval using the inversion method .............................................. 116
   6.4. Results and discussions .......................................................................................... 118
      6.4.1. Forward model validation .............................................................................. 118
      6.4.2. Data preparation for forward model ............................................................... 119
      6.4.3. Estimation of soil thermal properties .............................................................. 121
      6.4.4. The results of validation ............................................................................... 121
      6.4.5. Validation of estimated SSTA from satellite data ........................................... 123
      6.4.6. Sensitivity analysis of SSTA response to soil thermal properties .............. 125
      6.4.7. Parameter retrieval ....................................................................................... 126
      6.4.8. Validation of model inversion ....................................................................... 127
      6.4.9. Soil thermal properties maps ......................................................................... 129
   6.5. Conclusion .............................................................................................................. 132
7. Conclusions and recommendations ................................................................. 135
   7.1. Brief summary .......................................................................................... 135
   7.2. Results versus research objectives ......................................................... 136
   7.3. Recommendations .................................................................................. 139
8. References ........................................................................................................ 141
List of Figures

Figure 1.1: Global energy balance between the earth and the atmosphere (http://www.noaa.gov/). ................................................................. 2
Figure 1.2: Yearly LST time series, related yearly component and amplitude. .......... 6
Figure 1.3: Overview of the approach developed in this thesis to estimate the soil thermal properties using time series satellite observations of LST and soil heat flux. ................................................................. 10
Figure 2.1: Tibetan Plateau (http://www.zonu.com). ........................................... 13
Figure 2.2: Mean LST values in Tibetan Plateau during November 2008 (a) and 2008 year (b) ................................................................. 17
Figure 2.3: Histogram of mean LST data during November 2008 and 2008 year in TP. ................................................................................. 17
Figure 2.4: Hourly fraction of gaps (missing pixels) for the entire study area............. 18
Figure 2.5: Percentage of gaps in each pixel during November 2008 (a) and the year 2008 (b)................................................................. 19
Figure 2.6: The maximum gap length (hours) in time series of LST data during November 2008 (left) and the whole year 2008 (right). ................. 20
Figure 2.7: Land use map of the four stations on the Tibetan Plateau (W.BABEL et al., 2011). ................................................................. 21
Figure 3.1: A hourly LST time series with the short gaps (top), and long gaps (bottom). ................................................................. 24
Figure 3.2: A LST time series contaminated by gaps and outliers and HANTS output after filling the gaps and removing the outliers. ................................. 28
Figure 3.3: Old (left) and new (right) version of HANTS software analysis tab interface. ............................................................................. 29
Figure 3.4: Reconstructed one-month hourly LST time series applying HANTS with length of period = 72 hours. ................................................................. 29
Figure 3.5: A comparison of the performance of the new and old HANTS algorithm. .. 30
Figure 3.6: A time series and corresponding power spectrum (http://classes.yale.edu/). 32
Figure 3.7: Synthetic time series (top) and corresponding power spectrum (bottom). .. 33
Figure 3.8: Power spectrum of three days of hourly LST data. ............................... 36
Figure 3.9: The results of the new HANTS algorithm with different parameters during three days. ................................................................. 37
Figure 3.10: Amplitude vs. period for one-month LST time series with a minimum number of gaps.

Figure 3.11: Different curve-fitting of one-month hourly LST time series using the old version of HANTS (For clarity just ten days are shown).

Figure 3.12: Different curve-fitting of one-month hourly LST time series using new version of HANTS (For clarity just ten days are shown here).

Figure 3.13: The hourly ground measurements of LST data during January 2008 with and without gaps, and reconstructed data based on parameterizations in Table 3.6.

Figure 3.14: The original LST map of 1st January 2008 at 00:00 AM (top) and the reconstructed LST map for the same time using the old version of HANTS (bottom).

Figure 3.15: The RMSE (top) and MAE (bottom) map calculated based on valid observations and corresponding reconstructed signal by HANTS.

Figure 3.16: The histograms of the RMSE map (left) and MAE map (right) calculated based on valid observations and corresponding reconstructed signal by HANTS.

Figure 3.17: The original and reconstructed LST time series for a pixel in the study area with 78.63% gaps during January 2008.

Figure 3.18: The variation of RMSE against TNL% and CGS% to evaluate the HANTS gap-filling capability.

Figure 4.1: The original and reconstructed ground measurements LST data applying SSA with window size = 432 (every 10 minutes) and number of components = 7; the RMSE and MAE of reconstructed signal against original data are 2.75K and 2.13K respectively.

Figure 4.2: R-squared values of estimated and observed LST ground measurements as a function of the number of components (left), and as a function of the window size (right).

Figure 4.3: Gap-free reconstructed LST time series during January 2008 using SSA with Number of components (No.com) = 7 and window size = 72 hours.

Figure 4.4: Singular values spectrum of data with window size of 72 hours with 7 modes above the flat part of the spectrum.

Figure 4.5: Graph of normalized singular values with 72 hour window size.

Figure 4.6: Singular value spectrum of LST time series using various window sizes.
Figure 4.7: Monte Carlo SSA using the eigenvalue shape test ......................... 69
Figure 4.8: Monte Carlo SSA based on data and null-hypothesis EOFs test ... 70
Figure 4.9: SSA T-EOFs of components 1-2 (top), 3-4 (middle) and 6-7 (down) .... 71
Figure 4.10: Monte Carlo SSA based on hybrid and null-hypothesis EOFs test ........ 72
Figure 4.11: Reconstruction and outliers removal using M-SSA ......................... 73
Figure 4.12: Reconstruction of LST before and after outliers removal .................. 73
Figure 4.13: The original LST map for the 1st of January 2008 at 00:00 AM (top), the reconstructed LST map for the same time using M-SSA (bottom) .............. 74
Figure 4.14: The original and reconstructed LST time series for a pixel in the study area with 78.63% gaps during January 2008 using HANTS and SSA ............. 75
Figure 4.15: Validation of SSA gap-filling using ground measurements with the same pattern of gaps as the corresponding pixel in the satellite image data .................. 76
Figure 4.16: Algorithm for generating gaps (left) and noise (right) ...................... 77
Figure 4.17: The synthetic gappy data set (for clarity only 150 out of 18054 pixels are shown here); each column shows the variation of LST in each pixel in time (one month = 744 hours) and the white areas show the gaps patterns .................. 78
Figure 4.18: Schematic representation of Continuous Gap Size (CGS). Number of Gap pattern (NG) and Total Number of Losses (TNL) .......................... 79
Figure 4.19: The variation of RMSE against different gap patterns (TNL and CGS) ... 79
Figure 4.20: The evaluation of M-SSA for noise removal with different levels of red noise added to pure time series ................................................................. 80
Figure 4.21: The performance of M-SSA for different levels of noise and TNL ....... 81
Figure 4.22: The performance of M-SSA for different levels of noise and CGS ....... 81
Figure 5.1: A soil slab in the depth z ................................................................. 87
Figure 5.2: Original three-year hourly LST time series for a pixel in the study area (top); power spectrum of LST time series against period in days (middle); accumulated power in percentage against related period (bottom) .................................... 91
Figure 5.3: LST amplitude map for daily component (24 hours period) ............. 92
Figure 5.4: Land cover map of Tibetan Plateau for the 2000 year (http://landcover.usgs.gov/). ................................................................. 92
Figure 5.5: The workflow to estimate the daily amplitude of soil heat flux using the hourly ground measurements of $G_0$ and the instantaneous values of $G_0$ from satellite observations ................................................................. 96
Figure 5.6: Variation of index values for three ground stations ......................... 97
Figure 5.7: Amplitude map of soil heat flux for 24-hour period........................................98
Figure 5.8: The hourly, instantaneous and daily mean LST values. .................................99
Figure 5.9: Estimated amplitude using complete- and incomplete-LST data (left) and corresponding scatter plot (right). ................................................................. 100
Figure 5.10: Soil thermal admittance maps for daily (top) and yearly periods (bottom). ................................................................. ........................................ 101
Figure 5.11: Soil thermal admittance profile of some pixels in the study area.................102
Figure 6.1: Soil vertical profile of a three layered soil: a top layer of thickness $z_1$ and a middle layer of thickness $z_2$ overlay a homogeneous semi-infinite soil. T, temperature; $G$, heat flux; y, thermal admittance.................................................. 104
Figure 6.2: The histogram of STA: period = 3288 hours (left), and 3757 hours (right). ............................................................................................................................................ 111
Figure 6.3: Spectral soil thermal admittance using synthetic data.................................112
Figure 6.4: SSTA calculated using ground measurements at BJ (left) and using satellite data (right). ............................................................................................................................................ 113
Figure 6.5: The workflow showing the procedure to evaluate the capability of the inversion method to minimize the residuals between the observed and simulated SSTA. ............................................................................................................................................ 117
Figure 6.6: Comparison between observed SSTA from ground measurements against SSTA simulated using the retrieved soil thermal properties.................................118
Figure 6.7: The workflow for comparing the observed and simulated SSTA at BJ station. ............................................................................................................................................ 119
Figure 6.8: The SSTA estimated by model 1 using ground measurements of $G_0$ and LST at BJ station (blue line), by model 2 using the measured soil thermal properties (red line), by model 2 using estimated soil thermal conductivity from De Vries model (green line) and, by model 2 using estimated soil thermal conductivity from Yang model (method 3) (violet line); Mean daily STA for 12- and 24-hour periods. ............................................................................................................................................ 122
Figure 6.9: The SSTA estimated from the time series ground measurements and satellite data of $G_0$ and LST applying model 1................................................................. 123
Figure 6.10: The estimated and measured LST (left) and $G_0$ (right). .........................124
Figure 6.11: The estimated and measured daily amplitudes of LST (left) and $G_0$ (right). ............................................................................................................................................ 125
Figure 6.12: Measured vs. estimated daily soil thermal admittance for the 24-hour period. ................................................................. 125

Figure 6.13: Sensitivity of spectral soil thermal admittance assuming three-layer soil across all frequencies. ................................................................. 126

Figure 6.14: The workflow used to validate the inversion method. ................................. 127

Figure 6.15: Scatter plot of the retrieved and calculated soil thermal properties. .......... 128

Figure 6.16: Histogram of soil thermal conductivity of the first layer (λ₁) before (left) and after (right) smoothing. .............................................................................. 129

Figure 6.17: Histogram of soil heat capacity of the first layer (ρc₁) before (left) and after (right) smoothing. .............................................................................. 129

Figure 6.18: Soil thermal conductivity map of first (λ₁), second (λ₂) and third (λ₃) layer (left); soil heat capacity map of first (ρc₁), second (ρc₂) and third (ρc₃) layer (right) of a three-layer soil in Tibetan Plateau. .................................................. 133

Figure 6.19: The first layer soil thickness map (z₁). ......................................................... 134

Figure 6.20: The calculated and simulated classified map of SSTA (left); The MAE and RMSE between calculated and simulated SSTA profile (right). ............................. 134
List of Tables

Table 2.1: Spectral specification of the Single channel Visible and Infrared Spin Scan Radiometer (S-VISSR) sensor................................................................. 13
Table 2.2: Ground stations and their locations in Tibetan Plateau. ......................... 16
Table 2.3: Missing images during each month from 2008 to 2010. ........................... 18
Table 3.1: Defined parameters to apply the new and old version of HANTS. ............... 30
Table 3.2: Assigned parameters used with the new HANTS algorithm for three days LST time series and obtained RMSE values.............................................. 38
Table 3.3: Amplitude, power value and related frequencies of FFT test for a pixel. ..... 39
Table 3.4: Different parameterizations used by the old version of HANTS algorithm for gap-filling of one-month LST time series and obtained MAE and RMSE values. ........................................................................................................ 40
Table 3.5: Different values of parameters tested in the new version of HANTS software applied on one-month LST time series. ................................................. 41
Table 3.6: The HANTS parameters used for gap-filling of LST ground measurements during January 2008 and obtained goodness of fit values of RMSE, $R^2$, bias and MAE................................................................. 42
Table 3.7: The obtained goodness of fit values of RMSE, $R^2$, bias and MAE showing the HANTS ability for gap-filling (left) and reconstruction of signal (right). ... 43
Table 4.1: The statistical metrics values of estimated and observed LST ground measurements as a function of the number of components (left), and as a function of the window size (right)........................................................................ 65
Table 6.1: The thermal properties of some soil substances (De Vries, 1963). .......... 106
Table 6.2: The basic statistics of STA (Wm$^{-2}$k$^{-1}$) for each period and the entire study area.................................................................................................................. 111
Table 6.3: The total percentage of STA values at each frequency lower than the threshold indicated in the bottom row. ......................................................... 111
Table 6.4: Synthetic soil thermal properties assuming a three-layered soil. .......... 112
Table 6.5: Input parameters of the three-layered soil thermal admittance model. .... 115
Table 6.6: The measured soil organic carbon mass content, porosity, soil gravel mass, bulk density, and soil texture at BJ station (Chen et al. 2012)................. 120
Table 6.7: The yearly average of volumetric soil water content at BJ station during 2008-2010. ........................................................................................................ 120
Table 6.8: The soil parameters in different soil layers at the BJ station. ....................... 121
Table 6.9: The statistical measures of goodness of fit between SSTA estimated by model 1 and model 2. .................................................................................................................. 123
Table 6.10: The calculated and estimated soil thermal properties for a three layers soil. ........................................................................................................................................... 128
Table 6.11: Basic statistics of soil thermal properties maps and soil depth map. ......... 130
Table 6.12: The retrieved soil thermal conductivity using the inversion model and estimated soil thermal conductivity for a dry and saturated soil. ....................... 131
### List of Symbols

- $c$: Specific heat (Jkg$^{-1}$K$^{-1}$)
- $d$: Damping depth (m)
- $f_c$: Fractional vegetation cover (-)
- $g$: Shape factor of soil constituents (-)
- $G_0$: Soil heat flux (Wm$^{-2}$)
- $H$: Sensible heat flux (Wm$^{-2}$)
- $k_e$: Kersten number (-)
- $k_T$: Empirical coefficient (0.36)
- $m_g$: Gravel mass proportion (%)
- $m_{soc}$: Soil organic carbon mass content (%)
- $R_n$: Net Radiation (Wm$^{-2}$)
- $S_r$: Saturation degree (-)
- $y_0$: Surface soil thermal admittance (Wm$^2$K$^{-1}$)
- $y_1$: Thermal admittance of soil first layer (Wm$^2$K$^{-1}$)
- $z$: Soil thickness (m)
- $\alpha$: Thermal diffusivity (m$^2$s$^{-1}$)
- $\Delta G_0$: Soil heat flux amplitude (Wm$^{-2}$)
- $\varepsilon$: Emissivity (-)
- $\Gamma_c$: Ratio of soil heat flux to net radiation for full vegetation canopy (0.05)
- $\Gamma_s$: Ratio of soil heat flux to net radiation for bare soil (0.315)
- $\theta_{sat}$: Volumetric saturated water content (%)
- $\theta_m$: Porosity of mineral soils (%)
- $\lambda E$: Latent heat flux (Wm$^{-2}$)
- $\lambda$: Thermal conductivity (Wm$^{-1}$K$^{-1}$)
- $\lambda_{sat}$: Saturated bulk soil thermal conductivity (Wm$^{-1}$K$^{-1}$)
- $\lambda_{dry}$: Dry bulk soil thermal conductivity (Wm$^{-1}$K$^{-1}$)
- $\lambda_a$: Thermal conductivity of air (0.025 Wm$^{-1}$K$^{-1}$)
- $\lambda_w$: Thermal conductivity of water (0.6 Wm$^{-1}$K$^{-1}$)
- $\lambda_s$: Thermal conductivity of soil solid phase (Wm$^{-1}$K$^{-1}$)
- $\lambda_q$: Thermal conductivity of quartz (7.7 Wm$^{-1}$K$^{-1}$)
- $\rho$: Bulk density (kg m$^{-3}$)
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_m )</td>
<td>Bulk density of the soil mineral phase (kg m(^{-3}))</td>
</tr>
<tr>
<td>( \rho_o )</td>
<td>Bulk density of peat (130 kgm(^{-3}))</td>
</tr>
<tr>
<td>( \rho_p )</td>
<td>Mineral particle density (2700 kgm(^{-3}))</td>
</tr>
<tr>
<td>( \phi )</td>
<td>Phase (Radian)</td>
</tr>
<tr>
<td>( \omega )</td>
<td>Angular frequency (cycle/period)</td>
</tr>
</tbody>
</table>
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGS</td>
<td>Continuous Gap Size</td>
</tr>
<tr>
<td>DOD</td>
<td>Degree of Over Determinedness</td>
</tr>
<tr>
<td>EOF</td>
<td>Empirical Orthogonal Function</td>
</tr>
<tr>
<td>ET</td>
<td>Evapotranspiration</td>
</tr>
<tr>
<td>FET</td>
<td>Fit Error Tolerance</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FY-2C</td>
<td>Feng-Yun-2C</td>
</tr>
<tr>
<td>GSW</td>
<td>Generalized Split-Window</td>
</tr>
<tr>
<td>HANTS</td>
<td>Harmonic ANalysis of Time Series</td>
</tr>
<tr>
<td>K</td>
<td>Kelvin degree</td>
</tr>
<tr>
<td>LSE</td>
<td>Land Surface Emissivity</td>
</tr>
<tr>
<td>LST</td>
<td>Land Surface Temperature</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>MODIS</td>
<td>MODe rate resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>M-SSA</td>
<td>Multi-channel Singular Spectrum Analysis</td>
</tr>
<tr>
<td>NDVI</td>
<td>Normalized Difference Vegetation Index</td>
</tr>
<tr>
<td>NG</td>
<td>Number of Gap pattern</td>
</tr>
<tr>
<td>NOF</td>
<td>Number Of Frequencies</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>QTP</td>
<td>Qinghai-Tibet Plateau</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SA</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>SSA</td>
<td>Singular Spectrum Analysis</td>
</tr>
<tr>
<td>SSTA</td>
<td>Spectral Soil Thermal Admittance</td>
</tr>
<tr>
<td>S-VISSR</td>
<td>Single channel Visible and Infrared Spin Scan Radiometer</td>
</tr>
<tr>
<td>TNL</td>
<td>Total Number of Losses</td>
</tr>
<tr>
<td>TOA</td>
<td>Top Of Atmosphere</td>
</tr>
</tbody>
</table>
Chapter 1

1. General introduction

1.1. Introduction

The Earth’s surface and the atmosphere form a complex and dynamic system in which the matter and energy are being continuously exchanged between the elements of the Earth System at different temporal and spatial scales. Assuming the incoming solar energy at the top of the atmosphere as 100 units, around 46% of this energy passes through the atmosphere and is absorbed by the earth surface which is responsible for all the energy circulation on the surface (Trenberth et al., 2009). The rest is either absorbed or reflected by clouds, water vapour and aerosols in the atmosphere. The absorbed part by the earth surface is referred to as “net radiation” ($R_n$). To balance the difference between incoming and outgoing solar energy, three different processes take place at the surface (Iqbal, 1983; Pinker and Ewing, 1985; Sellers et al., 1990; Sellers, 1965; Wijk, 1964). A fraction of the absorbed energy is used for Evapotranspiration (ET) which is the “latent heat flux” ($\lambda E$), a portion is released directly by the surface and heat the atmosphere and is the “sensible heat flux” (H) and the remainder is conducted into the ground or water, the “soil heat flux” ($G_0$). The net radiation must always be equal to the sum of these three processes (first law of thermodynamics) (Figure 1.1). The surface energy balance formula when we consider instantaneous fluxes reads:

$$R_n = \lambda E + H + G_0$$

where $R_n$ is net radiation, $G_0$ is soil heat flux and $H$ is sensible heat flux all in (Wm$^{-2}$). $E$ is the rate of evaporation of water (kg m$^{-2}$ s$^{-1}$), and $\lambda$ is the latent heat of vaporization of water (J kg$^{-1}$).
Figure 1.1: Global energy balance between the earth and the atmosphere (http://www.noaa.gov/).

The soil heat flux is usually estimated as a portion of net radiation depending on the fractional vegetation cover (Bastiaanssen and Bandara, 2001; Choudhury et al., 1986; Murray and Verhoef, 2007). Compared to the other terms in Eq. 1.1, the soil heat flux contribution is generally small, and it is ignored in some cases. However, when the incoming and outgoing energy during a day and night are significant, neglecting the soil heat flux would lead to considerable errors in the estimation of short term (i.e. hourly) surface energy balance (Sauer and Horton, 2005). Soil heat flux effectively combines the energy transfer processes at the earth surface (surface energy balance) with the soil thermal regime, which leads to a change in Land Surface Temperature (LST).

Monitoring the exchange of available energy ($R_n - G_0$) between the land surface and the atmosphere (i.e. the latent and sensible heat flux) is essential to understand land surface processes and develop parameterization schemes in weather and climate models, water resource management, and for agricultural applications such as irrigation scheduling (Bowen, 1926; Brutsaert, 1982; Famiglietti and Wood, 1994; Menenti, 1984; Monteith, 1965; Morton, 1983; Penman, 1948; Priestley, 1972; Su, 1999). Accurate estimation of surface energy fluxes is a critical factor in hydrological modelling, weather forecasting models and agro-system models.

Soil heat flux varies over time from daily to seasonal periods. The ability of soil to store and to conduct heat determines how fast the LST changes during the day or over the seasons. For instance, diurnal LST variation is the consequence of daily variation of soil heat flux which, in turn, is mainly caused by solar intensity changes during a day. Longer – term variations of soil heat flux can be influenced by several factors such as solar intensity variation due to seasonality, rainfall, vegetation coverage and soil
moisture, etc. Therefore, the interaction between LST and soil heat flux over a long period is more complicated than on a daily basis (short term). The LST variation over time, during heat storage and release by soil, can be described by a combination of periodic functions (Carslaw and Jaeger, 1959; Wijk and De Vries, 1966). Although, it might not always be a pure periodic signal, there exist features due to the superposition of periodic functions (i.e. daily and yearly). This implies that time series of LST and soil heat flux can be described by combinations of different components of a periodic function (e.g. sine and cosine with various amplitude and phase at different frequencies). These significant periodic components can be identified to quantify the variation of soil heat flux over time by fitting a Fourier series (harmonic analysis) to the actual time series.

It is well understood that there is an explicit relationship between LST and soil heat flux. This relationship can be found as a travelling wave solution of the differential equation describing soil heat flow and lead to define the soil thermal admittance (Carslaw and Jaeger, 1959; Menenti, 1984). Soil thermal admittance is, in fact, a property of materials that is a measure of “impedance” of the earth surface to a given heat flux. The daily soil thermal admittance can be either estimated as the ratio of the daily amplitude of soil heat flux to the daily amplitude of LST or calculated using soil thermal properties (if they are available) (Coppola et al., 2007; Menenti, 1984; Wijk and De Vries, 1966). In the case both time series of LST and soil heat flux are more complex signals, they can be modelled by using a Fourier series with a few more terms (significant periodic components). The extension of soil thermal admittance concept can be used to estimate it at any frequency (period) (Menenti, 1984).

1.2. General problems

To estimate the soil thermal admittance, the LST and soil heat flux time series have to be decomposed into periodic components (by Fourier analysis) and the ratio of the amplitude of soil heat flux and LST at the same frequency calculated to obtain the frequency dependent soil thermal admittance. This spectral soil thermal admittance can be inverted to estimate soil thermal properties of different soil layers by inverse modelling. To obtain the soil thermal admittance at different frequencies (periods), time series data of LST and soil heat flux over longer period for instance one-year hourly time series data are required (e.g. as captured by geostationary satellite data). The LST
time series can be retrieved from thermal infra-red remote sensing, and then LST can be used to estimate other components of the land surface energy balance such as soil heat flux and net radiation.

Geostationary satellites can provide time series data with a sampling time from 15 minutes to an hour. These time series are commonly prone to missing data, noise and outliers (spatially and temporally) due to clouds, aerosols and algorithm malfunctioning which causes irregular observations. Missing data (gaps) means here no valid surface observations due to cloud coverage or failure of retrieval. A lot of effort has been put to detect and mask cloudy satellite image data (Henderson-Sellers, 1982; Saunders and Kriebel, 1988; Simpson and Gobat, 1996; Stowe et al., 1991) but the final result of those methods is image data on the surface with gaps and outliers. Gaps are characterized by their size, distribution and continuity, which can range from very short sparse gaps to long continuous gaps. Outliers are defined as abnormal values that deviate from the normal variability in the dataset. They are grouped into two categories of positive and negative outliers. Positive outliers may be caused by several reasons, for instance, sensor malfunctioning or retrieval algorithm failure. In this case, the measured value is either much higher than acceptable value for the variable being observed (e.g. NDVI > 1) or not comparable to nearby values in time (i.e. sudden increase of a LST value in a temporal profile consisting of hourly values of LST in a day). In thermal remote sensing, gases, clouds and atmospheric aerosols absorb part of the thermal energy emitted by the earth. They also emit thermal infrared energy at a much lower temperature than the underlying ground, so when the cloud masking algorithm does not detect clouds correctly, there will be some negative outliers instead of gaps in a LST time series.

1.2.1. Gaps and outliers in hourly LST time series data

Having consistent and continuous time series requires filling the gaps and identifying and eliminating the outliers from the original time series data. Removing the outliers from time series makes the problem worse by adding new gaps that should be filled. A number of approaches has been proposed to deal with gap filling and outlier removal from time series over the last decades (Amisigo and van de Giesen, 2005; Fang et al., 2007; Jia et al., 2011; Julien et al., 2006; Kondrashov et al., 2010; Menenti et al., 1993; Moffat et al., 2007; Roerink and Menenti, 2000a; Verhoef, 1996). These
approaches are mainly based on the information content which may exist in the time series (e.g. periodic time series). Such approaches, often referred to as temporal approaches, work reasonably well where the gaps are not very long and continuous. This is because the intrinsic periodicity of the time series can be exploited by these models using well-known algorithms, for instance, Fourier series analysis. When gaps are long and continuous, however, this information content may not be sufficient to yield an accurate reconstruction of the time series (Jia et al., 2011; Verhoef et al., 2005). Therefore, the application to datasets presenting long and continuous gaps has to be explored and evaluated.

An alternative is to use both the spatial and the temporal information content of the datasets together. There exist areas that present analogous temporal patterns scattered across the image dataset. Exploring these similarities can be used as a prior knowledge to constrain and help addressing the problem of gap filling. The model is fed by this prior knowledge as first guess and searches for parameters that fit best iteratively. However, despite their promising potential in gap filling, these models are computationally intensive and require considerable resources to fill long and continuous gaps in large image windows.

1.2.2. Estimation of soil thermal admittance at significant frequencies

To calculate the soil thermal admittance, the amplitude of soil heat flux and LST at same frequency are needed. Assuming a periodic behaviour, the LST and soil heat flux time series can be modelled as some harmonic functions of time around an average value. The actual time series, then can be decomposed to identify most significant periodic components that account for the variability in the time series. Having those periodic components, the related amplitudes, i.e. half of the difference between the maximum and minimum value, can be estimated. This can be achieved through fitting a Fourier series to the time series data (e.g. hourly LST) to estimate the amplitude and phase at significant frequencies. Figure 1.2, for instance, shows one-year hourly LST time series (blue line), yearly mean, and the smooth yearly component (red line) superimposed on actual data. In order to calculate the amplitude (of soil heat flux and LST) at significant frequencies, assuming periodic behaviour, we need continuous time series with sufficient temporal resolution (e.g. hourly). When the time series
contains only an instantaneous and a mean daily value (i.e. the soil heat flux data in this study), the estimation of amplitudes at dominant frequencies is troublesome.

![Yearly LST time series, related yearly component and amplitude.](image)

Soil thermal admittance is related to the surface and subsurface characteristics (e.g. soil compositions, texture), and near surface soil moisture (Carlson et al., 1981; Menenti, 1984; Palluconi and Kieffer, 1981). It controls the LST fluctuation, for instance, the daily or the yearly amplitude of LST. The amplitude of LST is a function of soil thermal properties at different depths, soil water content, albedo of the surface, surface roughness and meteorological variables (e.g. amount of received solar radiation, air temperature, wind, etc). Assuming a constant daily soil heat flux amplitude at a given surface, the soil thermal admittance difference is mainly controlled by soil water content (Xue and Cracknell, 1995). In that case, High (low) soil thermal admittance gives a small (high) daily LST amplitude.

### 1.2.3. Estimation of soil thermal properties via inverse modelling of spectral soil thermal admittance

Given a uniform soil profile, there is a clear analytical relationship between the soil thermal properties, LST and soil heat flux (Menenti, 1984). The soil thermal properties often vary by the solid soil fractions (i.e. mineral type, particle size, and organic matter), soil water content, and soil bulk density (Al Nakshabandi and Kohnke,
1965; De Vries, 1963; Farouki, 1981). These properties also change both spatially and temporally between soils and within soil layers. The soil thermal properties have a great influence on the partitioning of energy at the earth surface and determine the land surface temperature and exchange of soil heat flux across the soil profile. The soil thermal properties are conventionally measured in situ, though this is very resource demanding and costly and impossible at regional to global scales. Spectral soil thermal admittance is defined as the soil thermal admittance at different frequencies and contains information about soil thermal properties at different depths (Menenti, 1984). This implies that in principle such soil thermal properties at different depths can be retrieved from the spectral soil thermal admittance. To retrieve these properties a forward model (e.g. a physical model), simulating the spectral soil thermal admittance, has to be inverted against the observed values which can be obtained from time series of satellite observations of LST and soil heat flux. The number of soil thermal properties than can be retrieved by inversion of the forward model depends on the number of independent (non-correlated) observations. If the number of independent observations is less than the number of unknowns, the inversion becomes ill-posed. This implies that either the solution does not exist or it is not unique, or it is not continuous with the model conditions. To overcome the ill-posedness, we can regularize the problem by narrowing the solution space or using prior information about the solution. The prior information may be obtained by performing a sensitivity analysis quantifying the influence of model inputs on the model output. Sensitivity analysis will help identifying which variables may be retrievable and which ones can be fixed to a value in their domain without affecting the results significantly.

1.3. General objective

The main objective of the present investigation is to develop an approach to estimate soil thermal properties using time series satellite observations of LST and soil heat flux. The soil thermal properties are estimated at different soil depths through the inversion of spectral soil thermal admittance. To accomplish this, the following specific objectives will be addressed:
1- Gap-filling and outlier removal of hourly satellite observations of LST.
2- Estimate the amplitude of LST and soil heat flux at significant periodic components.
3- Estimate the soil thermal admittance at different frequencies to derive the spectral soil thermal admittance.
4- Estimate the soil thermal properties at different soil depths by inversion of the forward model.

1.4. Overview of approach

Figure 1.3 illustrates the workflow of the approach implemented in this study. The figure shows three major inter-related steps in which each step is addressed in a separate chapter. The first step regards data collection and pre-processing of raw data to generate gap-free time series data based on gap distribution (short and long gaps). This step is considered as the prerequisite for decomposing the LST and soil heat flux time series into periodic components. To do so, we need first to identify and remove the outliers and then fill the available gaps in the data as well as those generated by outlier removal. The HANTS (Menenti et al., 1993; Roerink and Menenti, 2000a; Sellers et al., 1994; Verhoef, 1996) and the M-SSA (Broomhead and King 1986a, 1986b; Broomhead et al. 1987; Elsner and Tsonis 1996) algorithms were used for gap-filling. In the HANTS algorithm, the periodic components are prescribed whereas the M-SSA algorithm does not require any prior assumptions (e.g. periodicity). However, due to the fact that the LST time series are periodic, the signal is reconstructed with periodic functions by M-SSA. The M-SSA works regardless of the nature of the time series (periodic or non-periodic) where a non-periodic time series, for instance, would not result in periodic components at the end. As the LST time series is relatively periodic the final results of both methods are similar. The second step is the estimation of amplitudes of LST and soil heat flux at dominant frequencies to obtain the spectral thermal admittance of the land surface. The last step is to invert the spectral thermal admittance to retrieve the soil thermal properties.
1.5. **Research questions**

- How to fill the gaps and identify and remove outliers from hourly time series satellite observations of LST to reconstruct continuous and gap-free data?
- How to estimate the amplitude of LST time series and soil heat flux at significant frequencies?
- How to estimate the spectral soil thermal admittance?
- How to estimate soil thermal properties by inversion of spectral soil thermal admittance?

1.6. **Outline of the thesis**

This research is composed of seven chapters, and its structure is briefly described below. **Chapter 1**, Introduction; a brief problem definition, research hypotheses, research questions and objectives. **Chapter 2** gives a brief overview of the study area, i.e. the Qinghai-Tibet Plateau and its importance for implementing the approach. Also, the data used in the study are described with a brief explanation about each sensor and data type with their specifications. **Chapter 3**, Gap-filling of time series satellite data of LST is described in this chapter with a focus on short gaps. **Chapter 4**, Gap-filling of long continuous gaps is explored in this chapter, and the evaluation of results is also addressed using time series of LST ground measurements. **Chapter 5** focuses on the estimation of soil thermal admittance over whole study area. **Chapter 6** describes the inversion of spectral soil thermal admittance to retrieve the soil thermal properties at different soil depths. **Chapter 7**, In this chapter, the final conclusions are drawn and recommendations listed for possible directions for future works.
Figure 1.3: Overview of the approach developed in this thesis to estimate the soil thermal properties using time series satellite observations of LST and soil heat flux.
Chapter 2

2. Description of the study area and data sets

2.1. Introduction on the CEOP-AEGIS project

CEOP-AEGIS which stands for "Coordinated Asia-European long-term Observing system of Qinghai–Tibet Plateau hydro-meteorological processes and the Asian-monsoon systEm with Ground satellite Image data and numerical Simulations" is a collaborative medium-scale focused research project financed by the European Commission under FP7 topic ENV.2007.4.1.4.2 "Improving observing systems for water resource management". It is motivated to support water resources management in South-East Asia (http://ceop-aegis.org/). The CEOP-AEGIS project was established to achieve the following goals:

- To integrate the ground measurements and satellite observations to deliver a prototype water monitoring system containing three years (2008-2010) time series data sets of water balance terms on different temporal resolutions from hourly to monthly.
- To observe the progress of water and snow coverage change, vegetation cover, soil surface moisture content and surface energy fluxes in order to analyse land-atmosphere interactions influencing the Asian Monsoon System (Yanai et al., 1992; Ye and Wu, 1998).

The CEOP-AEGIS lasted 60 months, and it built upon ten years of experimental and modelling research on the Tibetan Plateau carried out by a consortium of eighteen partners from eight countries (Menenti et al., 2009). Both data and modelling we addressed by the project using in-situ measurements, satellite observation of land surface properties and fluxes, soil moisture, precipitation, snow cover and water equivalent, modelling of land surface-atmospheric interactions, hydrological modelling
and remote sensing indicators useful for flood and drought early warning. The study described in this dissertation contributes to this project by focusing on land surface temperature and soil thermal properties. In the next section, first the study area of Tibetan Plateau and then the satellite data and ground measurements used in this dissertation will be described.

2.2. Study area

The study area is the Qinghai-Tibet Plateau (QTP), which is the highest plateau in the world, located in East Asia (Figure 2.1). The Tibetan Plateau covers a large area from subtropical to middle latitudes and 25 degrees of longitude and plays a significant role in the Asian Monsoon system (Ma et al., 2005). The Tibetan Plateau lies between the Himalaya Mountains to the South and the Taklimakan Desert to the North. The top-left corner of the study area is 39°19’39.37”N, 64°12’12.26”E and the bottom-right corner is 24°51’12.62”N, 107°2’48.11”E. The selected study area occupies an area of around 7.5 million square kilometers (~2000 km north to south and ~3500 km from east to west) at a mean elevation of 4,500 meters.

The monsoon climate in the river basins downstream of the TP, where around 40% of the world population lives, has a serious effect on human life and ecosystems of South and East Asia (Rasul, 2014). The estimation of water and energy balance in that region is of importance to predict the behaviour of those effects. Field observations on water and energy balance components cannot provide the required spatial coverage, temporal frequency and accuracy. So satellite observations of land surface properties, in combination with ground observations, are required to provide the necessary information on water resources and the interaction between the earth’s surface and atmosphere in Asian monsoon system.

2.3. Satellite data

In this study, we used Land Surface Temperature (LST) and soil heat flux estimated from the radiometric data collected by the Single channel Visible and Infrared Spin Scan Radiometer (S-VISSR) sensor on-board the Fengyun-2C (FY-2C) geosynchronous meteorological satellite (NSMC, 2012). The spectral configuration of S-VISSR is described in Table 2.1. The spatial (at the latitude of plateau) and temporal resolutions are 5×5 kilometers (totally 708 columns × 408 rows pixels) and hourly respectively. The data used in this study span a three years period (2008-2010) covering
the Tibetan Plateau and surrounding lower regions (26304 LST values for each pixel).

The time series satellite data are often contaminated by clouds causing gaps, especially when the data have high temporal resolution covering a large area with a high chance of cloud cover like the Qinghai-Tibet Plateau. Therefore, before starting any data analysis, we have to do some pre-processing to identify and remove non-pertinent information and fill the gaps in the time series data. This has been done with a procedure developed by Ghafarian et al. (2012) to get hourly gap-free data set over three years (see Chapter 3 and 4).

![Figure 2.1: Tibetan Plateau (http://www.zonu.com).](image)

**Table 2.1: Spectral specification of the Single channel Visible and Infrared Spin Scan Radiometer (S-VISSR) sensor.**

<table>
<thead>
<tr>
<th>Channel ID</th>
<th>Channel name</th>
<th>Spectral range(µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR1</td>
<td>Long wave infrared</td>
<td>10.3 – 11.3</td>
</tr>
<tr>
<td>IR2</td>
<td>Split window</td>
<td>11.5 – 12.5</td>
</tr>
<tr>
<td>IR3</td>
<td>Water vapor</td>
<td>6.5 – 7.0</td>
</tr>
<tr>
<td>IR4</td>
<td>Medium wave infrared</td>
<td>3.5 – 4.0</td>
</tr>
<tr>
<td>VIS1–4</td>
<td>Visible</td>
<td>0.55 – 0.90</td>
</tr>
</tbody>
</table>
2.3.1. Retrieval of emissivity and LST using FY-2C observations

The Land Surface Emissivities (LSEs) and LST of FY-2C observations were generated in the framework of the CEOP-AEGIS project by Tang and Li (2011) and Tang et al. (2008). There are many different algorithms to estimate LST from satellite observations which can be categorized broadly into three different approaches: the single channel algorithm (Jiménez-Muñoz and Sobrino, 2003; Ottlé and Vidal-Madjar, 1992; Wan and Dozier, 1989), the split window algorithm (Becker and Li, 1990; McMillin, 1975) and the triple windows algorithm (Sun and Pinker, 2003). The generalized split window algorithm was used to retrieved LST from the S-VISSR infrared channels. The theoretical basic algorithm used to retrieve the LST will be briefly described in the next section (Tang et al., 2008).

2.3.2. Retrieval algorithm of LST and LSE

As mentioned above, the split windows algorithm was used to estimate the LST based on the differential water vapor absorption in two different adjacent infrared channels (McMillin, 1975). The S-VISSR sensor onboard FY-2C has two neighboring thermal infrared channels (IR1 and IR2) (Table 2.1), the Generalized Split-Window (GSW) algorithm proposed by Wan and Dozier (1996) was adapted to estimate the LST from FY-2C satellite data (Tang et al., 2008). According to the GSW algorithm, the land surface temperature (LST) can be expressed as:

$$T_s = a_0 + \left( a_1 + a_2 \frac{1-\varepsilon}{\varepsilon} + a_3 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{T_{IR1} + T_{IR2}}{2} + \left( a_4 + a_5 \frac{1-\varepsilon}{\varepsilon} + a_6 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{T_{IR1} - T_{IR2}}{2}$$

with $\varepsilon = (\varepsilon_{IR1} + \varepsilon_{IR2}) / 2$ and $\Delta \varepsilon = \varepsilon_{IR1} - \varepsilon_{IR2}$

where $T_{IR1}$ and $T_{IR2}$ are the Top Of Atmosphere (TOA) brightness temperatures measured in channels IR1 (11.0 $\mu$m) and IR2 (12.0 $\mu$m), $\varepsilon_{IR1}$ and $\varepsilon_{IR2}$ are LSEs in channels IR1 and IR2; $\varepsilon$ is the average emissivity; $\Delta \varepsilon$ is the emissivity difference between the two adjacent channels; and the coefficients $a_0 - a_6$ proposed by Tang et al. (2008) were adopted.

LSEs ($\varepsilon_{IR1}$ and $\varepsilon_{IR2}$) of S-VISSR were estimated from the MODIS LSEs product (MOD11B1) of channels 31 (10.78-11.28 $\mu$m) and 32 (11.77-12.27 $\mu$m) (Tang et al., 2008). To do so, the emissivities in the two split-window channels of MODIS ($\varepsilon_{31}$ and $\varepsilon_{32}$) and S-VISSR ($\varepsilon_{IR1}$ and $\varepsilon_{IR2}$) were calculated as the convolution integral of the spectral emissivity with the channel response functions over the spectral range of the
channels (Tang et al., 2008). Then, a linear regression relationship between the emissivities in S-VISSR channels ($\varepsilon_{IR1}$ and $\varepsilon_{IR2}$) and in MODIS channels ($\varepsilon_{31}$ and $\varepsilon_{32}$) were established as follows:

$$\varepsilon_{IR1} = 0.0608 + 0.9356\varepsilon_{31} \quad (2.2)$$

$$\varepsilon_{IR2} = 0.1325 + 0.8611\varepsilon_{32} \quad (2.3)$$

### 2.3.3. Retrieval algorithm of soil heat flux

The other data set was the time series of instantaneous and daily mean soil heat flux ($G$) from 2008 to 2010. The instantaneous value of soil heat flux (at 14 Pm Beijing local time) and daily mean value over three years (each 1096 values per pixel) have been estimated using the energy balance equation (Eq. 1.1) and cover the entire study area (Faivre, 2014). The soil heat flux is often parameterized proportionally to the net radiation ($R_n$) arriving at the soil surface, therefore it is a function of the fractional vegetation cover (Bastiaanssen and Bandara, 2001; Choudhury et al., 1986; Murray and Verhoef, 2007). It can be expressed as:

$$G_b = R_n(\Gamma_c + (1 - f_c)(\Gamma_s - \Gamma_c)) \quad (2.4)$$

in which it is assumed that the ratio of soil heat flux to net radiation is $\Gamma_c = 0.05$ for a full vegetation canopy (Monteith, 1973) and $\Gamma_s = 0.315$ for bare soil (Kustas and Daughtry, 1990). An interpolation is then performed between these limiting cases using the fractional vegetation cover ($f_c$).

### 2.4. Ground station data

There are also four ground stations indicated as yellow stars in Figure 2.1 and their land cover maps are shown in Figure 2.7, where the soil heat flux has been measured every 30 minutes in 2008-2010 (Babel et al., 2011). The station names and their locations are shown in Table 2.2. The soil heat flux is calculated from ground measurements of surface temperature, soil temperature and soil moisture profiles using the method of Yang and Wang (2008). Among the available stations, we used the BJ station data which had a complete three years data series.
2.4.1. Nagqu (BJ) station

The BJ site (Nagqu station) is located in the Eastern Tibetan Plateau (Figure 2.1). It lies in a nearly flat area within a vast valley at 4502 m height. The vegetation cover is pasture and bare soil with a rather homogeneous landscape. It is surrounded by hill slopes from 1 km distance to the northeast to several km to the south and west. There is a small creek to the northeast of the station that goes together with wetland vegetation (see Figure 2.7).

Table 2.2: Ground stations and their locations in Tibetan Plateau.

<table>
<thead>
<tr>
<th>Station</th>
<th>E</th>
<th>N</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nagqu (Beijing)</td>
<td>-279574.1757</td>
<td>165257.0112</td>
<td>31° 20’ 57.59”N</td>
<td>91° 50’ 6.60”E</td>
</tr>
<tr>
<td>Nam Co</td>
<td>-364574.1757</td>
<td>95257.0112</td>
<td>30° 44’ 11.74”N</td>
<td>90° 54’ 24.95”E</td>
</tr>
<tr>
<td>Linzhi</td>
<td>-24574.1757</td>
<td>-29742.9888</td>
<td>29° 44’ 57.69”N</td>
<td>94° 43’ 40.21”E</td>
</tr>
<tr>
<td>Qomo</td>
<td>-744574.1757</td>
<td>-164742.9888</td>
<td>28° 21’ 1.67”N</td>
<td>86° 53’ 44.55”E</td>
</tr>
</tbody>
</table>

2.5. Exploration of the LST time series data

In order to get an overview of hourly LST time series data over the QTP, we explored the data set by mapping the mean LST variation over long periods (e.g. one year), percentage of missing data spatially and temporally, gap percentage in time series, maximum duration of continuous gaps. This data exploration helped us to understand better the variability of LST data over TP, the amount of gaps and outliers. As mentioned before, the hourly LST data are available from January 2008 until December 2010. Here we just used the data of November 2008 and the whole 2008 year.

2.5.1. Mean LST over study area during November 2008 and 2008 year

The mean LST values during November 2008 and the whole 2008 year in each pixel of the QTP have been plotted to visualize the spatial LST pattern (Figure 2.2). The maps show that there are two distinct areas: the southwest part of the QTP has a higher mean temperature than the other parts. The central part of the QTP has mean LST below 0 °C while the other parts are just above 0 °C.

The histogram of mean LST confirms the distinction of two areas having different LST mean values as there are two visible peaks in the graphs with one below and the other one above 0 °C (Figure 2.3.a,b). The logarithmic representation of the
histogram points clearly to the existence of some outliers in the data set. They appear as distinct peaks at the two tails of the histogram (Figure 2.3.c,d) assuming a valid data range between 220 K and 340 K.

Figure 2.2: Mean LST values in Tibetan Plateau during November 2008 (a) and 2008 year (b).

Figure 2.3: Histogram of mean LST data during November 2008 and 2008 year in TP.
2.5.2. Spatial extent of gaps

To determine the spatial distribution of gaps over the QTP during the November 2008 (consisting of 24 hours × 30 days = 720 images covering the TP) and 2008 (366 days × 24 hours = 8784 images covering the TP), the percentage of gaps (missing pixels) in each hourly LST image covering the whole QTP has been calculated (Figure 2.4). In November 2008, the missing pixels range from 10.4% up to 62.6% and from 1.4% to 95.6% in 2008. The yearly plot included some 100% values that belong to missing images during 2008. Table 2.3 shows the number of missing images in each month during 2008-2010. November 2008 and June 2010 have no missing images while April 2008 has 106 missing images. The results show that during the days of November 2008, the amount of gaps increased while during the night and afternoon the gaps decreased. It shows that during daytime the clouds covered most of the study area.

![Figure 2.4: Hourly fraction of gaps (missing pixels) for the entire study area.](image)

Table 2.3: Missing images during each month from 2008 to 2010.

<table>
<thead>
<tr>
<th>Period</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>1</td>
<td>14</td>
<td>65</td>
<td>108</td>
<td>3</td>
<td>1</td>
<td>43</td>
<td>9</td>
<td>70</td>
<td>37</td>
<td>0</td>
<td>2</td>
<td>353</td>
</tr>
<tr>
<td>2009</td>
<td>2</td>
<td>54</td>
<td>66</td>
<td>18</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>15</td>
<td>83</td>
<td>29</td>
<td>6</td>
<td>9</td>
<td>288</td>
</tr>
<tr>
<td>2010</td>
<td>1</td>
<td>4</td>
<td>60</td>
<td>28</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>9</td>
<td>61</td>
<td>71</td>
<td>12</td>
<td>45</td>
<td>309</td>
</tr>
</tbody>
</table>
2.5.3. Temporal extent of gaps

The gap-filling procedures usually use the temporal correlation in the data, i.e. the amount of gaps and their distribution in time is critical. In order to see the temporal pattern of missing observation per time-series (in each pixel) in the entire study area, the percentage of gaps with respect to the total number of observations in each pixel has been calculated (Figure 2.5). The percentage of gaps in November 2008 range from 0% to 78.9%, while in 2008, this is from 14.8% to 77%. It implies that in some area (i.e. the Eastern QTP) the amount of gaps in a single pixel time series is more than half of observations in that time series. It shows that the gap-filling procedures are necessary for having continuous and consistent time series.

Figure 2.5: Percentage of gaps in each pixel during November 2008 (a) and the year 2008 (b).

2.5.4. Maximum continuous gap size in LST time series

To illustrate the worse cases of gaps in the LST time series, the percentage of longest continuous gap size in each pixel was calculated (Figure 2.6). To do so, first the length of longest continuous gap pattern in each pixel was found, and then it was divided by the total length of time series. The results show that, for November 2008, the range of maximum gap size changes from 0% to 48.19% (347 hours ~14 days), and for the whole year of 2008 from 0.90% (79 hours) to 4.7% (410 hours ~ 17 days). Those are the cases that will be very challenging for gap-filling procedures only relying on temporal correlation. No information on the signal during those periods is available and the results of reconstruction are not reliable (see Chapter 3).
2.6. Conclusion

The statistical investigation of time series LST data shows that the data has different kinds of gaps based on the time duration, continuity, location and size. Besides, the size of data for hourly time series LST data during 2008-2010 is enormous and large computational and data storage resources are needed to process such data sets.
Figure 2.7: Land use map of the four stations on the Tibetan Plateau (W.BABEL et al., 2011).
Chapter 3

3. Filling short gaps with the HANTS algorithm

3.1. Introduction

The geostationary satellites, which were used in this study, provide frequent temporal observations (i.e. 15 minutes to hourly observations) and allow multiple surface measurements on a daily basis. The quality, spatial and temporal consistency of time series of remotely sensed data (e.g. Land Surface Temperature; LST) are degraded by missing data with different number, size, distribution and continuity of gaps, so that certain areas may only be observable for a very short time (i.e. long continuous gaps). This problem is intensified by the presence of gaps and outliers together. The number of gaps in a time series can be categorized into short or long gaps. A short (long) gap is shorter (longer) than the period of time sampled by half of the observation points (Jia et al., 2011; Roerink and Menenti, 2000a). For example, Figure 3.1 shows two hourly LST time series during one month with 744 observations in which the valid observations and gaps are colored as vertical blue lines and the white spaces between blue lines, respectively. The top time series has in total 65 missing data while the bottom time series has 470 missing data. Therefore, based on above definition, the top (bottom) time series has short (long) gaps.
Time series data can be either (quasi) periodic or non-periodic. Depending on the periodicity of the time series, Fourier theory (Fourier, 1818) (for periodic data) or wavelet theory (Ricker, 1953) (for non-periodic data) can be applied to analyse the time series (Verhoef, 1996). Fourier series analysis or harmonic analysis is a common method in many different applications of earth science especially in time series analysis and has capabilities to enhance our understanding of periodic phenomena (Bloomfield, 2000). Remote sensing time series of LST are among the periodic time series due to the fact that they are influenced by the diurnal and yearly course of the sun and periodic changes in weather conditions, which yield periodic components of the time series. Determining these components is crucial to reconstruct a given periodic time series having gaps and outliers and to understand the response of the land surface to weather and climate forcing.

To fill the short gaps and remove the outliers, a number of methods have been developed and successfully applied in the literature. For instance, a Fast Fourier Transform algorithm (FFT) was applied to reconstruct gapless time series data of Normalized Difference Vegetation Index (NDVI) (Menenti et al., 1993; Sellers et al., 1994; Verhoef, 1996). The FFT algorithm needs data without gaps and equidistant in time. Harmonic ANalysis of Time Series (HANTS) algorithm (Roerink and Menenti, 2000a; Verhoef et al., 1996) was developed to remove the outliers and fill the gaps simultaneously in time series data of NDVI with irregular sampling interval and to extract phenological information. Jia et al. (2009) applied the HANTS algorithm to create gap-filled Evapotranspiration (ET) time series estimated using MODIS data. Julien et al. (2006) utilized HANTS for time series of yearly mean LST to obtain cloud-free time series and climatological information. Their results confirmed the usefulness of this approach for LST time series analysis. HANTS has been used on daily MODIS
The HANTS algorithm was designed to execute two tasks: (i) identifying and removing of outliers and cloudy observations, and (ii) gap-filling of the remaining valid observations by temporal interpolation. Besides that, HANTS can be used to decompose a periodic time series into its components that describe the time series and then identify the significant periodic components. The main conceptual framework of HANTS is based on modelling a time series by calculating a Fourier series. HANTS handles the Fourier analysis as a curve fitting problem using a weighted least squares solution. In the HANTS algorithm, each observation in the time series is assigned a weight of one or zero for good and bad (outliers) data, respectively. In order to find outliers and cloud contaminated observations, HANTS performs the curve fitting iteratively. In the first step, the least square curve fitting is performed using all data in the series. In the second step, observations are compared to the curve determined in the 1st iteration. Observations that deviate more than a pre-defined threshold are removed by assigning zero weight to them. The remaining data are used to compute the least square curve fitting again, and the outliers are identified and removed again using the same threshold as in the first step. This iteration procedure is repeated until either all the remaining observations are within the pre-defined threshold or the number of remaining data becomes less than the number of parameters by which the curve is described.

HANTS has been mostly applied to relatively long sampling intervals time series data (e.g. 10 days) and rarely to daily sampling interval time series (Alfieri et al., 2013) in the literature. In this case, in particular, when the time series analysis objective is only gap-filling, the selection of the dominant periodic components becomes less critical. This is due to the fact that a few frequencies can describe the data set and since the sampling interval is one day or longer, the components that are shorter than a day cannot be captured in the analysis. However, this is not the case in this study since we used hourly LST time series data with a complex gap and outlier distribution for a long period of time from 2008 to 2010 over the Tibetan Plateau. A further challenge is the computational load and memory usage required for processing such a huge data set (i.e. hourly LST time series; see Chapter 2).

The objective of this chapter is then, to evaluate the utility of HANTS algorithm to identify and remove the outliers and then fill both the gaps that already exist in the data set and the ones generated by removing the outliers. This leads to reconstruct gap-
free hourly time series of LST estimated from FY-2C geostationary satellite (see Chapter 2). First, the theoretical basis of HANTS algorithm is given. Second, the Fast Fourier series (FFT) algorithm and the power spectrum analysis are explained. Third, the most significant periodic components in the time series of hourly LST data are identified using the power spectrum analysis for short and long periods (three days and one month respectively). Different parameterizations are then used for gap-filling of short gaps in LST data using HANTS algorithm. Finally, the obtained results are validated against ground measurements and synthetic LST data and challenges and shortcomings of the algorithm are discussed.

3.2. **HANTS algorithm**

HANTS algorithm was proposed by Verhoef (1996) to fill the missing or cloudy observations and eliminate the outliers in time series data with periodic behaviour. HANTS algorithm is based on the concept of discrete Fourier transform (Menenti et al., 1993; Roerink and Menenti, 2000a; Verhoef, 1996; Verhoef et al., 1996) to model time series of satellite data.

A temporal sequence of \( N \) observations \( y_i, i = 1 \) to \( N \) can be described by a Fourier series as:

\[
y_i = a_0 + \sum_{j=1}^{M} a_j \cos \left( \omega_j t_i - \phi_j \right)
\]

where \( \omega_j \) is the frequency of \( j^{th} \) harmonic term in the Fourier series, \( t_i \) is the time at which the \( i^{th} \) sample was taken, \( M \) is the number of frequencies of the Fourier series \( (M \leq N) \), \( a_j \) and \( \phi_j \) are the amplitude and the phase of the \( j^{th} \) harmonic term respectively. Because the zero frequency has no phase, the amplitude related to the zero frequency, \( a_0 \), is equal to the average of all \( N \) observations of \( y \). The harmonic frequencies are a base frequency (i.e. \( \omega_1 = 2\pi / N \)) and all integer (i.e. \( i = 1 \) to \( N \)) multiples of the base frequency:

\[
\omega_i = (2\pi / N) \times i, \quad \text{where} \ i = 1, 2, \ldots, N
\]

In the HANTS algorithm after selecting the number of frequencies \( (M) \) and the frequencies \( (\omega_j) \), the unknown parameters of the Fourier series are the amplitudes \( a_j \) and the phase \( \phi_j \) values, which are determined by fitting the time series of observations.

In order to create a reliable model of the signal with HANTS, there are parameters that should be defined by users (Figure 3.3):
1. Valid data range: the acceptable range of observed values. The observations out of this range are removed at the first stage by assigning zero weight to them.

2. Period: number of time samples in each periodic component in the Fourier series.

3. Number Of Frequencies (NOF): the number of harmonic terms. NOF determines the amount of details that can be captured in the reconstructed signal. Low NOF creates a smoother signal than a high NOF value. The number of frequencies is counted from the base period onwards (numbered one).

4. Direction of outliers: indicates the direction of outliers with reference to the current model of the signal. For example, cloud contamination causes lower LST values, therefore, the direction of outliers should be selected as low when applying HANTS algorithm to LST data.

5. Fit Error Tolerance (FET): specifies which absolute deviation from the current curve in the selected direction is still acceptable. After each iteration, observations that have a deviation greater than FET are set as outliers and removed from the calculation by assigning a zero weight to them. Iterations stop when the deviation of all remaining observations becomes smaller than FET.

6. Degree of Over Determinedness (DOD): minimum number of extra data points which have to be used in curve fitting. The number of valid observations must always be greater than the number of parameters required to describe the signal (2×NOF-1). In order to get a reliable result more data points than the necessary minimum should be included which is indicated by DOD. The iteration is terminated if the number of remaining points becomes less than DOD + 2×NOF-1, if it was not already terminated because the FET criterion was met.

As there is no direct way to determine these parameters, some preliminary tests can help to get some idea about the parameters like NOF (determined, for instance, from an earlier Fast Fourier Transform (FFT) analysis) which will be described in the next section.

Before starting to explain the actual results of this study and to have an image about the gaps and outliers in actual LST data, we present an example of LST time series in which there are some gaps and also outliers and the result of using HANTS to reconstruct gap-free data (Figure 3.2). In the next section, we will explain in detail how we can determine the most significant periodic components in the LST time series data.
3.2.1. The evolution of HANTS over time

HANTS originally was designed to be used for time series with short gaps. There are two versions of HANTS; we named them the old and new version, which have a different way of defining the frequencies (the new version can be downloaded from http://gdsc.nlr.nl/gdsc/en/tools/hants). The visual interface of the old and new version of HANTS is shown in Figure 3.3.

In the old version shown on the left side of Figure 3.3, the length of the period is the longest period considered in the Fourier series. This is the same as the base period in the new version (Figure 3.3, Right). As an example, we want to apply HANTS on one-month of hourly LST data in January 2008 which contains 24 hours × 31 days = 744 hourly images and thus our period length is 744 hours. If 744 hours is set as the length of the longest period, then HANTS applies the NOF parameter to construct the Fourier series. It is also possible to select different, arbitrary values as length of period like 72 hours. The latter, however, implies that harmonics at lower frequency (longer period) are not included in the series, while they might be significant. Moreover, the reconstructed curve replicates itself after every 72 hours (Figure 3.4).

The main difference between new and old version of HANTS software is the possibility to select the most significant periods independently from each other in the old version.
3.2.2. Definition of significant frequencies to model time series with HANTS

The definition of significant periodic components is a key-step to construct a representation of a time series. This is considered differently in the old and in the new version of HANTS, and this paragraph describes how to proceed in either case.

The desired frequencies (e.g. the most significant periods) can be determined by the user in the old version of HANTS. The new version, however, determines the periods (frequencies) according to the number of frequencies (NOF parameter) and the selected base period. In this way, the base period \(N\) is divided by a sequence of numbers starting from 1 to the number of frequencies \(M\) that may not all be significant. To illustrate this, for a base period of 744 hours and NOF = 12, both versions delivered identical results (Figure 3.5) as the given frequencies in the old version were designed according to second row of Table 3.1.
Table 3.1: Defined parameters to apply the new and old version of HANTS.

<table>
<thead>
<tr>
<th>HANTS version</th>
<th>Valid data range (K)</th>
<th>Base period (hour)</th>
<th>NOF</th>
<th>Period (hour)</th>
<th>Direction of Outliers</th>
<th>FET(K)</th>
<th>DOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>New HANTS</td>
<td>220-335</td>
<td>744</td>
<td>12</td>
<td>744,372,248,186,149,</td>
<td>LOW</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Old HANTS_1</td>
<td>220-335</td>
<td>744</td>
<td>12</td>
<td>124,106,93,83,74,68,62</td>
<td>LOW</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Old HANTS_2</td>
<td>220-335</td>
<td>744</td>
<td>16</td>
<td>28, 24, 18, 14, 12, 10, 8, 6, 2</td>
<td>LOW</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 3.5: A comparison of the performance of the new and old HANTS algorithm.

It should be noted that the selected parameters (Old HANTS_1 and New HANTS in Table 3.1) lead to reconstructed signals different from the original time series, because not all significant frequencies have been selected. But, when the significant periods were selected in the old version (Old HANTS_2 in Table 3.1), the result was similar to original data (Figure 3.5). So from the present test, the following conclusions can be drawn:

1- The old version of HANTS can consider dominant frequencies which may be identified independently (e.g. if the base period is $N$ and the number of frequencies are $M$, then $M$ frequencies can be selected among the harmonic frequencies).

2- In the new version of HANTS, If the base period ($N$) is equal to the total number of samples in a dataset containing long temporal coverage (e.g. one-month hourly LST time series), we would need a very large NOF with a serious computational load to reconstruct the time series. For instance, in the given test above, using the new version of HANTS, the NOF should be at least 31 to include 24 hours period in the modelling (i.e. period = 744/31 = 24 hours in above test). The possibility of selecting significant frequencies independently from each other is, therefore, important.
3- The new version of HANTS can be used for a long time series (e.g. one-month hourly LST time series) with less computational load, when the base period is selected shorter than the total number of samples in the dataset. For instance, in this study, after doing several tests (results are not shown), 72 hours (three days) was selected as the base periods to be applied with the new HANTS for one-month hourly LST time series.

3.3. **Fourier series analysis of LST time series data**

A decomposition of a periodic function into its components provides useful insight on the processes determining the observed signal and on their relative weight. A multi-annual hourly LST time series has clear diurnal and annual cycles that are strongly related to the annual and diurnal cycles of radiative forcing. Other smaller but significant cycles may be difficult to observe visually, and the proposed method provides estimates of the amplitude and phase of all periodic components, including smaller ones.

3.3.1. **Fast Fourier Transform (FFT) and power spectrum analysis**

Fourier series analysis can be used to decompose a complex signal into its sine or cosine components. The Eq. 3.1 can be written in matrix form as:

\[
\begin{bmatrix}
    y_1 \\
    \vdots \\
    y_N 
\end{bmatrix} = 
\begin{bmatrix}
    f_1(t_1) & \cdots & f_M(t_1) \\
    \vdots & \ddots & \vdots \\
    f_1(t_N) & \cdots & f_M(t_N) 
\end{bmatrix} \begin{bmatrix}
    a_1 \\
    \vdots \\
    a_M 
\end{bmatrix} \tag{3.2}
\]

If we multiply both sides by the transpose of F we get:

\[
F^T y = F^T F a, \quad \text{or} \quad a = (F^T F)^{-1} F^T y \tag{3.4}
\]

then the coefficients \(a\) are estimates of the amplitude and phase of each component obtained by the least squares method.

The Fast Fourier transform (FFT) is an algorithm for computing the matrix multiplication in Eq. 3.4 using relatively few arithmetic operations. A special so-called mixed radix FFT has been developed by Menenti et al. (1993). In this case, the length of a time series of \(N\) observations can be factored with the radix numbers of 2, 3, 4 and 5, and the samples should be equidistant in time. For instance, a time series with 360
observations can be decomposed as $2 \times 4 \times 3^2 \times 5 = 360$, and this saves a lot of processing time. The radix number 4 is not strictly necessary, but it is included because a single transform step with radix 4 is faster to execute than two with radix 2 (Roerink and Menenti, 2000b). The results of the FFT with length N consist of the amplitudes and phases of all frequencies. Having all the amplitudes of a Fourier series, one can identify the most significant periodic components by applying power spectrum analysis (Bloomfield, 2000; Chatfield, 1995). The power spectrum is a plot of power $P(f)$ at frequency $f$ ($P(f) = A(f)^2$ where $A(f)$ is the amplitude as a function of frequency ($f$)). The larger values in a power spectrum indicate the dominant frequencies in the time series. As an example, Figure 3.6 (left) shows an input time series (blue colour line) composed of two different sinusoidal time series (red and green line) and related power spectrum shows the amplitude of these two components (right).

![Figure 3.6: A time series and corresponding power spectrum (http://classes.yale.edu/).](http://classes.yale.edu/)

To illustrate how the FFT identifies periodic components in a time series, we created a synthetic periodic time series (Figure 3.7 top) composed of predefined frequencies and amplitudes. This time series consists of hourly data with 0.5, 1, 10 and 36 days period (i.e. frequencies of 2, 1, 0.1 and 0.0278 cycle per day and amplitude 2, 2, 5, 5.1 respectively). The time series formula with all terms explicit reads:

Number of days = 360 days
Number of samples per day = 24

$N = \text{Number of days} \times \text{Number of sample per day} = 8640$ samples

$X = i / (\text{Number of sample per day}); i = 0, 1, 2, N-1$

Period = 1 day = 24 hours
\[ Y = 2 \sin \left( \frac{2\pi X}{\text{Period}} \right) + 2\sin \left( \frac{2\pi X}{\text{Period}} \right) + 5.1\sin \left( \frac{2\pi X}{10 \times \text{Period}} \right) + 5\sin \left( \frac{2\pi X}{36 \times \text{Period}} \right) \]  

(3.5)

The corresponding power spectrum is given in Figure 3.7 (bottom). In the power spectrum graph, the highest frequencies (i.e. short oscillations) will appear on the right side and the pre-defined periodic components will appear as clear peaks on the left side of the power spectrum. The power spectrum of our synthetic time series shows correctly the pre-defined periodic components at the low frequencies of 0.0278, 0.1, 1 and 2 cycles per day.

Figure 3.7: Synthetic time series (top) and corresponding power spectrum (bottom).

This test shows that the FFT algorithm as implemented identifies correctly the periodic components in a time series. This is required prior to apply HANTS for gap-filling.
3.3.2. Goodness of fit criterion

An important consideration in modelling and reconstruction procedures is the selection of a measure of goodness of fit between actual data and modelled ones. In this thesis, we applied Mean Absolute Error (MAE), R-squared value ($R^2$), and Root Mean Square Error (RMSE), and Bias as metrics for validation. MAE is defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i|$$  \hspace{1cm} (3.6)

where $x_i$ is actual data and $y_i$ is modelled data; $i = 1, 2, ..., n$.

R-squared value, often called the coefficient of determination, is used as another measure of goodness of fit in our study. R-squared is defined as follows:

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}}$$  \hspace{1cm} (3.7)

where,

$$SS_{err} = \sum_{i=1}^{n} (x_i - y_i)^2$$

$$SS_{tot} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

The $SS_{err}$, $SS_{tot}$ and $\bar{x}$ are the sum of squares of residuals, the total sum of squares, and the mean value respectively. RMSE is also applied, but as RMSE is more sensitive to the extreme values, the MAE was also considered in some cases. The RMSE is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}$$  \hspace{1cm} (3.8)

The Bias is defined as follows:

$$Bias = \frac{\sum_{i=1}^{n} (x_i - y_i)}{n}$$  \hspace{1cm} (3.9)

when the Bias is negative (positive) the model overestimates (underestimates) the results.
3.4. Results and discussions

As noted before, there is no clear way to select the proper parameters to be used in HANTS. By doing some trial and error tests (e.g. FFT analysis), one can determine these parameters with acceptable results at the end. Since the objective of this chapter is just gap-filling and not getting information on climatology, the HANTS algorithm was applied on a monthly basis. The other reason is that running the analysis over the entire three years data is not practical due to limited computer memory resource and computational time. Therefore, based on the conclusions in Section 3.2.1 and 3.2.2, the FFT algorithm was first applied to 3 days and then one-month of hourly LST time series (considered as short and long time series relatively) with 72 (three days) and 744 (one month) samples to find the dominant periodic components in these periods. The significant periods longer than one month (e.g. seasonal and yearly components) will be identified in Chapter 5 where all three-year LST time series data (from 2008 to 2010) will be used to estimate the amplitude of significant periodic components.

The previous tests (Section 3.2.2) suggested that first, the old version of HANTS is in principle better than the new version, because of the ability of the old version to assign the dominant frequencies independently. Second, we need to know whether the two versions of HANTS are equally accurate in terms of reconstructing short gaps. Thus, the new and old HANTS algorithm were used to fill the gaps in some experiments. Different parameterizations were considered, and the results were evaluated using the MAE and RMSE metrics to find the best setting of HANTS parameters in either case. The validation of HANTS when applied to hourly LST observations was done by applying it to the time series of ground measurements and synthetic data of LST.

3.4.1. Analysis of a three days LST time series with FFT and HANTS

In order to select the proper parameters and to determine the accuracy of HANTS reconstructed time series after gap-filling, the FFT and HANTS were first applied to a three days LST time series. Selecting three days was due to fact that based on conclusion drawn in Section 3.2.2, we selected 72 hours (three days) as base period to be used in new HANTS, therefore, the dominant periodic components in this period (three days) have to be known. Moreover, it is necessary to select the same parameterization in the new and old version when the results of both are compared. The
accuracy of reconstruction was evaluated by comparison between all the valid actual and reconstructed LST time series on the basis of the RMSE values. As FFT needs data with equal sampling interval and no gaps, a LST time series with no gaps in the study area was selected for FFT analysis. This time series consists of three days LST data (72 hourly LST observations). Figure 3.8 shows the most important frequencies for three days of LST based on power values estimated with the FFT (red bar).

Based on applying the FFT analysis on 3 days of LST data, three frequencies (with periods of 24, 12 and 6 hours) are selected as the most significant periodic components to give NOF = 3 in HANTS because they have relatively higher power value than the remaining and nearby components. To do so, the power spectrum can be smoothed by simply grouping the power values ordinates in sets of size p (e.g. p = 12 or 6 in this test) and finding their mean values (Chatfield, 1995).

![Figure 3.8: Power spectrum of three days of hourly LST data.](image)

We applied the new HANTS software (based on conclusion 3 in Section 3.2.2) to the same three days LST time series with the parameterizations presented in Table 3.2 and Figure 3.9. The results show that the NOF = 3 and 5 give higher RMSE than NOF = 10 and 12 (Table 3.2). It is because when we use more periodic components, the reconstructed time series capture more details in the actual time series. Based on the power spectrum in Figure 3.8, when the NOF is selected as 3, the periods involved in curve fitting are 72, 36 and 24 hours, then the reconstructed curve is composed of lower frequencies which indeed creates a smooth curve showing just the daily variation of LST data. But when the NOF is equal to 12, the corresponding periods are 72, 36, 24, 18, 14, 12, 10, 9, 8, 7, 6.5 and 6 hours which make the reconstructed series more
accurate (RMSE = 2.6K against 6.8K for NOF = 12 and NOF = 3 respectively). There was no significant difference between FET = 3 and 5 when NOF = 10 and 12 were selected. Based on the results in Figure 3.9 and the RMSE values in Table 3.2, HANTS parameters can be set as below for three days LST time series.

1- NOF frequencies = 10-12
2- FET (fit error tolerance) = 5K
3- Valid data range = 240-330K (For different month these values could be changed)
4- DOD (Degree of over determinedness) = 5

3.4.2. Identification of most relevant periodic components in one month LST data

As mentioned in the Section 3.2.2, the advantage of the old version of HANTS over the new one is the flexibility in choosing the most significant frequencies for curve fitting. As done earlier, we applied the FFT algorithm to identify the most significant periodic components in one-month LST data time series and then applied the old and new HANTS version for gap-filling and time series reconstruction.

Figure 3.9: The results of the new HANTS algorithm with different parameters during three days.

One of the outputs of the FFT is the amplitude of Fourier components, which can be used to identify the most important ones. Table 3.3 shows a small part of the FFT output for one-month LST time series with very few gaps.
### Table 3.2: Assigned parameters used with the new HANTS algorithm for three days LST time series and obtained RMSE values.

<table>
<thead>
<tr>
<th>Valid data range (K)</th>
<th>Base period (hour)</th>
<th>NOF</th>
<th>FET</th>
<th>DOD</th>
<th>RMSE (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>240-330</td>
<td>72</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>6.8</td>
</tr>
<tr>
<td>240-330</td>
<td>72</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>7.3</td>
</tr>
<tr>
<td>240-330</td>
<td>72</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>2.6</td>
</tr>
<tr>
<td>240-330</td>
<td>72</td>
<td>10</td>
<td>3</td>
<td>5</td>
<td>2.7</td>
</tr>
<tr>
<td>240-330</td>
<td>72</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>2.6</td>
</tr>
</tbody>
</table>

The first column contains the amplitude values, and the second one is the power values (i.e. squared amplitudes), while Figure 3.10 shows the amplitude vs. period (for the sake of clarity, here, we plotted amplitude vs. period instead of power value vs. period). The gray – highlighted lines in Table 3.3 are some of the dominant components to be used for gap – filling with old HANTS.

### 3.4.3. **HANTS results applied on one month LST time series**

The old HANTS algorithm was applied to one – month time series of LST using the parameters shown in Table 3.4. The length of period is 744 hours except for Rec_9 that is 74 hours, and the NOF changes from 7 up to 16. Figure 3.11 shows the results most useful to find the most accurate reconstructed signal on the basis of the MAE and RMSE values. The Rec_1 test in Table 3.4 represents the result of applying the old version of HANTS with NOF = 12 and using mostly the low frequency components. As low frequencies only are included in the calculation, the accuracy of Rec_1 was not satisfactory (higher MAE and RMSE among others). Since 24 hours was not one of the selected periods, the result does not match well the daily variations. In the Rec_2 test, we added some higher frequencies, i.e. periods = 24, 10, 6, and 2 hours to capture more details of the signal. In this case, the accuracy of reconstruction was higher (MAE = 8.2K and RMSE = 11K). There is no significant difference between Rec_2, Rec_3 and Rec_8. When the 72 hours was selected as Length of period in Rec_9 test, the result was slightly improved compare to other tests mentioned yet. But, if the time series consist of one – year data, the seasonal and yearly components cannot be captured, when the base period is 72 hours. The better result was achieved, when the NOF = 16 and base period = 744 were selected in Rec_12 with MAE = 4.2K and RMSE = 5.7K.
Table 3.3: Amplitude, power value and related frequencies of FFT test for a pixel.

<table>
<thead>
<tr>
<th>Amplitude</th>
<th>Power value (squared amplitude)</th>
<th>Harmonic component</th>
<th>Period (744 ÷ harmonic component)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.3</td>
<td>177.8</td>
<td>1</td>
<td>744</td>
</tr>
<tr>
<td>8.8</td>
<td>77.4</td>
<td>2</td>
<td>372</td>
</tr>
<tr>
<td>2.8</td>
<td>7.7</td>
<td>3</td>
<td>248</td>
</tr>
<tr>
<td>8.5</td>
<td>70.6</td>
<td>4</td>
<td>186</td>
</tr>
<tr>
<td>8.9</td>
<td>79.8</td>
<td>5</td>
<td>148.8</td>
</tr>
<tr>
<td>7.9</td>
<td>63.5</td>
<td>6</td>
<td>124</td>
</tr>
<tr>
<td>8.4</td>
<td>69.4</td>
<td>7</td>
<td>106.3</td>
</tr>
<tr>
<td>7.9</td>
<td>62.9</td>
<td>8</td>
<td>93</td>
</tr>
<tr>
<td>6.4</td>
<td>40.4</td>
<td>9</td>
<td>82.7</td>
</tr>
<tr>
<td>3.9</td>
<td>14.7</td>
<td>10</td>
<td>74.4</td>
</tr>
<tr>
<td>2.5</td>
<td>6.2</td>
<td>11</td>
<td>67.6</td>
</tr>
<tr>
<td>7.1</td>
<td>49.3</td>
<td>12</td>
<td>62</td>
</tr>
<tr>
<td>12.2</td>
<td>145.3</td>
<td>13</td>
<td>57.3</td>
</tr>
<tr>
<td>5.3</td>
<td>29.1</td>
<td>14</td>
<td>53.2</td>
</tr>
<tr>
<td>6.2</td>
<td>38.6</td>
<td>15</td>
<td>49.6</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>16</td>
<td>46.5</td>
</tr>
<tr>
<td>3.4</td>
<td>11.6</td>
<td>17</td>
<td>43.7</td>
</tr>
<tr>
<td>4.1</td>
<td>17.2</td>
<td>18</td>
<td>41.3</td>
</tr>
<tr>
<td>6.9</td>
<td>48.4</td>
<td>19</td>
<td>39.2</td>
</tr>
</tbody>
</table>

Figure 3.10: Amplitude vs. period for one-month LST time series with a minimum number of gaps.

The new Version of HANTS was also applied to the same time series to compare the results with the old version of HANTS (Figure 3.12).

Based on MAE and RMSE values in Table 3.5 the following conclusions were drawn:

1- Increasing base periods from 24, 72 and 744, the MAE and RMSE increase (Rec_1, _2, and _3). By increasing the number of frequencies from 7 to 10; the MAE decreased (Rec_1 and Rec_6; Rec_2 and Rec_5).

2- The best results were achieved with base period = 72 hours, NOF = 10 and FET = 5K.
3- With base period = 744 hours, we need at least NOF = 31 to capture the diurnal component of LST.

The comparison between the new and old HANTS results applied on one-month LST time series shows that with the same base period and different NOF, the old version is more accurate than the new version. This is due to the fact that, in the old version, the most significant periodic components to be used in the signal reconstruction can be selected by the user.

Table 3.4: Different parameterizations used by the old version of HANTS algorithm for gap-filling of one-month LST time series and obtained MAE and RMSE values.

<table>
<thead>
<tr>
<th>Test</th>
<th>valid data range (k)</th>
<th>Length of period (hour)</th>
<th>NOF</th>
<th>Period (hour)</th>
<th>FET(K)</th>
<th>DOD</th>
<th>MAE (K)</th>
<th>RMSE (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rec_1</td>
<td>220-335</td>
<td>744</td>
<td>12</td>
<td>744, 372, 186, 149, 124, 106, 93, 62, 57, 39, 21, 14</td>
<td>5</td>
<td>5</td>
<td>25.9</td>
<td>32.8</td>
</tr>
<tr>
<td>Rec_2</td>
<td>220-335</td>
<td>744</td>
<td>10</td>
<td>744, 372, 106, 57, 39, 28, 24, 10, 6, 2</td>
<td>5</td>
<td>5</td>
<td>8.2</td>
<td>11</td>
</tr>
<tr>
<td>Rec_3</td>
<td>220-335</td>
<td>744</td>
<td>12</td>
<td>744, 372, 106, 72, 57, 39, 24, 14, 10, 6, 3</td>
<td>5</td>
<td>5</td>
<td>8.9</td>
<td>11.7</td>
</tr>
<tr>
<td>Rec_8</td>
<td>220-335</td>
<td>744</td>
<td>12</td>
<td>744, 186, 149, 124, 83, 62, 57, 24, 14, 11, 6, 2</td>
<td>5</td>
<td>5</td>
<td>8.3</td>
<td>11.1</td>
</tr>
<tr>
<td>Rec_9</td>
<td>220-335</td>
<td>72</td>
<td>7</td>
<td>72, 36, 24, 18, 12, 8, 6</td>
<td>5</td>
<td>5</td>
<td>6.9</td>
<td>8.8</td>
</tr>
<tr>
<td>Rec_12</td>
<td>220-335</td>
<td>744</td>
<td>16</td>
<td>744, 372, 345, 165, 106, 57, 39, 28, 24, 18, 14, 12, 10, 8, 6, 2</td>
<td>5</td>
<td>5</td>
<td>4.2</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Figure 3.11: Different curve-fitting of one-month hourly LST time series using the old version of HANTS (For clarity just ten days are shown).
Figure 3.12: Different curve-fitting of one-month hourly LST time series using new version of HANTS (For clarity just ten days are shown here).

Table 3.5: Different values of parameters tested in the new version of HANTS software applied on one-month LST time series.

<table>
<thead>
<tr>
<th>Test</th>
<th>valid data range (k)</th>
<th>Base period(hour)</th>
<th>NOF</th>
<th>FET(K)</th>
<th>DOD</th>
<th>MAE (K)</th>
<th>RMSE (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rec_1</td>
<td>220-335</td>
<td>24</td>
<td>7</td>
<td>5</td>
<td>50</td>
<td>6.54</td>
<td>8.5</td>
</tr>
<tr>
<td>Rec_2</td>
<td>220-335</td>
<td>72</td>
<td>7</td>
<td>5</td>
<td>50</td>
<td>6.68</td>
<td>8.6</td>
</tr>
<tr>
<td>Rec_3</td>
<td>220-335</td>
<td>744</td>
<td>7</td>
<td>5</td>
<td>50</td>
<td>23.93</td>
<td>27.73</td>
</tr>
<tr>
<td>Rec_5</td>
<td>220-335</td>
<td>72</td>
<td>10</td>
<td>5</td>
<td>50</td>
<td>5.26</td>
<td>6.9</td>
</tr>
<tr>
<td>Rec_6</td>
<td>220-335</td>
<td>24</td>
<td>10</td>
<td>5</td>
<td>50</td>
<td>5.9</td>
<td>7.8</td>
</tr>
</tbody>
</table>

3.4.4. HANTS performance and validation with ground measurements

In order to evaluate the HANTS algorithm gap-filling accuracy, an independent set of observations is required. The independent set of observations is created by artificially introducing gaps into the actual observations. The HANTS performance can be evaluated either (i) by considering only the gap-filled observations or (ii) by considering the valid observations that have not been removed. The former addresses the ability of HANTS to fill the gaps, and the latter would give a measure of the HANTS performance in reconstructing the signal when the actual data is valid. The gap-filling accuracy is dependent on the number, size, location and distribution of gaps.

To evaluate the HANTS gap-filling capability, the LST ground measurements were used. The LST was measured at BJ station (see Chapter 2) every 10 minutes from 2008 to 2010. The LST data in January 2008 was used and hourly values of LST were obtained by linear interpolation over time (an hour) since the FY-2C LST data are instantaneous measurements at the hour. Then, the gap pattern of concurrent satellite LST
observations was replicated in the ground measurements to create the independent observations. The created gappy data set has in total 63.3% of gaps from which the longest gap size, defined as a longest period with continuously missing data, is 7.66% (shown as black arrow in Figure 3.13, i.e. 57 hours / 744 hours × 100 = 7.66%). The new and old HANTS versions were used to fill the gaps with different parameterizations to evaluate the performance (Figure 3.13 and Table 3.6). The red area in Figure 3.13 indicates the gaps imposed on ground measurements (blue area). The goodness of fit between the actual ground LST data and the reconstructed by HANTS is presented in terms of four statistical metrics; RMSE, MAE, bias and $R^2$. Table 3.7 lists the detailed results for the two aforementioned evaluation schemes.

![Figure 3.13: The hourly ground measurements of LST data during January 2008 with and without gaps, and reconstructed data based on parameterizations in Table 3.6.](image)

Table 3.6: The HANTS parameters used for gap-filling of LST ground measurements during January 2008 and obtained goodness of fit values of RMSE, $R^2$, bias and MAE.

<table>
<thead>
<tr>
<th>Test</th>
<th>HANTS version</th>
<th>valid data range (K)</th>
<th>Length of period (hour)</th>
<th>NOF</th>
<th>Periods (hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>test_1</td>
<td>New</td>
<td>230-330</td>
<td>72</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>test_2</td>
<td>New</td>
<td>230-330</td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>test_3</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>12</td>
<td>744, 372, 106, 72, 57, 39, 28, 24, 14, 10, 6, 3</td>
</tr>
<tr>
<td>test_4</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>12</td>
<td>744, 372, 106, 72, 57, 39, 28, 24, 14, 10, 6, 3</td>
</tr>
<tr>
<td>test_5</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>9</td>
<td>744, 148, 124, 57, 21, 14, 11, 8, 3</td>
</tr>
<tr>
<td>test_6</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>12</td>
<td>744, 186, 149, 124, 83, 62, 57, 24, 14, 11, 6, 2</td>
</tr>
<tr>
<td>test_7</td>
<td>Old</td>
<td>220-330</td>
<td>72</td>
<td>7</td>
<td>72, 36, 24, 18, 12, 8, 6</td>
</tr>
<tr>
<td>test_8</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>14</td>
<td>744, 372, 248, 148, 124, 93, 82, 74, 67, 57, 31, 24, 12, 8</td>
</tr>
<tr>
<td>test_9</td>
<td>Old</td>
<td>220-330</td>
<td>744</td>
<td>16</td>
<td>744, 372, 345, 165, 106, 57, 39, 28, 24, 18, 14, 12, 10, 8, 6, 2</td>
</tr>
</tbody>
</table>
Table 3.7: The obtained goodness of fit values of RMSE, $R^2$, bias and MAE showing the HANTS ability for gap-filling (left) and reconstruction of signal (right).

<table>
<thead>
<tr>
<th>Test</th>
<th>RMSE (K)</th>
<th>$R^2$</th>
<th>Bias</th>
<th>MAE (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>test_1</td>
<td>11.4</td>
<td>0.25</td>
<td>-1.5</td>
<td>8.8</td>
</tr>
<tr>
<td>test_2</td>
<td>11.4</td>
<td>0.18</td>
<td>-0.41</td>
<td>8.6</td>
</tr>
<tr>
<td>test_3</td>
<td>11.7</td>
<td>0.43</td>
<td>-0.15</td>
<td>9.2</td>
</tr>
<tr>
<td>test_4</td>
<td>11.51</td>
<td>0.43</td>
<td>-1.6</td>
<td>9.2</td>
</tr>
<tr>
<td>test_5</td>
<td>12.9</td>
<td>0.14</td>
<td>-0.82</td>
<td>11.2</td>
</tr>
<tr>
<td>test_6</td>
<td>13.6</td>
<td>0.26</td>
<td>-0.5</td>
<td>11</td>
</tr>
<tr>
<td>test_7</td>
<td>11.4</td>
<td>0.18</td>
<td>-0.43</td>
<td>8.6</td>
</tr>
<tr>
<td>test_8</td>
<td>10.6</td>
<td>0.49</td>
<td>1.2</td>
<td>8.3</td>
</tr>
<tr>
<td>test_9</td>
<td>11.7</td>
<td>0.19</td>
<td>1.5</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>RMSE (K)</th>
<th>$R^2$</th>
<th>Bias</th>
<th>MAE (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>test_1</td>
<td>8.8</td>
<td>0.4</td>
<td>-5.5</td>
<td>6.7</td>
</tr>
<tr>
<td>test_2</td>
<td>8.3</td>
<td>0.4</td>
<td>-4.5</td>
<td>6.3</td>
</tr>
<tr>
<td>test_3</td>
<td>7.7</td>
<td>0.4</td>
<td>-3.5</td>
<td>6.2</td>
</tr>
<tr>
<td>test_4</td>
<td>9.1</td>
<td>0.3</td>
<td>-5.5</td>
<td>7.1</td>
</tr>
<tr>
<td>test_5</td>
<td>11.1</td>
<td>0.2</td>
<td>-6.5</td>
<td>8.5</td>
</tr>
<tr>
<td>test_6</td>
<td>8.8</td>
<td>0.4</td>
<td>-4.6</td>
<td>6.6</td>
</tr>
<tr>
<td>test_7</td>
<td>8.3</td>
<td>0.4</td>
<td>-4.5</td>
<td>6.3</td>
</tr>
<tr>
<td>test_8</td>
<td>7.5</td>
<td>0.4</td>
<td>-1.5</td>
<td>5.2</td>
</tr>
<tr>
<td>test_9</td>
<td>6.6</td>
<td>0.5</td>
<td>-2.5</td>
<td>5.1</td>
</tr>
</tbody>
</table>

The reconstructed time series using new HANTS (test_2) and old HANTS (test_9) are shown in Figure 3.13 as examples of results with lower RMSE and MAE (Table 3.7, right). The test_1 and test_2 are the results of using new HANTS with NOF = 10 and 12 and gave RMSE = 8.8K and 8.3K and MAE = 6.7K and 6.3K respectively. This is due to the fact that the longest period (lower frequency) defined for using the new HANTS was 72 hours (3 days) while the gap size sometimes is longer than 36 hours, especially at the end of time series, therefore the reconstructed time series deviates from the real observations. The test_3 up to test_9 are the results of using the old HANTS version with different parameter settings. The lowest RMSE = 6.6K and MAE = 5.1K among other results in Table 3.7 (right) was achieved by the old version with NOF = 16. The comparison between the calculated RMSE and MAE of HANTS results for signal reconstruction and gap-filling (shown in Table 3.7) illustrates that those two metrics are higher when HANTS gap-filling is considered. This is because the number of valid observations which participate in parameter estimation of the model is less than half of total observations. The other reason is that the gaps are mostly distributed at the end of the time series which makes signal reconstruction very difficult. Furthermore, as there is no information about the daily amplitude at the end of the time series, the daily amplitude is estimated using the observations in the first part of the time series. This causes large error towards the end part of the reconstructed time series.

The results show that when there are long and continuous gaps especially at the beginning and/or end of time series, and the number of gaps is more than half the period defined in HANTS the accuracy of HANTS reconstructed time series becomes poor (Jia et al. 2011). However, the results of reconstruction where the gaps are short (the first half of the time series in Figure 3.13) were satisfactory with RMSE = 3.9K and MAE = 3.1K.
3.4.5. **HANTS performance over the entire study area**

Based on the results from previous tests, the old version of HANTS was applied on hourly LST time series covering the entire study area on a monthly basis from 2008 to 2010 using the same parameters as Rec_12 (with lower MAE and RMSE among other results) in Table 3.4. As an example, Figure 3.14 (top) shows the original LST map for the 1\textsuperscript{st} of January 2008 at 00:00 AM. The black areas in the original LST map display the spatial distribution of gaps (zero values) covering mostly the East, North and South West of the original image with a total 37.5\% spatial coverage of entire study area. The LST values in the original image vary from 220K to 287K, with LST increasing from the North to the South. As an example, HANTS result after gap-filling for the 1\textsuperscript{st} of January 2008 at 00:00 AM is shown in Figure 3.14 (bottom).

![Figure 3.14: The original LST map of 1\textsuperscript{st} January 2008 at 00:00 AM (top) and the reconstructed LST map for the same time using the old version of HANTS (bottom).](image)

The reconstructed LST map shows that the gaps are filled by HANTS with the same pattern as the original image, i.e. the spatial pattern after reconstruction is rather smooth locally (no “salt and pepper” appearance). The ability of HANTS to reconstruct the LST
time series (based on the perspective (ii) defined in Section 3.4.4) has been evaluated by calculating RMSE and MAE maps (as an example, for 1\textsuperscript{st} January 2008 at 00:00 AM) over the entire study area (Figure 3.15). The RMSE and MAE maps show that in most of the cases the values are lower than 4K, therefore, it implies that HANTS can reconstruct the complicated hourly LST time series. This is confirmed by looking at the histograms of those maps presented in Figure 3.16.

There are two visible strip lines in the north part of Figure 3.14 (bottom) which could be due to the sensor malfunction. They belong to pixels where gaps in the original LST time series profiles amount to more than half of all data points in one-month. Clearly, HANTS cannot correctly reconstruct the signal in those areas. This is an intrinsic drawback of HANTS where the gaps are very long and continuous.

Figure 3.15: The RMSE (top) and MAE (bottom) map calculated based on valid observations and corresponding reconstructed signal by HANTS.
Figure 3.16: The histograms of the RMSE map (left) and MAE map (right) calculated based on valid observations and corresponding reconstructed signal by HANTS.

Figure 3.17: The original and reconstructed LST time series for a pixel in the study area with 78.63% gaps during January 2008.

To illustrate this problem, the temporal profile of a pixel representing the distribution of long gaps is shown in Figure 3.17. The selected pixel belongs to one of the pixels in the strip line in the North part of reconstructed image. This pixel has in total 78.63% gaps during one-month, (i.e. 585 hours (total number of gaps in one-month) / 744 hours (total number of observations in one-month)×100 = 78.63%).

3.4.6. Evaluation of the HANTS gap-filling using synthetic data

Apart from the evaluation using ground truth data (Section 3.4.4), we evaluated the HANTS gap-filling performance using synthetic data as well. The synthetic data creation is based on the approach that will be described in the next chapter (see Section 4.8.8), therefore we avoid repeating the approach here. The RMSE between the gap-filled data and the synthetic data is given versus the percentage of the Total Number
of Losses (TNL%) and the length of longest continuous gaps (i.e. Continuous Gap Size CGS%) in Figure 3.18. A smooth trend is seen in the left graph in Figure 3.18, showing an increase in the RMSE values with the increase of TNL. However, the increase of CGS does not correspond with the RMSE increasing in all cases and there are some irregularities in the right graph. A possible explanation for such irregularities is that the distribution of gaps is not identical over the time series. For instance, with the same TNL, if the gaps are mostly located at the beginning or at the end of the time series, the reconstructed signal is less reliable than a case with gaps identically distributed over the time series. The error bars in the figure denotes the standard deviation in each bin.

![Figure 3.18: The variation of RMSE against TNL% and CGS% to evaluate the HANTS gap-filling capability.](image)

### 3.5. Conclusion

HANTS algorithm was used to fill the gaps in the hourly LST time series observed by the FY-2C satellite. Based on the experiments described in this chapter, HANTS shows the capability to fill short gaps. In this study, we used both old and new version of HANTS in gap-filling of three days and one-month hourly LST time series as short and long time series respectively. The old version of HANTS with its ability to choose significant periodic components independently from each other to reconstruct a given time series is better than the new version of HANTS. One of the difficulties in using HANTS is that there is no clear way to select parameters and it needs some testing to find the most suitable parameter setting for a specific data set. One way is to use FFT to find the dominant periodic components in time series. But, even though the FFT helps to identify the significant periodic components, to obtain accurate results different combinations of parameters have to be tested. The validation of HANTS reconstruction was conducted both by using ground measurements of LST and synthetic data in which...
the gaps were artificially imposed. The evaluation of gap-filling by HANTS applied to these gappy data with different parameterizations showed that by increasing the number of frequency (NOF) the accuracy was improved. However, it should be noted that when the total number of gaps in the gapped data set was higher than half of data points and the longest continuous gaps are located at the beginning or at the end of the time series, the reconstructed signal is not reliable. The results of applying HANTS on synthetic data show that when the total number of gaps reach 50-60% of total number of observations, the RMSE will be around 4K, and the same RMSE will be achieved, when the longest continuous gaps are up to 20-30%. As HANTS just considers temporal correlation for each pixel separately, a method considering spatial and temporal correlation to fill the gaps can help especially for continuous gaps. In the next chapter, we will use a different methodology to use both spatial and temporal correlation at the same time to fill the gaps, especially, when the gap size is greater than half the length of a time series.
4. Filling long gaps with the singular spectrum analysis

4.1. Introduction

In Chapter 1, we discussed the importance of having continuous, consistent and accurate time series of Land Surface Temperature (LST) to support studies of climate change, vegetation dynamics and land-atmosphere interactions. Harmonic ANalysis of Time Series (HANTS), as one of the most widely used methods for gap-filling of time series data, was used to reconstruct gap-free LST data (see Chapter 3). The HANTS performance for gap-filling showed that when the gaps are short and distributed across the time series, the accuracy of reconstructed time series was acceptable with the RMSE $\approx 3-4K$ (Figure 3.18). Even though, the HANTS algorithm has been tested for different applications, still there are some limitations when it deals with long continuous gaps. Jia et al. (2011) used the Temporal Similarity-Statistics (TSS) method to find some initial values for HANTS when long continuous gaps exist in MODIS NDVI data using available historical data for each pixel. However, the historical data may also have gaps, and the TSS method does not solve the problem completely, because it considers only temporal correlation to fill the gaps.

In this chapter, the iterative form of the Singular Spectrum Analysis (SSA) is proposed to be used for both single channel variables (considering segments of a single pixel time series) and Multi-channel (M-SSA) which considers segments of a time series at multiple pixels simultaneously. This is considered as an advanced methodology which uses both temporal (SSA) and spatio-temporal (M-SSA) correlation to fill the gaps, especially for long gaps.

SSA (Singular Spectrum Analysis) is an advanced technique for time series analysis that uses the fundamentals of multivariate statistic and geometry, linear algebra mathematics and signal processing (Golyandina and Zhigljavsky, 2013). SSA is
principally a data-adaptive technique. The primary aim of SSA is to decompose a time series into some simpler, interpretable components such as a trend, various oscillatory components (periodic and quasi-periodic) and noise. The main concept in studying the SSA properties is “separability”, which describes how well different components can be separated from each other (Golyandina. et al., 2001).

SSA was proposed by Broomhead and King (1986a, 1986b) and Broomhead et al. (1987). From that time, it has attracted a lot of attention by many people dealing with methodological aspects and applications of SSA (Allen and Smith, 1996; Danilov and Zhigljavsky, 1977; Ghil and Taricco, 1977; Vautard et al., 1992; Yio et al., 2000). An elementary introduction to SSA can be found in the book by Elsener and Tsonis (1996). SSA has been very successful, and has already become a standard tool in the analysis of climatic, meteorological and geophysical time series; see, for example, Vautard and Ghil (1989), Ghil and Vautard (1991), and Yio et al., (1996).

In a dynamic system (e.g. diurnal variation of LST), individual pixel values in time represent the outcome of the interactions among all processes controlling the state of the system (e.g. diurnal and yearly course of the sun). Therefore, the evolution of the observations in time often includes both regular (cycles) and irregular (noise) components. Using this idea, SSA uses Empirical Orthogonal Functions (EOFs), a Principal Component Analysis (PCA) in the time domain, that extracts information from short and noisy time series without initial knowledge of the dynamic processes affecting the underlying time series (Vautard et al., 1992). The structure of the time series in a defined period of time (i.e. the “window size”) is described as a sum of simple, elementary series which is used to define features of time series such as trend, various oscillations (e.g. same as Fourier series analysis in Chapter 3) and noise. The main objective of PCA is to transform a set of (often) time-dependent variables into a smaller set of uncorrelated variables that account for most of the variance in the data in the first few components (Jolliffe, 1986). As in many time series data, a few leading components (significant components) capture most variance in the observations, while additional components may be considered as noise.

The objective of this chapter is to evaluate the usefulness of SSA and M-SSA to identify and fill gaps and remove outliers from time series of hourly satellite observations of LST from 2008 to 2010 covering the entire Tibetan Plateau to construct
high-quality gap-free hourly LST time series. The main question is whether it is possible:

- to utilize SSA and M-SSA to reconstruct the hourly LST time series;
- to identify and fill the gaps;
- to identify and remove the outliers;
- to validate the results.

Like in Chapter 3, the main focus here is just gap-filling and as the data set is huge the gap-filling was applied on monthly segments to reduce the processing time and required memory. Therefore, to extract information on response to periodic climate forcing components (e.g. yearly component) is not the main issue in this chapter. We will identify both climate forcing components (with low frequency) and other significant components (with high frequency) in Chapter 5 using Fourier series analysis applied to the three-year gap-filled LST data, where we estimate the spectral soil thermal admittance.

In this chapter first of all, we will explain the mathematical concept and the main algorithm of SSA. The mathematical background that includes linear algebra and matrix manipulation can be found in Appendix A. Then, the significant components of the LST time series (with \( N \) observations) within a predefined period of time referred to as window size \( (m; m < N) \) will be identified. To ensure that selected components are statistically significant, many realizations of the original time series will be created by assuming different noise scenarios. The results of each scenario will be described, and the reconstruction based on the selected most significant components will be evaluated. After selecting the number of significant components (modes) and window size, they will be used to reconstruct gap-free LST time series data with (M)-SSA. Validation of results using ground measurements of LST will be shown, and finally the evaluation of (M)-SSA performance for gap-filling and noise reduction in different scenarios will be addressed at the end of this chapter.

### 4.2. The singular spectrum

The singular spectrum analysis is based on the spectral (eigenvalue) decomposition of a matrix \( A \) into its set (spectrum) of eigenvalues (Elsner and Tsonis, 1996). The eigenvalues (\( \lambda \)'s) are the numbers that make the matrix \( A - \lambda I \) singular (\( I \) is...
the identity matrix, see Appendix A, Eq. 4). A matrix is singular when it is impossible to factorize it (decompose into simpler matrices) (Appendix A). This is the case when there is at least one zero value in the main diagonal elements (pivot) of a matrix.

The spectral decomposition (Appendix A), which SSA is based on, is a reduction in the dimensionality. The reduction in the dimensionality implies a simpler explanation of the underlying physics described by the components of the time series. SSA is a linear approach to analyse and predict the behaviour of a time series. It is a nonparametric (data – adaptive) method, and this is one of the advantages of SSA over classical spectral methods (e.g. Fourier analysis). Fourier analysis is a model based method which works well when the time series is clearly a combination of periodic components, and when the time series includes some aperiodic components it requires a lot of frequencies to reconstruct time series. In Fourier analysis, the periodic components are imposed, while in SSA, the elementary components are extracted from the data. This implies that if the time series contains intrinsically periodic components, the SSA decomposition will be periodic as well. The data-adaptive nature of SSA makes it suitable for analysis of nonlinear dynamics because there is no assumption about the underlying physical processes which govern the observed time series, so it can be used for any type of time series with different complexities.

As it is noted before, the main purpose of SSA is the spectral decomposition. Linear algebra is a fundamental mathematical method which is used to describe linear functions. There are two kinds of matrix simplification; i) elimination, ii) spectral decomposition.

The elimination method for the simplification of matrices produces diagonal and triangular matrices (Appendix A, Eq. 1), and is usually used for the solution of linear systems.

4.3. The SSA algorithm

The workflow of the SSA algorithm is illustrated as follows based on Musial et al. (2011):

Step1: A single scalar time series \( F(t); t = 1, 2, \ldots, n \) is embedded into a multidimensional trajectory matrix of lagged vectors \( X = [f_1, \ldots, f_k] \) where \( k = n - m + 1 \) and each lagged vector is defined as \( X_j = (f_j, \ldots, f_{j+m-1}); j = 1, \ldots, k \). This trajectory matrix contains the complete record of patterns present within a time window of size
Increasing the window size increases the spectral coverage of SSA and, more information about the basic pattern of the time series will be captured, while decreasing the window size enhances the statistical confidence of the final results (Elsner and Tsonis, 1996), since the structure of a time series will be captured repeatedly (Ghil et al., 2002). The final form of the trajectory matrix \( X \) is a rectangular matrix of the form:

\[
X = \begin{pmatrix}
  f_1 & f_2 & f_3 & \cdots & f_k \\
  f_2 & f_3 & f_4 & \cdots & f_{k+1} \\
  f_3 & f_4 & f_5 & \cdots & f_{k+2} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  f_m & f_{m+1} & f_{m+2} & \cdots & f_n
\end{pmatrix}
\]  

(4.1)

Step 2: The next step is the decomposition of the trajectory matrix \( X \) of size \( m \times k \) using the Singular Value Decomposition (SVD) method (Appendix A) which yields:

\[
X = DLE^T
\]  

(4.2)

where \( D \) and \( E \) are left and right singular vectors of \( X \) with \( m \times m \) and \( k \times k \) size respectively, and \( L \) is a rectangular diagonal matrix of size \( m \times k \). The elements of \( L \), called singular values, are the square roots of the eigenvalues of the lagged – covariance matrix \( S = XX^T \) of size \( m \times m \). The lagged covariance matrix is a symmetric matrix (i.e. \( S = S^T \)), and the elements of \( S \) are proportional to the linear correlation between all pairs of snapshots (patterns appearing in the \( m \)-window). The columns of matrix \( D \) are the eigenvectors of \( S \) also known as Empirical Orthogonal Functions (EOFs). The rows of \( E^T \) are eigenvectors of matrix \( X^TX \). If the time series is naturally periodic and corresponding eigenvalues have high covariance elements along the diagonal values of the lagged-covariance matrix, the eigenvectors will include the periodic components of the time series. As periodic patterns in the time series will result in some segments being in phase and others out of phase, high covariance elements aligned along the diagonals of the lagged covariance matrix \( S \) will indicate oscillations in the time series (Elsner and Tsonis, 1996). If we plot the singular values in descending order, one can often distinguish between an initial steep slope, representing a signal, and a (more or less) flat floor, representing the noise level (Vautard et al., 1992). Then any subset of \( d \) eigenvectors (EOFs), \( 1 \leq d \leq m \), for which the related eigenvalues are positive provides the best representation of the matrix \( X \) as a sum of matrices \( X_i, \ i = 1, 2, \ldots, d \) (Golyandina et al., 2001).
Step 3: Partitioning $d$ eigentriples into $p$ distinct subsets. Then, summing all the components inside each subset such that;

$$X = \sum_{n=1}^{p} X_{n}, \text{ where, } X_{n} = \sum_{i=1}^{n} X_{i}$$

(4.3)

The matrices $X_{n}$ have the form of a Hankel matrix (Appendix A) in an ideal case and consequently fit the trajectory matrices.

Step 4: Since the ideal case described in step 3 is not usually the case, the $X_{n}$ matrices should be transformed into the form of a Hankel matrix to fit the trajectory matrices. This step is known as diagonal averaging. In this sense, the original matrix can be reconstructed as the sum of these matrices.

$$x_{t} = \sum_{n=1}^{p} x_{n}^{n}, \ t = 0, 1, \ldots, n-1$$

(4.4)

where for each $p$, the series $x_{n}^{n}$ is the result of the diagonal averaging of the matrix $X_{n}$.

The multi-channel SSA (M-SSA) is an extension of SSA that is used when time series of maps exist (e.g. the time series of hourly LST maps) (Broomhead and King, 1986b). In this context, M-SSA utilizes a $L$ number of spatial time profiles ($L \leq m$, where $m$ is the window size chosen in SSA) and uses the spatial information (i.e. a few leading S-PCA components of spatial time series) along with temporal information (T-PCA) to reconstruct the time series more accurately especially, where the gaps are long.

If we have more than one time series of given observations, $X_{lt}$, where $i = 1, 2, \ldots, N$ and $l = 1, 2, \ldots, L$; the generalization of SSA to construct the multi-variable lagged covariance matrix ($T$) will be as follows:

$$T = \begin{bmatrix}
T_{1,1} & T_{1,2} & \cdots & T_{1,L} \\
T_{2,1} & T_{2,2} & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
T_{L,1} & T_{L,2} & \cdots & T_{L,L}
\end{bmatrix}$$

where $T_{ll'}$ is the lagged-covariance matrix between channel $l$ and $l'$. The $T_{ll}$ is the same as $S$ for single channel SSA with just one time series with $l = 1$. By diagonalizing matrix $T$, $L \times m$ eigenvectors and eigenvalues of matrix $T$ are created by M-SSA. Each eigenvector ($E^{k}$), which is called Space-Time Principal Components (ST-PCs), consists of $L$ successive $m$-long segments. Same as SSA, by selecting $p$ ST-PCs, the original time series can be reconstructed.
The (M)-SSA software, manuals and help can be freely downloaded from the following website: http://web.atmos.ucla.edu/tcd//ssa/guide/guide4.html (Ghil et al. 2002).

The SSA workflow for gap filling consists of several steps which are explained below.

For a given window size \((m)\) the original time series is centered by computing the unbiased value of the mean, and the missing data is set to zero. The first leading EOF is found by an iterative procedure which applies the SSA algorithm on the zeroed and centered set. The missing values are updated based on the reconstructed component of the current EOF, obtained by projecting the EOF onto the original time series. The SSA algorithm is applied again on this updated set to re-compute the EOF and the missing values are reconstructed again. The process is repeated until a convergence test has been satisfied. The normalized RMSE (as a criterion) between the new reconstructed data in the gappy part and the known values of gaps from the previous iteration (excluding available data) is calculated until the difference between two successive values of normalized RMSE is less than 2.5% (Kondrashov et al., 2010). Then the iteration starts for the second leading EOF (keeping the first one fixed) until convergence has been achieved for the second EOF. This process is repeated for the selected number of EOFs, each time keeping the previous ones fixed. To find the optimal value for the window size and the number of dominant SSA modes to fill the gaps, cross–validation is applied. A portion of the available data (selected at random) is flagged as missing, and the RMSE error of the reconstruction is computed to find the best value for the window size and number of EOFs. There is no direct way in SSA to eliminate outliers, but as few leading components are used in reconstruction, the reconstructed signal will be smoothed, and the outliers are eliminated indirectly in the smoothed final signal. But, outliers still have an impact on the reconstruction.

4.4. Different types of noise in time series analysis

SSA can be used to investigate the structure of a time series. In this case, observations are checked by SSA to find regularities. If they are found, it is possible to build a hypothesis about the physical mechanisms that cause these regularities in the observed signal.

For judging about the application of a technique like SSA to find out regularities, statistics and the role of assumptions regarding noise are fundamental. When we use SSA to find regularities in a data set, we need to apply a careful statistical evaluation to
prove whether these regularities are significant or not. For the first step, we begin with noise processing.

4.4.1. White noise

White noise is a random signal with a flat power spectral density. White noise is a stochastic component and it is independent of the signal. By this definition, we can write the lagged-covariance matrix of the process as:

\[
S = S_{\text{signal}} + S_{\text{noise}}
\]  

(4.5)

where \(S_{\text{signal}}\) and \(S_{\text{noise}}\) are the lagged covariance matrix of signal and noise component of the process respectively. A random vector \((w)\) is a white noise vector if and only if it has the following properties:

\[
\mu_w = E(w) = 0
\]  

(4.6)

where \(\mu_w\) is the mean value of vector \(w\), and the covariance matrix of \(w\) is:

\[
R_{ww} = E(ww^T) = \delta^2 I
\]  

(4.7)

in which, \(\delta^2\) is the variance of vector \(w\) and, \(I\) is the identity matrix. The autocorrelation matrix or lagged covariance matrix of a white noise reads:

\[
S_{\text{noise}} = \delta^2 I
\]  

(4.8)

If we store the eigenvectors of \(S\) as columns of the matrix \(E\), then since \(S\) is symmetric, and we can write:

\[
E^T S E = E^T (S_{\text{signal}} + \delta^2 I) E = \Lambda
\]  

(4.9)

where \(\Lambda\) is the diagonal matrix containing the eigenvalues of \(S\). As the eigenvectors are orthonormal then we have:

\[
E^T E = I
\]  

(4.10)

then we can write:

\[
E^T S_{\text{signal}} E = \Lambda - \delta^2 I
\]  

(4.11)

So, assuming the noise in the time series is white, \(S\) and \(S_{\text{signal}}\) share the eigenvectors, and the relationship of eigenvalues becomes:

\[
\sum_{k=1}^{m} \lambda_k = \sum_{k=1}^{m} \lambda_{\text{signal}}^k + m \delta^2
\]  

(4.12)
So under the assumption of white noise, separation of signal from noise would be easier than with other types of noise. The limitation of SSA for filtering noise does not depend on the complexity of signal, but the type of noise is crucial. If we cannot assume only white noise is present in the observed time series, then detecting the signal will be more difficult.

4.4.2. Autocorrelated noise (red noise)

Autocorrelated noise is a noise in which the noise in a particular observation is related to the noise in nearby (in time) observations. This property can be illustrated by a first order autoregressive process (AR1) such as:

\[ x_t = ax_{t-1} + \varepsilon_t \]  

(4.13)

where \( x_t \) is some observation at time \( t \), \( x_{t-1} \) is a lag-one shifted copy of same variable at time \( t-1 \), \( a \) is lag-one covariance between \( x_t \) and \( x_{t-1} \), \( \varepsilon_t \) is a random error (white noise) with zero mean and variance \( \delta^2 \). Both \( a \) and \( \delta^2 \) are called process parameters.

4.5. Signal detection

One of the main goals of SSA in time series analysis is to identify the signal in the observed record against a background of noise. If we do not have some physical reason to assume the existence of a particular regularity in our time series, then we need to estimate a confidence level to decide whether such regularity (e.g. a periodic component) has appeared by chance. To do so, like Allen (1992) suggested, we can relate a random occurrence to the stochastic part of a data record (noise) and then we need to calculate the probability of capturing such component in our data, and if it is exceeding a threshold, we can assess whether it is random or based on an underlying physical cause.

In a natural system, observed signals are complicated by natural variability and contain some stochastic components. We may assume a white noise as stochastic component, but the observations of e.g. the diurnal variation of LST, relate to the temporal variability of factors affecting LST like soil moisture, thus giving the LST variation a memory. As a result, it is more likely that the diurnal variation of LST will include fluctuations at different frequencies with different amplitudes. Therefore, noise is more likely to be red than white and this should be taken into account when
separating signal and noise. There are two statistical methods to assess the nature of noise: analytical and Monte Carlo (MC). The analytical approach assumes the random component is normally distributed, which will be used as the statistical test. However, the statistics on random variables coming from SSA are non-Gaussian, therefore, in SSA, it is difficult to describe the random variables analytically (Elsner and Tsonis, 1996). The Monte Carlo (MC) approach is applied by creating a surrogate data record assuming red noise.

4.6. Monte Carlo SSA

The Monte Carlo SSA was originally proposed by Allen (1992) to separate signal and noise (Allen and Smith, 1996; Elsner and Tsonis, 1996; Ghil and Taricco, 1977), assuming red noise (Section 4.4.2). As we mentioned in the previous section, the assumption of red noise will be reasonable in most cases. In order to define a red noise, we need to estimate the parameters (e.g. $a$ and $\delta^2$ in Eq. 4.13) which describe it. Recalling the definition of red noise, the noise value at time $t$ depends on the value at time $t-1$, and the parameters can be estimated from the time series itself. If we select a set of parameters that cause rejection of the red noise null hypothesis, while another set of parameters make the red noise null hypothesis likely, then our decision will be inconclusive. Allen (1992) used maximum likelihood criteria to estimate unbiased red noise parameters. The simplest way to calculate the red noise parameters is by using the lag-one auto-covariance of time series as $a$ and the variance of time series as $\delta^2$ (Eq. 4.13). Having the red noise parameters, one can generate different realizations of the red noise by assuming an initial value ($x_0$), and different white noise sets. The idea is that if we take different realizations of the red noise, they would have the same expected variance and lag-one covariance as those of the time series itself. So in this case, we have for example $P$ ensembles of surrogate time series, each containing the same data points. For each surrogate (surr) record, the lagged covariance matrix ($S_{\text{surr}}$) and related eigenvalues ($\lambda_{g_{\text{surr}}}$), $g = 1, 2, \ldots, m$ are calculated by diagonalization of $S_{\text{surr}}$ as follows:

$$\hat{\lambda}_{\text{surr}} = E_{\text{surr}}^T S_{\text{surr}} E_{\text{surr}}$$

(4.14)

Then, the collection of $P$ eigenvalues for each mode (i.e. $g = 1, 2, \ldots, m$) is used to form a distribution function by which the appropriate significance level percentiles can be calculated. Having plotted the eigenvalues of data ($\lambda_{\text{data}}$) as a function of modes ($g = 1, 2, \ldots, m$) and selected a significance level (e.g. 95%) estimated by using
surrogate data, one can decide which eigenvalue and related eigenvector is significant at
the selected level. If an eigenvalue lies above the percentile of the $\lambda_{g}^{surr}$, this indicates
that it is significant at this level, and the null hypothesis of red noise has been rejected
and that mode is related to a particular signal with its frequency and period. There are
three different EOF tests for the Monte Carlo signal – noise separation which will be
explained in the next sections.

4.6.1. Eigenspectrum-shape Monte Carlo test

The Eigenspectrum-shape Monte Carlo method was first proposed by Elsener
and Tsonin (1994). In this context, the ranked eigenvalues of data are plotted along with
chosen percentile of ranked eigenvalues of the noise realizations. To do so, first the
eigenvalues of data are calculated using the diagonalized lagged covariance matrix of
data, and then for each noise realization the eigenvalues are calculated. Finally, the
graph of ranked eigenvalues of data and eigenvalues of noise can be plotted. The
following formulas show how eigenvalues of data and realizations are calculated:

$$\lambda^{surr} = E^{surrT} S^{surr} E^{surr}$$
$$\lambda^{data} = E^{dataT} S^{data} E^{data}$$

(4.15)

4.6.2. Data-based and null-hypothesis-based Monte Carlo test

In the data-based Monte Carlo test, the lagged covariance matrix of data ($S^{data}$)
and ensemble of realizations of red noise ($S^{surr}$) are projected onto EOFs of the data
lagged covariance matrix (Allen, 1992; Allen and Smith, 1994). That is, we have

$$\lambda^{surr} = E^{dataT} S^{surr} E^{data}$$
$$\lambda^{data} = E^{dataT} S^{data} E^{data}$$

(4.16)

This method has two advantages. First, it uses the same transformation for both data and
surrogates. Second, as the $\lambda_{surr}$'s of surrogates have been estimated from a single set of
eigenvalues of data ($E^{data}$), there is no need to decompose the lagged covariance matrix
of surrogates, and it allows a shorter computational time. The drawback of this method
is that since the eigenvalues of surrogates ($\lambda^{surr}$) are calculated based on EOF's of data
($E^{data}$), the $\lambda^{surr}$ will have nonzero off – diagonal elements which contain no useful
information (Elsner and Tsonis, 1996).
Another method proposed by Allen and Smith (1996b) is based on using the eigenvectors of the null-hypothesis. If the lagged covariance matrix of data \( (S_{\text{data}}) \) and ensemble of realizations \( (S_{\text{surr}}) \) are projected onto the EOFs of the probable covariance matrix of red noise (null-hypothesis), and then added to any selected signal EOFs which is identified beforehand, it is called null-hypothesis-based (Allen and Smith 1994). The assumption behind this method is that the null-hypothesis is true until we prove it false. So if the assumption is that the time series data include red noise (null-hypothesis), we try to represent the data using the eigenvectors that we expect for red – noise affected data. To do so, we define the eigenvectors of the null-hypothesis \( (E_{\text{null}}) \) as the eigenvectors of the matrix \( S_{\text{null}} \) which is the average of matrices \( S_{\text{surr}} \) generated by the null-hypothesis. The null-hypothesis-based method would be either related to pure noise or a hybrid of pure noise plus certain signal EOFs (so called “included EOFs”). In the latter case, the result would be more reliable because the noise parameters estimation would be more reliable (Allen and Smith 1994). The advantage of this method over the eigenspectrum-shape test is that the spectrum of eigenvalues is plotted against the dominant frequencies of corresponding eigenvectors (T-EOFs) and it is better for interpretation, because in this case, we can see which eigenvalue belongs to which frequency. The matrix notation of the pure noise null-hypothesis test is as follows:

\[
\lambda_{\text{surr}} = E_{\text{null}}^T S_{\text{surr}} E_{\text{null}} \\
\lambda_{\text{data}} = E_{\text{null}}^T S_{\text{data}} E_{\text{null}}
\]

(4.17)

The hybrid case of the null-hypothesis is when we have identified the components of a time series related to signal, then we can assess whether the remaining components should be attributed to noise (Allen 1992; Allen and Smith 1996). In this method, first a filtered version of the trajectory matrix is calculated as follows:

\[
X' = X E_{\text{data}} \left( I - K \right) E_{\text{data}}^T = X C
\]

(4.18)

where \( K \) is a \( m \times m \) diagonal matrix with \( K_{kk} = 0 \) if eigenvector \( K \) is identified as a signal and \( K_{kk} = 1 \) elsewhere. \( C \) is called signal projection matrix. Then, an explicitly filtered reconstructed version of the original time series \( (F') \) is created by diagonal averaging of \( X' \) and a new data set which consists of the reconstructed signal \( (F') \) plus pure noise AR(1) is created. We called the new data set as composite or hybrid surrogate \( (c.surr) \). The composite surrogates are averaged to estimate the composite null hypothesis \( (c.null) \) lagged covariance matrix \( (S_{c.surr}) \) and by diagonalizing it the
eigenvectors can be calculated ($E_{\text{null}}$). Then, this new data set and many realizations of AR(1) noise are projected onto EOFs of a new data set which consist of pure noise and signal as follow:

$$\lambda_{\text{null}}^c = E_{\text{null}}^T S_{\text{null}} E_{\text{null}}^c$$

$$\lambda_{\text{data}} = E_{\text{null}}^T S_{\text{data}} E_{\text{null}}$$

(4.19)

4.7. Trend analysis and stationarity

Almost all physical systems exhibit some degree of unpredictability and randomness, though the behaviour of the system for different time series of observations of such a system is not exactly the same. But, still there are some deterministic components in the time series of observations which can be extracted using the SSA technique. In order to do that, we need to discriminate between periodic and quasi ‒ periodic components with background noise. This can be done by looking at power spectra of a time series using Fourier series (see Chapter 3). A trend can be defined as a component of a time series with slowly varying magnitude, i.e. it may represent periodic variations over a period of time much longer than the one spanned by available observations. Since the power spectra illustrate the mean distribution of power as a function of frequency over the entire time series, if the time series changes its structure over time, the spectrum is not easily interpretable. So the interpretation of the results of SSA and spectral analysis is more reliable when the assumption of stationarity holds true. But, as noise always exists in time series of observations on natural system (due to measurement error and natural variability), it is impossible to have a time series having perfectly stationary properties (no trend and no varying periodic components). In fact, the only requirements for interpreting a time series’ spectra are weak stationary properties in data (constant mean and variance) (Priestley, 1981).

If the time series is long enough to be considered a combination of all possible realizations, it is described as ergodic (Priestley, 1981). If there are different realizations of a system, it is possible to check the stationarity, but if there is only one realization it is impossible to distinguish between trend or non-stationarity and very-low frequencies in the data (Vautard et al., 1992). If we want to study the high ‒ frequency component or create a power spectrum of data to reveal periodic components in our time series, it is necessary to remove the trend or very low frequency components from our data. There are many de-trending methods like, pre-whitening (Jenkins and Watts, 1968; Weedon,
2003), polynomial fits (Pestiaux et al., 1988) and spline (Yiou et al., 1991). It has been shown that the SSA can be used as a data – adaptive de-trending procedure (Ghil and Vautard, 1991). In another study, Vautard et al. (1992) used the nonparametric test of Mann-Kendall (Kendall and Stuart, 1968) for global trend identification. The Kendall method works as follows:

Let \( x_i, i = 1, 2, \ldots, n \) be a time series, then \( k_r \) is the number of times that \( x_{i_1} < x_{i_2} \) for all \( i_1, i_2 = 1, 2, \ldots, n \) such that \( i_1 < i_2 \). A coefficient (\( \tau \)) can be defined as:

\[
\tau = \frac{4 k_r}{n(n-1)} - 1
\]

which is distributed normally with zero mean and standard deviation (\( \delta \)), i.e.:

\[
\delta = \sqrt{\frac{2(2n+5)}{9n(n-1)}}
\]

The statistical significance test is carried out based on null hypothesis of no trend (i.e. stationarity) and the hypothesis will be rejected when the value of \( \tau \) is outside the interval of \((-1.96\delta, +1.96\delta)\), with 95% confidence level.

4.8. Results and discussions

As mentioned in the introduction, hourly LST time series (used in this thesis) often have a significant number of missing data and outliers (positive and negative). In Chapter 3 we used the HANTS algorithm to fill the gaps and remove the outliers in LST time series. The results were satisfactory when the gaps were short (less than half of observed data points). But this is not always the case, and the LST time series data also has continuous long gaps. To overcome this problem, in this chapter, we evaluate the capability of (M)-SSA to fill the gaps, especially long continuous gaps. The hourly LST time series were reconstructed on a monthly basis from 2008 to 2010, covering the entire Tibetan Plateau. The main reason for selecting a monthly basis is because of the computational considerations. The number of observations per pixel over three-year data set is \( n = 24 \) (hours) \( \times 365 \) (days) \( \times 3 \) (years) = 26280. In this case, for example, selecting a window size of one-year (i.e. \( m = 8760 \) observations) while assuming 10 significant periodic components, the trajectory matrix (\( X \)) would be a matrix of size \( m \times k \), where \( k = n-m+1 = 17521 \). Such a matrix has to be inverted in order to compute the eigenvalues and eigenvectors and this process is repeated iteratively for each component (for instance 10 iteration for each component). On a regular computer (e.g.
64bit, 16GB RAM and CPU 2.8 Dual Core), the inversion takes more than 3 minutes per component. Consequently, the processing time over the entire study area during the 2008 – 2010 period will take more than hundreds of years, which is highly impractical. Therefore, we did the gap-filling procedure on the monthly basis where the study area was divided into 16 equally-size blocks, with $m = 72$ hours and number of components $d = 7$ for the sake of computational efficiency.

In the following sections, we first investigate the capability of SSA in reconstructing the hourly LST time series as well as its ability to remove positive and negative outliers. We then validate the performance of SSA using both LST ground measurements and synthetic data.

### 4.8.1. Reconstruction of LST time series using SSA

Here we want to see whether it is possible to use SSA technique to reconstruct LST time series. As we have hourly LST time series, each 24 hours we can see a clear oscillation which corresponds to the daily cycle of irradiance. As described in Section 4.3, the window size ($m$) and main SSA dominant modes ($d$) are two crucial parameters for SSA. So first we want to select the optimum window size and number of relevant components in this window size. As the nature of hourly LST time series is periodic, the significant modes identified by SSA are also periodic functions.

### 4.8.2. Window size and most relevant periodic components

In order to find the most appropriate window size and significant periodic components in our data, we used LST data measured at a ground station (i.e. BJ station; see Chapter 2) measured every 10 minute during January 2008 with total 4464 observations and also satellite observations of hourly LST using just one pixel during January 2008 with 744 hourly observations. The ground measurements have no gaps and the selected pixel from satellite observation has no outliers but a few gaps which are to be filled by SSA and the original data in the non-gappy part was kept unchanged. This is because we need an error-free data set to see the effect of different window size and components in the reconstructed time series compared to the original one.
4.8.2.1. The most relevant periodic components and window size in the ground measurements data

We first used all ground measurements of LST to explore the effect of window size and number of periodic components in the representation of the original data. Cross-validation is then used to determine the optimum number of leading SSA principal modes and the window size. This was done by creating artificial gaps in the ground measurements and applying SSA with different window sizes and number of components to this gappy data set (e.g. Figure 4.1). The resulting RMSE, MAE, R-squared (R²) and Bias values between reconstructed and original data are used to find the optimum SSA parameters based on a trade-off between a low R-squared value and calculation time (Table 4.1).

![Figure 4.1: The original and reconstructed ground measurements LST data applying SSA with window size = 432 (every 10 minutes) and number of components = 7; the RMSE and MAE of reconstructed signal against original data are 2.75K and 2.13K respectively.](image)

Figure 4.2 (left) shows that increasing the Number of components (No.com) from 7 to 28, causes the R² to increase from 89.4% to 91.3%. However, the calculation time for reconstruction increases considerably when going from 7 to 28 components. Since the R² does not change that much (just 1.9%), it was decided to use seven components for further analysis. The same test for selecting optimum window size was conducted (Table 4.1, right) and as Figure 4.2 (right) shows the optimum window size is 432 which is equal to three days or 72 hours LST data. These values (i.e. the number of components = 7 and window size = 72 hours) were used as main SSA parameters for the representation of time series of LST satellite observations. After selecting those parameters, we assessed the ability of SSA for gap-filling of a LST temporal profile of the corresponding pixel of the ground station (Figure 4.3). This result shows that, SSA was able to create a gap-free data set representing correctly the original satellite data.
The quantitative evaluation of SSA gap-filling capability with different gaps and noise level will be presented in Section 4.8.8.

Table 4.1: The statistical metrics values of estimated and observed LST ground measurements as a function of the number of components (left), and as a function of the window size (right).

<table>
<thead>
<tr>
<th>No.com</th>
<th>RMSE</th>
<th>MAE</th>
<th>R² %</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>2.75</td>
<td>2.13</td>
<td>89.4</td>
<td>-0.01</td>
</tr>
<tr>
<td>14</td>
<td>2.63</td>
<td>1.74</td>
<td>90.4</td>
<td>-0.19</td>
</tr>
<tr>
<td>28</td>
<td>2.51</td>
<td>1.49</td>
<td>91.3</td>
<td>-0.18</td>
</tr>
<tr>
<td>50</td>
<td>2.49</td>
<td>1.38</td>
<td>91.4</td>
<td>-0.18</td>
</tr>
<tr>
<td>100</td>
<td>2.48</td>
<td>1.30</td>
<td>91.5</td>
<td>-0.18</td>
</tr>
<tr>
<td>200</td>
<td>2.48</td>
<td>1.25</td>
<td>91.5</td>
<td>-0.18</td>
</tr>
<tr>
<td>300</td>
<td>2.48</td>
<td>1.23</td>
<td>91.5</td>
<td>-0.18</td>
</tr>
<tr>
<td>432</td>
<td>2.48</td>
<td>1.20</td>
<td>91.5</td>
<td>-0.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Window size</th>
<th>RMSE</th>
<th>MAE</th>
<th>R² %</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>144</td>
<td>3.0</td>
<td>1.6</td>
<td>88.6</td>
<td>-0.7</td>
</tr>
<tr>
<td>432</td>
<td>2.5</td>
<td>1.5</td>
<td>91.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>864</td>
<td>3.4</td>
<td>2.0</td>
<td>84.9</td>
<td>0.5</td>
</tr>
<tr>
<td>1296</td>
<td>3.4</td>
<td>2.2</td>
<td>85.7</td>
<td>0.7</td>
</tr>
<tr>
<td>1728</td>
<td>4.3</td>
<td>2.8</td>
<td>79.7</td>
<td>1.0</td>
</tr>
<tr>
<td>2160</td>
<td>5.0</td>
<td>3.2</td>
<td>77.4</td>
<td>0.7</td>
</tr>
<tr>
<td>2592</td>
<td>5.0</td>
<td>3.2</td>
<td>77.6</td>
<td>0.5</td>
</tr>
<tr>
<td>3024</td>
<td>5.2</td>
<td>3.4</td>
<td>74.2</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Figure 4.2: R-squared values of estimated and observed LST ground measurements as a function of the number of components (left), and as a function of the window size (right).

Figure 4.3: Gap-free reconstructed LST time series during January 2008 using SSA with Number of components (No.com) = 7 and window size = 72 hours.
4.8.2.2. The most significant periodic components and window size in the satellite data

Selecting the window size in not an easy task in SSA analysis. As we mentioned earlier, a large window size can capture longer period oscillations (low frequency components) in a data set, but a smaller window size, in contrast, increases the statistical significance due to fact that low and high frequency components will not compete for limited available variance (Elsner and Tsonis, 1996). SSA cannot determine periods longer than the given window size \( m \) (Vautard et al., 1992). To capture periods of an arbitrary length, it is better to choose longer \( m \). The length of window size \( m \) should not exceed one third of the number of data in time series \( m \leq \frac{1}{3}n \) and SSA is typically successful at analysing periods in the range \((1/5m, m)\) (Vautard et al., 1992). In our data set, we expected to have a dominant 24-hour period. However, to ensure capturing longer periods we selected a window size = 72 hours. On the one hand the difference between window size of 24, 48 and 72 hours is not meaningful and on the other hand an even longer window size makes the computational time very long (see the explanation in Section 4.8) without significant changes in the final results of reconstruction (Figure 4.2). We compared the SSA singular value spectrum of lagged-covariance matrix for different window sizes (Figure 4.6). It shows that the variations of the window size only compress or stretch the spectrum, but the relative magnitude of each individual singular value is unchanged.

4.8.3. Finding the most significant periodic components

The selection of the number of eigenvalues that capture the significant parts of eigenvalue spectrum can be done using the visual interpretation of eigenvalues spectrum and the Monte Carlo test. We will explain the results of the above-mentioned methods in the next sections.

4.8.4. Visual interpretation of singular value spectrum

As we mentioned in Section 4.3, if we plot the singular values of a time series in descending order, usually, we can distinguish between an initial steep slope, representing signals and a more or less flat tail related to noise. Figure 4.4 shows the singular values spectrum of a pixel as representative of data with 72 hours window size. In this graph, the horizontal axis shows the number of modes and the vertical axis shows the logarithm of the variance of related mode. As the window size was set to 72 hours,
SSA decomposes the time series into 72 components and thus we have 72 singular values. The mode numbers 1 to 7 are distinguishable from the flat tail of the graph and they represent the signal part of the spectrum, and they explain ~ 92% of the variance in the data (Figure 4.5).

The pair of SSA modes that are nearly equal and have the related EOFs which are in quadrature phase shift can efficiently represent an oscillation in time series (Vautard and Ghil, 1989). This is because a pair of eigenmodes of a data-adaptive method like SSA can capture the periodicity in data having oscillations. In Figure 4.4, we can distinguish three pairs of modes that have almost the same amount of magnitude (modes 1,2; 3,4 and 6,7). These three pairs belong to three different oscillations (periodic components) in the time series, but we should confirm that statistically. Using the Mann-Kendall trend test (Section 4.7), the singular value related to mode number 5 relates to the trend. To confirm whether they are significantly related to signal, we do the Monte Carlo test.

![Figure 4.4: Singular values spectrum of data with window size of 72 hours with 7 modes above the flat part of the spectrum.](image)

4.8.5. **Monte Carlo SSA**

The Monte Carlo test uses a first order linear autoregressive (AR1) process to create surrogate time series based on parameters estimated from the data (Section 4.4.2). Then for each realization, the SSA has been used to calculate the singular values and
then estimate the error at 95% confidence level to assess the significance of each singular value. The singular values which lie above the error bar are significant and represent the signal part of the spectrum. The following section will represent the results of different Monte Carlo tests (the theory of Monte Carlo SSA was explained in Section 4.6) separately.

Figure 4.5: Graph of normalized singular values with 72 hour window size.

Figure 4.6: Singular value spectrum of LST time series using various window sizes.
4.8.5.1. Results of Eigenspectrum-shape Monte Carlo test

We described the concept of this test in Section 4.6.1. In Figure 4.7, the graph shows the ranked eigenvalues and the error bars of surrogate noises with specific percentile of 97.5%. It is clear that just the eigenvalues of 1 and 2 are significant at this level, and the rest are not distinguishable from the red noise background. The drawback of this method is that the eigenvalues are not related to the associated frequency, and the result does not say anything about the structure of related eigenvectors because there is no unique eigenvectors (Allen and Smith, 1996). So this test just compares the overall shape of the spectrum of data and noise, without considering corresponding frequencies. In this method, the error bars depend on the null-hypothesis of red noise, and they are independent of data except in the calculation of the null-hypothesis parameters. This test gives an overview of the probable significant components, but to investigate the existence of other significant eigenvalues and related eigenvectors we will do another test in the next section.

![Monte Carlo SSA graph](image)

Figure 4.7: Monte Carlo SSA using the eigenvalue shape test.

4.8.5.2. Data-basis and null-hypothesis-basis Monte Carlo test results

The results of data and null-hypothesis EOF-s for pure red noise hypothesis are shown in Figure 4.8. The graphs show the power in the data eigenvalues and surrogates data bars vs. the frequency associated to EOFs. Obtaining a single frequency for each
EOF is not straightforward due to the fact that the EOFs resulting from SSA for an intrinsically periodic time series are not purely sinusoid unless they are composed of completely purely periodic components (e.g. synthetic periodic time series). To do so, Allen and Smith (1996) associated a frequency to each EOFs by maximizing the R-squared ($R^2$) with a pure sinusoid for better visualization.

![Figure 4.8: Monte Carlo SSA based on data and null-hypothesis EOFs test.](image)

The top graph of Figure 4.8 shows the result of the data-based Monte Carlo test that was described in Section 4.6.2. It is clearly observable that eigenvalues related to EOFs 1-2, 3-4 and 6-7 are almost superimposed to each other at frequencies equal to ~ 0.0417, 0.083 and 0.124 cycle/hour (which belong to periods of 24, 12 and 8 hours) (Figure 4.9). The test indicates that these EOFs have a larger amplitude in the data series than related ones in the noise series. The error bars in Figure 4.8 (top) show the 2.5th and 97.5th percentiles of the eigenvalues of surrogates ($\lambda_{\text{surr}}$). If the eigenvalues related to EOFs of data ($\lambda_{\text{data}}$), shown as red points in Figure 4.8, lie above the null-hypothesis error bars, they do not belong to the red noise series at 97.5% significance level. There are three visible significant oscillations pairs shown as black arrows in the Figure 4.8 (top). The null-hypothesis EOFs test which is visible as lower graph of Figure 4.8
confirms the significance of these pairs. The null-hypothesis-based test has the advantage of having a lower probability of selecting a noise component as a signal component, which help to identify significant signal better and thus the result is more reliable (Allen and Smith, 1996). This result should be tested with other Monte Carlo tests which contain better articulated null hypotheses which will be shown in the next paragraph.

Figure 4.9: SSA T-EOFs of components 1-2 (top), 3-4 (middle) and 6-7 (down).

The result of hybrid and null-hypothesis EOFs for pure red noise plus signal is shown in Figure 4.10. The top graph shows that again the eigenvalues of 1-2, 3-4 and 6-7 are superimposed to each other and they are above the significant error bar of 97.55%. So we can conclude that with 97.5% significance the corresponding eigenvectors of mentioned eigenvalues are periodic oscillatory signals and we used them to represent the LST time series.
4.8.6. Positive and negative outliers removal

The existence of outliers (negative and positive) besides gaps impose additional challenges to reconstruct LST time series. In many time series data sets which have periodic components like diurnal variation of LST, the broad, slow variations that offer some degree of periodicity (signals) are of greater interest than the fast changes, which often appear as random, unpredictable events (noise), such as uncertainties in the observations. When few components which belong to the signal are selected to be used by SSA, the reconstructed time series will be smoothed out and the noise and outliers will be removed (Figure 4.11).

However, the effect of existing outliers in the calculation causes some deviation from the original data set. In order to remove this effect, we first applied M-SSA to the original time series of LST (18054 pixels in each block) and reconstructed the time series (e.g. Figure 4.11, red line) with predefined M-SSA parameters. Then the absolute differences between the original time series values and the first reconstructed result have been calculated. As explained before, negative outliers mainly relate to the
temperature of clouds. Clouds temperature varies with their altitude, type and thickness. Clouds are generally expected to have a lower temperature than the background surface (nearly 1 degree per 100 meters altitude). Therefore, a value of 10K has been defined as the threshold between the deviation of the original observations and the reconstructed data. This sets the observations with deviations > 10K as zero values. This procedure is similar to define FET (Fit Error Tolerance) in HANTS algorithm. Then we apply again M-SSA on the new time series in which the outliers have been removed. The results show that the reconstructed values become more similar to the original data after removing the outliers (e.g. Figure 4.12, red line).

Figure 4.11: Reconstruction and outliers removal using M-SSA.

Figure 4.12: Reconstruction of LST before and after outliers removal.
Based on the results of previous tests, M-SSA was applied on monthly segments of hourly LST time series covering the entire study area from 2008 to 2010 with window size = 72 hours and number of components = 7. The entire study area was divided into 16 blocks, each block had 177 columns ×102 rows, and M-SSA was run for each block separately. This is due to the fact that loading and processing the entire study area for one-month of hourly LST need a huge amount of memory and processing time.

![Figure 4.13](image_url)

*Figure 4.13: The original LST map for the 1st of January 2008 at 00:00 AM (top), the reconstructed LST map for the same time using M-SSA (bottom).*
To compare the results of HANTS with M-SSA, the same map shown in Figure 3.14 which is the LST map on January 1\textsuperscript{st} 2008 at 00:00 AM will be presented here (Figure 4.13, top). M-SSA was applied on January 2008 to fill the gaps, and, as an example, the result after gap-filling for the 1\textsuperscript{st} of January 2008 at 00:00 AM is shown in Figure 4.13 (bottom). The reconstructed LST map shows that the gaps are filled by M-SSA with the same pattern as the original LST map on January 1\textsuperscript{st} 2008 at 00:00 AM. In Chapter 3, we mentioned that HANTS is not capable of filling the long continuous gaps. But, M-SSA results show that this problem was solved by M-SSA capability to combine spatial and temporal information of the time series to fill the long continuous gaps. As an example, the spectral profile of the same pixel represented in Figure 3.17 is shown in Figure 4.14. The graph shows that even with long gaps the M-SSA filled the gaps correctly.

Figure 4.14: The original and reconstructed LST time series for a pixel in the study area with 78.63\% gaps during January 2008 using HANTS and SSA.

To ensure that the reconstructed monthly signals of LST time series were connected correctly, we did some experiments. We calculated the deviations between the last observation of the first month (e.g. January 2008) and the first observation of the second month (e.g. February 2008) of original and reconstructed observations separately. If the deviations are rather small, we can conclude that the monthly consecutive segments of observations are connected well. The results showed that on average around 74\% deviations of original and reconstructed observations are within 3K which is comparable with the error of reconstruction and it implies that consecutive months are well connected.
4.8.7. Validation of SSA using ground measurements of LST

In order to validate the results of SSA applied on LST time series, we used the cross-validation method. In cross-validation, we use the data set itself and create some randomly artificial gaps in the time series and then compare the results of reconstruction with the original data set. In this case, we need to have a gap-free original data set. As our original satellite data contain error and noise, cross-validation of results by creating some artificial gaps, is not very reliable. In order to validate results of SSA to reconstruct our LST time series, we used time series of ground measurements of LST at the BJ station (see Chapter 2) in the same time period (January 2008) as satellite observations. Now we have actual and gap-free LST measurements, and we can use them to validate the performance of SSA in gap-filling. To do so, we created random and artificial gaps. As the time coverage of both ground and satellite observation is the same and also in order to have randomly distributed gaps, we imposed the same gap pattern detected (using HANTS) in satellite observations on ground measurements. Then SSA was applied on gappy ground measurements and the result was compared with actual data, i.e. the ground measurements removed to create the gaps. In Figure 4.15, the red areas show the gaps while the blue areas belong to the retained ground measurements. The black line shows the results of applying SSA to fill the gaps. The results imply that even with 63% of gaps, \( R^2 = 0.83 \) with MAE = 2.25K.

![Figure 4.15: Validation of SSA gap-filling using ground measurements with the same pattern of gaps as the corresponding pixel in the satellite image data.](image-url)

In order to find out to what extent we can rely on the results of M-SSA, we try to evaluate the performance of M-SSA in gap-filling and noise reduction of hourly time series of LST. M-SSA needs multivariable time series data (more than one time series), and for evaluating its performance in gap-filling and noise reduction, we need to have a pure data set with no noise, gaps and outliers as a reference and then, add gaps and/or noise and finally compare the results of M-SSA with the original data. This data set could be from ground measurements or synthetic data. The SSA was already evaluated using ground measurements (Section 4.8.7), but for the evaluation of M-SSA, we used synthetic image data.

![Algorithm for generating gaps (left) and noise (right).](image)

Figure 4.16: Algorithm for generating gaps (left) and noise (right).

We used the reconstructed actual LST time series which have been already gap-filled as synthetic data (pure time series) which consists of 177 column × 102 rows ×
744 hours so we will have $102 \times 177 = 18054$ realizations of LST time series. Then, different gap scenarios based on Total Number of Losses (TNL), Number of Gap patterns (NG), Continuous Gap Size (CGS) (Figure 4.16) with or without a autoregressive lag-1 (AR1) time series as red noise, were imposed on pure time series to create gappy and noisy time series (Figure 4.17). For instance, Figure 4.18 shows a time series with 48 total observations. It has 9 missing observations distributed across the time series, one in the beginning, six at the middle and two at the end of time series. TNL is calculated as the percentage of the total number of gaps related to the total number of observations which is $(9/48)\times100 = 18.7\%$. NG is defined as the number of places (patterns) in the time series where there is no data (i.e. three in this example, Figure 4.18). CGS is referred to as the percentage of longest continuous gaps related to total number of observations in the time series, which in this example is $(6/48)\times100 = 12.5\%$.

![Figure 4.17: The synthetic gappy data set (for clarity only 150 out of 18054 pixels are shown here); each column shows the variation of LST in each pixel in time (one month = 744 hours) and the white areas show the gaps patterns.](image)

The performance of M-SSA has been evaluated by comparing the reconstructed time series with the original time series. Root Mean Square Error (RMSE) was used as a measure of the goodness of fit.
4.8.8.1. Performance of M-SSA with only gaps

The synthetic data which we call “reference” data, were created by imposing gaps with different patterns and continuity based on the algorithm shown in Figure 4.16 (left). Then, M-SSA was applied to the created gappy data to fill the gaps. The RMSE between original data and the reconstructed ones has been calculated. The Figure 4.19 illustrates the variation of RMSE with TNL and CGS. It can be seen that with increasing the TNL and CGS up to 60-70% the RMSE = 3-5K. It also shows that even if we have more than half of time series with CGS = 60%, the M-SSA still can reconstruct the time series with reasonable accuracy RMSE = 5K. The Figure 4.19 shows that the RMSE increases more when we have higher CGS than TNL. It means that if we have the same TNL in two time series but one with higher CGS, the RMSE is larger than in the first case.

Figure 4.18: Schematic representation of Continuous Gap Size (CGS), Number of Gap pattern (NG) and Total Number of Losses (TNL).

Figure 4.19: The variation of RMSE against different gap patterns (TNL and CGS).
4.8.8.2. Performance of M-SSA with only noise

As already mentioned in the Section 4.4.2, the red noise or auto-correlated noise is generated by calculating the first lag correlation coefficient between a time series \( x_i, i = 1, 2, \ldots, n-1 \) and one lag shift of the same time series \( x_i, i = 2, 3, \ldots, n \) using the Eq. 4.13. The created red noise for each realization then is scaled to create four groups between 1 to -1 (noise_1), 3 to -3 (noise_3), 5 to -5 (noise_5) and finally all possible variations of red noise which we referred to it as noise_100. The created noises were added to the original time series and then four different time series for each realization were created. M-SSA was used to remove the noise in all these realizations. The RMSE between M-SSA results (noise-free time series) and original time series has been calculated and then the mean and error bar (showing the standard deviation) of RMSE for each noise level is shown in Figure 4.20. Figure 4.20 implies that by increasing the noise level from noise_1 (± 1) to noise_100 (all possible red noise change), the RMSE increases. The results show that even having red noise up to ±5K, the RMSE remains in a reasonable range (RMSE = 1.5 – 3.5K) and it shows that M-SSA has the ability to remove the noise up to ±5K.

![Graph showing RMSE for different levels of red noise.](image)

Figure 4.20: The evaluation of M-SSA for noise removal with different levels of red noise added to pure time series.

4.8.8.3. Performance of M-SSA with gaps and noise

In this section, we added the red noise and gaps together to the original data to evaluate the performance of M-SSA in gap filling and noise removal simultaneously. In this case, for each realization, there are four different time series based on the algorithm shown in Figure 4.16 (four levels of noise described in Section 4.8.8.2). Then, M-SSA was applied to fill the gaps and remove the noises. The results of M-SSA and the original data were used to calculate the RMSE. The average and standard deviation of
RMSE were calculated for each level of TNL from 0-10%, 10-20, .., 90-100%. Figure 4.21 represent the evaluation of M-SSA in gap filling and noise removals. The results show that by increasing the noise level and TNL, the RMSE increases. For each level of gaps starting from 0-10% up to 90-100%, there are four levels of noise starting from ± 1 up to noise_100 (all possible variations of red noise). The results show that having TNL up to 60-70% with the noise level up to ± 5, the RMSE increases from 0.8K to 4.2K.

Figure 4.21: The performance of M-SSA for different levels of noise and TNL.

Figure 4.22: The performance of M-SSA for different levels of noise and CGS.
Figure 4.22 shows the variation of RMSE against the four levels of noise and different CGS. The results show that when the CGS is up to 50-60 % with the noise level up to ± 5, the RMSE will change between 2.5K and 6K. This is the situation where the time series is degraded by long cloudy days, but even in these situations, M-SSA is still capable of reconstructing the time series with acceptable accuracy.

4.9. Conclusion

The (M)-SSA algorithm was used for gap-filling, noise and outlier removal of hourly LST time series estimated from satellite observations during 2008-2010 covering the entire Tibetan Plateau. The LST time series has long continuous gaps (up to 70-80% of data points in some cases), and the results of applying (M)-SSA show the capability and usefulness of this method to fill the gaps with acceptable accuracy even when long gaps are present. Different methods have been tested to identify the significant periodic components and window size in the LST time series. The performance of (M)-SSA was evaluated using both LST ground measurements and synthetic data with different gap number, gap size and location of gaps. The evaluation results can help a user to assess the expected accuracy before applying the M-SSA in a time series with a known gaps distribution. The drawback of the M-SSA is the costly computational time which is obvious since the M-SSA executes a huge number of calculations iteratively (e.g. applying the M-SSA to a monthly multi-channel time series with 18054 channels, each containing 744 data, takes 2.5 hours on a regular computer with 64bit, 16GB RAM and CPU 2.8 Dual Core). The gap-filling procedure was applied on a monthly basis on three years of hourly LST time series covering the entire study area. In the next chapter, the gap-free LST and soil heat flux time series will be used to estimate the soil thermal admittance at different significant frequencies.
Chapter 5

5. Estimation of the Spectral Soil Thermal Admittance (SSTA)

5.1. Introduction

Land Surface Temperature (LST) is one of the key variables required to accurately model the surface energy balance (see Chapter 1). A significant body of work shows that the land surface temperature can be estimated by thermal infra-red remote sensing (Dash et al., 2002; Li et al., 2013; Norman and Becker, 1995; Norman et al., 1995; Sobrino et al., 2004; Sobrino et al., 1994). The thermal emission of the earth, measured in the 3-14 µm wavelength region of the electromagnetic spectrum by a space-borne radiometer, is used to retrieve the LST, which can then be used to estimate net radiation and soil heat flux (see Chapter 1). The soil heat flux is one important component of the surface energy balance and determines heat flow into the soil and changes in LST, modulated by soil thermal properties. The evolution of soil heat flux and LST over time from daily to seasonal periods can be described by a combination of periodic functions due to daily and seasonal variations of solar irradiance, and atmospheric forcing (Carslaw and Jaeger, 1959). These significant periodic components can be identified in the time series of LST and soil heat flux by Fourier series analysis (Chapter 3). Long time series of soil heat flux and LST (i.e. longer than one-year hourly LST time series) is needed to identify significant periodic components caused by seasonal and multi – annual forcing. There is a clear analytical relationship between LST and soil heat flux obtained by travelling wave solution of the differential equation of soil heat flow, assuming periodic heat flux at the surface as boundary conditions (Carslaw and Jaeger 1959). The relationship between the amplitudes of the periodic components of LST and soil heat flux is the soil thermal inertia (admittance) (Carslaw and Jaeger, 1959; Menenti, 1984).
It was shown that when the surface soil heat flux densities \( (G_0) \) and land surface temperature (LST) of a homogeneous semi-infinite soil are sinusoidal functions of time with angular frequency \( \omega \) (1 / 86400 Hz for the daily cycle), the soil thermal inertia, \( I \), can be estimated using remote sensing data on LST and \( G_0 \) as follows (Coppola et al., 2007; Menenti, 1984; Wijk and De Vries, 1966):

\[
I = \frac{\Delta G_0}{\Delta T_0} \sqrt{\omega}
\]

(5.1)

where \( \Delta G_0 \) and \( \Delta T_0 \) are the (e.g. daily) amplitudes at the surface of \( G_0 \) and LST, respectively. The soil thermal inertia is related to the soil thermal properties (assuming a semi-infinite homogeneous soil) as:

\[
I = \sqrt{\rho c \lambda}
\]

(5.2)

where \( \lambda \) is the thermal conductivity (Wm\(^{-1}\)K\(^{-1}\)), \( \rho \) is the bulk density (kg m\(^{-3}\)) and \( c \) is the specific heat (Jkg\(^{-1}\)K\(^{-1}\)) of the soil. Thermal inertia is a key thermal property of materials that accounts for both the capacity and the rate of heat storage and release. Since it relates a flux density to a state variable of materials, it can be considered an impedance of materials (Cracknell and Xue, 1996).

Soil thermal inertia is an useful property of solid bodies, it is accessible to radiometric measurements and is related to soil thermal properties in general and soil water content in particular (Carlson et al., 1981; Coppola et al., 2007; Palluconi and Kieffer, 1981; Wang et al., 2010). Materials with a high thermal inertia will have small changes in temperature, for a given change in heat flux, while low thermal inertia leads to large changes in temperature for the same rate of heat transfer (Xue and Cracknell, 1995).

The first study of thermal inertia dates back to the work of Jaeger (1953) who studied the surface temperature of the moon. Menenti (1984) introduced the concept of soil thermal admittance using an analytical solution on heat transfer in porous media previously proposed by Carslaw and Jaeger (1959). This approach is based on finding a nonlinear relationship between surface soil heat flux and surface temperature, assuming the periodic heating at different frequencies of a semi-infinite (layered) porous medium. The concept of soil thermal admittance will be discussed in Section 5.2. In another analytical approach, Price (1977) developed an algorithm using land surface temperature and reflectance to estimate net radiation and from that soil heat flux to map thermal inertia and showed the contrast between irrigated and desert areas in the north.
part of the Gulf of California. The potential of mapping thermal inertia to support geological interpretation was explored by many researchers (Cassinis et al., 1984; El-Shazly et al., 1973). For instance, Cassinis et al. (1984) compared the estimates based on satellite data with ground truth data and found that thermal inertia mapping is useful for discrimination of lithological units in Sardinia, Italy, but constrained by the low temporal resolution of satellite data.

Direct estimation of thermal inertia using Eq. 5.2 needs parameters which cannot be measured on the ground in large areas and other planets. Remote sensing then can be used to estimate it using radiometric observations and the analytical solution of heat energy balance equation (see Section 5.2). Some examples of using remote sensing observations to estimate thermal inertia of Mars can be found in (Jakosky et al., 2000; Putzig and Mellon, 2007). Coppola et al. (2007) used airborne remote sensing data of infrared and visible part of spectrum to estimate the soil water content by calculating soil thermal inertia from estimated daily amplitudes of land surface temperature and soil heat flux. They integrated remote sensing information into a hydrological model (Campbell, 1985) and a stochastic framework to estimate the soil water content from thermal inertia. Following the work of Carslaw and Jaeger (1959), Sellers (1965) has introduced a simple analytical relationship between diurnal/seasonal amplitude of the soil surface temperature and soil heat flux using a sinusoidal function to represent LST(t) and to estimate soil thermal inertia. This method has been used by Wang et al. 2010 to estimate the thermal inertia using simultaneous ground measurements of soil heat flux and soil surface temperature assuming sinusoidal variation of LST. One common problem in mapping thermal inertia is the lack of proper day and night LST data in diurnal cycles.

The primary objective of this chapter is to estimate soil thermal admittance for each significant periodic component of LST(t). This is done using the gap-filled hourly LST time series, daily mean and instantaneous soil heat flux time series, which are, in turn, estimated from the FY-2C geostationary satellite between 2008 and 2010 over the entire Tibetan Plateau (Chapter 2). To do so, first, the significant periodic components of hourly LST time series over the three years will be identified by power spectrum analysis (Chapter 3). Then, the amplitudes of LST and $G_0$ of those significant periodic components are estimated. Finally, maps of the soil thermal admittance at significant frequencies are produced for the entire Tibetan Plateau. Furthermore, the Spectral Soil Thermal Admittance (SSTA), which shows the variation of soil thermal admittance
against significant frequencies, will be presented for some selected points. The qualitative interpretation of the spectral soil thermal admittance maps will be discussed in Chapter 6.

5.2. Theory and method

Heat transfer takes place by three different processes, namely, conduction, convection, and radiation. In a porous solid, conduction is the dominant process, while convection and radiation are negligible. In liquids and gases, convection and radiation are the most important modes of transferring heat (Carslaw and Jaeger, 1959). In this section, we will summarize the general theory of heat conduction in the soil.

As mentioned above, heat is mainly transported in soils by conduction. The analytical solution we applied to relate the amplitude of heat flux at the surface to the amplitude of surface temperature holds under the following assumptions:

1- Heat flow is one dimensional in the vertical direction.

2- No heat source or sink is present in the soil.

Assuming above conditions, the heat flux density at depth z (Figure 5.1) is directly proportional to the vertical temperature gradient existing at that depth:

$$ G = -\lambda \left( \frac{\delta T}{\delta z} \right) $$

(5.3)

where $G$ is the soil heat flux density (Wm$^{-2}$) (positive and downward if the temperature decreases with depth), $\lambda$ is the thermal conductivity (Wm$^{-1}$K$^{-1}$) and $\delta T/\delta Z$ is the temperature gradient (Km$^{-1}$). The thermal conductivity and other soil thermal properties will be discussed in details in Chapter 6.

The amount of heat stored in the same slab per unit of time can be written as $\rho c (\delta T/\delta t) dz$. The $\rho c$ is the volumetric soil heat capacity and it is defined as the amount of heat required to increase by 1 K the temperature of a unit volume of soil and $\delta T/\delta t$ is the temperature change per unit time.

We assumed that no heat source or sink is present, so that the continuity equation reads:

$$ \frac{\delta G}{\delta z} = -\rho c \frac{\delta T}{\delta t} $$

(5.4)

and by substitution of Eq. 5.3 in Eq. 5.4 it becomes:
\[
\frac{\delta}{\delta z} \left( \lambda \frac{\delta T}{\delta z} \right) = \rho c \frac{\delta T}{\delta t} \quad (5.5)
\]

Assuming constant \( \lambda \) and \( \rho c \) in a soil slab, the vertical heat flow in the soil the Eq. 5.5 becomes:

\[
\lambda \frac{\delta}{\delta z} \left( \frac{\delta T}{\delta z} \right) = \rho c \frac{\delta T}{\delta t} \quad (5.6)
\]

\[
\frac{\delta T}{\delta t} = \alpha \frac{\delta^2 T}{\delta z^2} \quad (5.7)
\]

where \( \alpha = \frac{\lambda}{\rho c} \) (m\(^2\)-s\(^{-1}\)) is soil thermal diffusivity (the soil thermal properties will be discussed in Chapter 6).

![Figure 5.1: A soil slab in the depth z.](image)

The Eq. 5.7 can be solved to have the temperature of soil at any depth and any time if the proper initial and boundary condition are given (Carslaw and Jaeger, 1959; Menenti, 1984; van Wijk and De Vries, 1963).

In order to solve the Eq. 5.7 for a homogeneous semi-infinite soil, we need to define the boundary condition, If we assume a periodic surface temperature \( T(0,t) \) at time \( t \) and depth \( z = 0 \):

\[
T(0,t) = \bar{T} + A \sin(\omega t + \varphi) \quad (5.8)
\]

where \( \bar{T} \) is the mean surface temperature, \( A \) is the amplitude of the surface temperature, \( \varphi \) is the phase and \( \omega \) is the angular frequency (2\(\pi\)/\(p\) and \( p \) is the period, i.e., 2\(\pi\)/86400 = 7.27\times10^{-5} \text{ sec}^{-1} \) for diurnal variation).

The traveling wave solution of Eq. 5.7 can be sought with the boundary condition Eq. 5.8 (Carslaw and Jaeger, 1959; Horton and Wierenga, 1983; Menenti, 1984; van Wijk and De Vries, 1963) as:
\[ T(z,t) = \bar{T} + A \exp(-z/d) \sin(\omega t + \varphi - z/d) \]  
(5.9)

where \( d \) (m) is damping depth defined as:

\[ d = \sqrt{\frac{2a}{\omega}} \]  
(5.10)

When the soil heat flux passes through the soil surface, the change in LST and soil heat flux is not in phase and soil heat flux has a phase shift of \( \pi/4 \) relative to LST (i.e. G leads T by 3 hours for diurnal cycle and 1.5 month for the annual cycle) (Carslaw and Jaeger 1959). The soil heat flux \( G(z, t) \) for a sinusoidal variation of temperature can be obtained from Eq. 5.3 and Eq. 5.9 as follows (Sellers, 1965):

\[ G(z, t) = A(\lambda \rho c \omega)^{1/2} \exp(-z/d) \sin\left(\alpha t + \varphi - z/d + \frac{\pi}{4}\right) \]  
(5.11)

The final relationship between soil heat flux and soil temperature is illustrated by (Menenti, 1984) as follows:

\[ G(z, t) = \lambda' T(z, t) \]  
(5.12)

The complex variable \( \lambda' \) can be written as:

\[ \lambda' = (\lambda \rho c \omega)^{1/2} \exp(i\pi/4) \]  
(5.13)

\[ \gamma = (1 + i) \left( \frac{\omega}{2\alpha} \right)^{1/2} \]  
(5.14)

At the soil surface, the parameter \( \lambda' \) is defined as the thermal admittance \( y_0 \) (Wm\(^{-2}\)K\(^{-1}\)) at the frequency \( \omega \). The term \( \exp(i\pi/4) \) accounts for the phase shift between flux and temperature waves of period \( P \), while the modulus \( (\lambda \rho c \omega)^{1/2} \) accounts for the ratio between surface amplitude of soil heat flux and LST (Eq. 5.15) (Menenti, 1984).

The soil thermal admittance \( y_0 \) can be calculated based on work of (Idso et al., 1976; Menenti, 1984; Wang et al., 2010) as follows:

\[ y_0 = \frac{\Delta G_0}{\Delta T_0} \]  
(5.15)

where \( y_0 \) (Wm\(^{-2}\)K\(^{-1}\)) is the thermal admittance at the soil surface, \( \Delta G_0 \) (Wm\(^{-2}\)) is the surface amplitude of soil heat flux and \( \Delta T_0 \) (K) is the surface amplitude of land surface temperature. As defined above, the surface soil thermal admittance is the ratio between the amplitude of soil heat flux and soil surface temperature. The same relationship can be derived by Eq. 5.9 and Eq. 5.11, now at any frequency \( \omega \) and assuming again a semi–infinite homogenous soil, if we use just the amplitudes as follows:
\[ y_0(\omega) = \frac{\Delta G_0}{\Delta T_0} = (\lambda \rho c \omega)^{1/2} \] (5.16)

Since the soil thermal admittance depends on the frequency \( \omega \), we can estimate a frequency dependent (spectral) soil thermal admittance using the amplitudes of soil heat flux and soil surface temperature obtained by time series analysis of \( G_0 \) and LST data. As shown by Eq. 5.10, periodic components of decreasing frequency will penetrate to increasing soil depths and the \( y_0 \) in Eq. 5.16 is an “apparent” soil thermal admittance since it relates to thermal properties of increasingly deep soil layers with decreasing frequency of the periodic components. Accordingly, the spectral soil thermal admittance conveys information about the soil thermal properties of different layers (Menenti, 1984). The relationship between the spectral soil thermal admittance at the surface and the soil thermal properties of different soil layers is developed and applied in Chapter 6.

In this study, we want to estimate the spectral soil thermal admittance in the whole study area using time series data of LST and \( G_0 \). The Eq. 5.9 and Eq. 5.11 assume a homogeneous semi-infinite soil whose surface is heated in a periodic (sinusoidal) manner that corresponds to the daily, annual or other significant periodic heating cycles. The soil surface temperature and soil heat flux variations are not a pure periodic function of time around an average values (van Wijk and De Vries, 1963), but yet the periodic model can be applied successfully to represent a time series of observations using a Fourier series. If we have the amplitudes of soil heat flux and soil surface temperature at significant frequencies, we can obtain the thermal admittance at these frequencies. It is necessary first to identify the most significant periodic components in LST time series, and then amplitudes of LST and soil heat flux can be estimated at those frequencies. How to identify most significant periodic components was already explained in detail in Chapter 3 and 4.

In the next section, we will illustrate the results obtained by applying the approach described above to calculate the Spectral Soil Thermal Admittance (SSTA).

5.3. Results and discussions

As explained above, in order to estimate the spectral soil thermal admittance, we need to determine the amplitudes of the dominant periodic components of LST and \( G_0 \) time series. First, we identified the significant periodic components in hourly LST time series over three-year from 2008 to 2010. Then, the amplitudes of LST and soil
heat flux time series data at these frequencies were estimated. Finally, the (apparent) soil thermal admittance at different significant frequencies was estimated covering the entire study area.

5.3.1. **Finding the most significant periodic components in LST time series**

As explained in Chapter 3, the periodic components in a time series can be identified by power spectrum analysis. Figure 5.2 (top) shows the LST time series for a pixel in the study area from 1\textsuperscript{st} of January 2008 till 31\textsuperscript{th} December 2010. The power spectrum of this pixel is shown in Figure 5.2 (middle). A base-10 log scale is used for the Y axis of the middle graph in Figure 5.2. The visual interpretation of the calculated power spectrum is not an easy task, since some of the components have very high power value (i.e. daily and yearly components) and some of them have very low power value. As we are looking for the most significant periodic components having relatively higher power values than others, we first sorted all power values in descending order. Then, each power value was normalized by normalizing each power value to the total value of powers in the whole power spectrum and then multiplying it by 100 to get a percentage. The cumulative power spectrum graph was then created by summing up the relative power from the highest value to the lower ones to highlight the most significant frequencies. The first few higher values are the most significant periodic components (Figure 5.2, bottom).

The results show that the most significant periods are 365 and 1 days respectively with total power of 82.5%. We selected 11 most significant periods which contain around 88.4% of total power. They are yearly, daily, 1252 hours (1.74 months), 2190 hours (~3 months), 2630 hours (~3.65 months), 2920 hours (~4 months), 3288 hours (~4.5 months), 3757 hours (~5 months), 4384 hours (~6 months), 5260 hours (~7 months) and 6576 hours (~9 months).

5.3.2. **Estimating the amplitude and phase of LST data at significant frequencies**

The most significant periodic components in a time series explain most variability and contain information about the physical processes that determine that variability. We calculated the amplitude and phase of the dominant components of gap-filled LST time series (the gap-filling procedure was explained in Chapter 3 and 4) for each pixel of the study area. To do so, we used the Fourier analysis and least square
method (described in Chapter 3). Given the significant frequencies, the output of the Fourier series analysis is amplitude and phase at those frequencies. As we selected 11 frequencies as the dominant components (Section 5.3.1), the resulting maps consist of 11 images of amplitude and 11 images of phase.

![Figure 5.2: Original three-year hourly LST time series for a pixel in the study area (top); power spectrum of LST time series against period in days (middle); accumulated power in percentage against related period (bottom).]
Figure 5.3: LST amplitude map for daily component (24 hours period).

Figure 5.4: Land cover map of Tibetan Plateau for the 2000 year (http://landcover.usgs.gov/).
As an example, the LST amplitude map of the daily component (i.e. 24 hours) is shown in Figure 5.3. The amplitude values vary from 1.5K to 24K. It clearly shows some patterns in the study area with distinguishable land cover. The high amplitude values with brown colours are located in north and central part of the study area. These areas in the north of study area are mostly covered with sand as shown in the land cover map in Figure 5.4. The high daily amplitude values in these areas might be due to low soil water content in the soil surface which in turn causes the higher LST variation, since energy absorbed by the surface cannot be dissipated by evaporation. The water bodies, highlighted by black circles in the amplitude map, have lowest daily LST amplitude because of the high thermal capacity of water and high evaporation.

The glaciers with pink and white colour in the bottom and west parts of study area also have low amplitude values. This is because of high reflectance of ice and snow which causes lower heat flux at the surface. Furthermore, the amount of energy absorbed by the surface is dissipated by melting or sublimating of ice, i.e. with small temperature changes during a day. In the land cover map (Figure 5.4), the areas covered by evergreen forest and croplands have higher soil water content at the soil surface and the canopy temperature is less than bare soil which, in turn, shows lower amplitude values in the range 1.5 - 10K.

5.3.3. Estimating the amplitude of soil heat flux

In order to estimate the amplitude of soil heat flux at identified dominant frequencies (Section 5.3.1), the data set should have a consistent and long data record (e.g. like the hourly LST time series) showing the evolution of $G_0$ during short (e.g. daily) and long periods of time (e.g. yearly). However, the available soil heat flux data (see Chapter 2) consists of just a daily mean and an instantaneous value on each day estimated at 14 pm (Beijing local time) from 2008 to 2010. The soil heat flux was estimated by Faivre (2014) using the gap-filled LST time series data over three years from 2008 to 2010. The same soil heat flux was used by Faivre (2014) to estimate the sensible (H) and latent heat ($\lambda E$) flux and these estimates were evaluated against eddy covariance measurements of H and $\lambda E$ at four ground stations in Tibetan Plateau mentioned in Chapter 2. The evaluation results showed that the RMSE of estimated H and $\lambda E$ were 33.88 (Wm^{-2}) and 35.66 (Wm^{-2}), respectively and estimations were reliable.
In this section we describe a procedure to estimate the amplitude of soil heat flux at predefined significant frequencies having a partial and discontinuous time series of $G_0$ observations. We will use first the ground measurements of soil heat flux to estimate the amplitude of soil heat flux for the daily component and then estimate the amplitude for the lower frequency components.

5.3.3.1. Daily amplitude of soil heat flux

We need to estimate the daily amplitude of a soil heat flux time series that contains only the daily mean and an instantaneous value in each day. The Fourier series method used for LST time series is not applicable, since it needs observations equally spaced in time. Moreover, we cannot reconstruct a hourly time series with one observation per day, since the shortest observable period is twice the sampling time (Nyquist theorem).

We have assumed that the ratios between the soil heat flux values at each hour (e.g. 10 pm) in each day to an instantaneous value ($G_{inst}$) at a given time on that day is constant during three years. We have used a time series of hourly $G_0(t)$ ground measurements to estimate the 24 hourly ratios, i.e. a normalized $G_0(t)$ over 24 hours. The 24 ratios have been first calculated for each day during the three years. Average of each ratio at each hour during three years will create the required time series with 24 observations. The maximum and minimum values of this time series are the ratios between the maximum and minimum of the original time series to the instantaneous value. Now if we have an instantaneous value at the same time on any given day, we can multiply it by the maximum and minimum of the created 24 hour time series to estimate the maximum and minimum on each day of the original time series. Then the average of the half daily difference between maximum and minimum values gives us an estimate of the daily amplitude of that time series. The workflow in Figure 5.5 illustrates this approach in a simple way.

Under the above assumptions, we want to use the ground measurements of soil heat flux at three ground stations (i.e. BJ, Namco and Qomo, see Chapter 2) to estimate the daily amplitude of soil heat flux having just one instantaneous value on each day calculated from satellite observations. The ground measurements consist of $G_0$ every 30 minute. The average of every two half hour values will yield an hourly data set, which consists of 26304 (1096 days in three years × 24 hours) values in three years (2008-
There are three steps to estimate daily soil heat flux amplitude using the abovementioned method as follows:

1- The hourly values of ground measurements of soil heat flux in each day \((G_{ij} ; i = 1, 2, 3, \ldots, M\) where \(M = 24\) is the number of hours in a day and \(j = 1, 2, 3, \ldots, N\) where \(N = 1096\) is the number of days in three years) are divided by the instantaneous value on that day \((G_{\text{inst},j})\) which is the 9th observation of \(G_0\) in each day (each day contains 24 measurements). In matrix form this leads to:

\[
A = \begin{bmatrix}
G_{11} & G_{12} & \cdots & G_{1M} \\
G_{\text{inst}_1} & G_{\text{inst}_1} & & \\
G_{21} & G_{22} & \cdots & G_{2M} \\
G_{\text{inst}_2} & G_{\text{inst}_2} & & \\
& \vdots & \ddots & \vdots \\
G_{N1} & G_{N2} & \cdots & G_{NM} \\
G_{\text{inst}_N} & G_{\text{inst}_N} & & \\
\end{bmatrix}
\]

(5.17)

where each row of the matrix \(A\) is the evolution of \((G_{ij} / G_{\text{inst},j})\) ratio in each day and the columns of \(A\) represent the \((G_{ij} / G_{\text{inst},j})\) ratio in each hour over three years.

2- Then, the average of each column of the matrix \(A\) was calculated resulting in a vector (i.e. vector \(B\) in Eq. 5.18) with \(M = 24\) averaged values:

\[
B = \frac{1}{N} \left[ \sum_{j=1}^{N} \frac{G_{1j}}{G_{\text{inst},j}}, \sum_{j=1}^{N} \frac{G_{2j}}{G_{\text{inst},j}}, \ldots, \sum_{j=1}^{N} \frac{G_{Mj}}{G_{\text{inst},j}} \right]
\]

(5.18)

Figure 5.6 shows the evolution of vector \(B\) for three ground stations created using the same method as described above. The average of three graphs was calculated and shown as a black line in Figure 5.6. This graph is assumed to represent the daily evolution of \(G_{ij} / G_{\text{inst},j}\) in the entire study area.

3- The minimum (-0.432) and maximum (1.202) values of \(B\) averaged over the ground stations (i.e. the black line in Figure 5.6) were multiplied by the daily instantaneous values of soil heat flux estimated from satellite \((G_{\text{inst(sat)},j})\) data to obtain the maximum and minimum values of \(G_0\) in each day. The average of amplitudes in each day during three years give us the daily amplitude \((\Delta G)\) of soil heat flux for a pixel in the study area as follows:
\begin{align*}
G_{\text{max}_j} &= 1.202 \times G_{\text{inst}(\text{sat})_j} \\
G_{\text{min}_j} &= -0.432 \times G_{\text{inst}(\text{sat})_j} \\
\Delta G &= \left( \sum_{j=1}^{1096} \frac{G_{\text{max}_j} - G_{\text{min}_j}}{2} \right) / 1096
\end{align*}

(5.19)

where \( j = 1,2,3,\ldots,1096 \) (number of days in three years).

This method can be used to estimate the daily amplitude of soil heat flux having just the instantaneous value in each day. However, we need also to estimate amplitudes of other significant periods identified in Section 5.3.1. This element will be explained in the next section.

Figure 5.5: The workflow to estimate the daily amplitude of soil heat flux using the hourly ground measurements of \( G_0 \) and the instantaneous values of \( G_0 \) from satellite observations.
5.3.3.2. Estimation of amplitude of soil heat flux for all significant periods

The method described in Section 5.3.3.1 can be used to estimate the amplitude of 24-hour period, but we need to estimate amplitudes of longer periods (e.g. yearly period). The soil heat flux time series is assumed to be a periodic function of time (same as hourly LST time series), that can be represented as a combination of the same significant periodic components as LST(t) with a Fourier series (see Chapter 3 for details).

As described in theory and method, LST and soil heat flux have a $\pi/4$ phase shift. This implies that $\pi/4$ can be added to the phase values of the LST periodic components to obtain the phase values of the soil heat periodic components at the same frequencies.

Taking into account that the time series of soil heat flux consists of $N = 1096$ (days) $\times 24$ (hours) $= 26304$ hourly equidistant samples of $G_0$ and that $G_{\text{inst}}$ indicates the $i^{\text{th}}$ observation, then the periodic model implies that the time series $G_0$ can be described by means of a Fourier series as follows:

$$G_{\text{inst}} = G_{\text{mean}} + \sum_{j=1}^{11} a_{\text{amp}_j} \sin \left( \frac{2\pi t_i}{p_j} - (\phi_j + \frac{\pi}{4}) \right)$$  \hspace{1cm} (5.20)

where $G_{\text{inst}}$ is the instantaneous soil heat flux (W m$^{-2}$) for ($i = 1, 2, \ldots, 1096$), $G_{\text{mean}}$ is the mean value of $G_0$ over three years, $a_{\text{amp}_j}$ is the amplitude of period $p$ (hours) ($j = 1, 2, 3, \ldots, 11$), $\phi_j$ is the phase of LST at the significant periods, $p_j$ are the significant periods ($p = 24, 1252, 2190, 2630, 2920, 3288, 3757, 4384, 5260, 6576, 8768$ hours), and $t$ is the time stamp at which $G_{\text{inst}}$ was observed ($t = 9, 27, \ldots, 1081$). The Eq. 5.20
can be solved to estimate the amplitudes at significant periods (as only unknown variable) using least square method as follows:

\[ G_i = G_{inst} - G_{mean} = a_{amp_i} \sin \left( \frac{2\pi t_i}{p_j} - \left( \phi_j + \frac{\pi}{4} \right) \right) \]  

(5.21)

The matrix-vector representation of the above equation becomes:

\[
\begin{bmatrix}
G_1 \\
G_2 \\
\vdots \\
G_i \\
\vdots \\
\end{bmatrix} = 
\begin{bmatrix}
f_{t_1}t_1 & f_{t_2}t_1 & \cdots & f_{t_1}t_1 \\
f_{t_1}t_2 & f_{t_2}t_2 & \cdots & f_{t_1}t_2 \\
\vdots & \vdots & \ddots & \vdots \\
f_{t_1}t_i & f_{t_2}t_i & \cdots & f_{t_1}t_i \\
\vdots & \vdots & \ddots & \vdots \\
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
\vdots \\
a_i \\
\vdots \\
\end{bmatrix} = Fa
\]

(5.22)

where \( a_j \) is the amplitude of the significant periodic components of soil heat flux.

The method just described was applied to estimate the amplitudes of \( G_0 \) of significant periodic components for all pixels of study area. As we identified 11 significant periods, the results were 11 maps each showing the amplitude of soil heat flux at a significant period. As an example, Figure 5.7 shows the daily amplitude map of soil heat flux. As the original soil heat flux time series (estimated by Faivre (2014)) had no data values in the west and southwest of the study area, the amplitude maps also contain NaN values in these areas.
5.3.4. Validation of amplitudes estimated using incomplete soil heat flux data

To validate the accuracy of this method, the estimated amplitudes of a hourly LST time series over three years (we call it here as complete-data) at significant periods will be compared with the estimated amplitudes of the time series that just consists of the daily average and an instantaneous value in each day (we call it here as incomplete-data) to replicate the characteristics of the soil heat flux time series. Figure 5.8 shows the complete data as blue points, the daily averages as green points, and the instantaneous values as red points. For the sake of clarity, we just show some parts of the LST time series in Figure 5.8.

We first estimated the amplitudes of complete-data at significant periods and then the amplitudes of the incomplete-data set using the method in section 5.3.3.2. If the complete time series is a pure sinusoidal time series, the estimated amplitudes using both complete- and incomplete-data set will be the same. However, the hourly LST time series is not a pure periodic time series. To compare the results, we calculated the RMSE, Mean Absolute Error (MAE) and coefficient of determinations ($R^2$) between estimated amplitudes of complete- and incomplete-LST data. The results show the RMSE = 1.8K, MAE = 1.18K and $R^2 = 0.99$ (Figure 5.9, right). Since there is no information about 12-hour period in the incomplete-data set, we ignored it and started from 24-hour period. Figure 5.9 (left) shows the estimated amplitudes using incomplete- and complete-LST data for the significant periods.

The comparison between the results indicates that even having incomplete-data set we can still estimate the amplitudes of significant periodic components with a good accuracy. The discrepancies between the results can be explained by the fact that the time series of LST is not a pure periodic time series.

![Figure 5.8: The hourly, instantaneous and daily mean LST values.](image_url)
5.3.5. Soil thermal admittance

The soil thermal admittance is defined as the ratio between the amplitude of soil heat flux to land surface temperature (Eq. 5.15). It is a frequency – dependent variable which can be calculated at the frequencies of the periodic components of a time series if the amplitudes of $G_0$ and LST at these frequencies are available. The graph of soil thermal admittance against the frequencies is called the Spectral Soil Thermal Admittance (SSTA).

As the amplitudes of $G_0$ and LST at significant periods were already estimated, the soil thermal admittance at those periods was calculated for the entire study area (Figure 5.10). The maps show clearly the variation of thermal admittance in the study area. SSTA varies from ~2 to 11 Wm$^{-2}$K$^{-1}$ for the daily period and 0.14 to 6 Wm$^{-2}$K$^{-1}$ for the yearly period. Assuming constant daily soil heat flux amplitude, the areas having higher thermal admittance values are areas which may have higher soil water content or vegetation area with lower LST fluctuations (low LST amplitude), while in contrast areas having lower values of soil thermal admittance belong to areas having lower soil water content which cause higher LST fluctuations (high LST amplitude). This result is confirmed by comparison between daily LST amplitude (Figure 5.3) and daily soil thermal admittance (Figure 5.10, top).

The variation of soil thermal admittance with the period is given for some pixels in the study area in Figure 5.11. The spectral soil thermal admittance is explicitly related to the soil thermal properties at different soil depths (Menenti, 1984). The interpretation of spectral soil thermal admittance will be presented in the next chapter.
Figure 5.10: Soil thermal admittance maps for daily (top) and yearly periods (bottom).
5.4. Conclusion

This study shows the feasibility of using time series of satellite observations of LST and soil heat flux to estimate the soil thermal admittance. The method developed here is simple, and it needs minimum input data compared to other methodologies. The data needed for estimation of soil thermal admittance are amplitude of LST and $G_0$ which can be estimated assuming the periodicity of both time series by means of Fourier series and least square method. When the available time series is incomplete (i.e. time series of $G_0$) and consists of just the daily mean and instantaneous value in each day, it is possible to derive the amplitude of time series components with good accuracy and estimate the soil thermal admittance. The soil thermal admittance values at different frequencies give us the Spectral Soil Thermal Admittance (SSTA) which contains valuable information about the soil thermal properties in different soil layers. The thermal properties of soil layers can be estimated indirectly using spectral soil thermal admittance which will be described in detail in the next chapter.
Chapter 6

6. Retrieval of soil thermal properties by inverse modeling of spectral soil thermal admittance

6.1. Introduction

The three main soil thermal properties (heat capacity \((\rho c)\), thermal conductivity \((\lambda)\) and thermal diffusivity \((\alpha)\)) are of great importance for a variety of applications. Road and pipeline network design especially in cold environments, estimation of soil temperature and energy balance at the surface of the ground (Alkhaier et al., 2012; Peters-Lidard et al., 1998), agriculture and plant growth rate (Nagai and Makino, 2009), meteorology (Lynch et al., 1998), soil chemical processes (Hopmans et al., 2002) and geology are examples of these applications. Such applications require quantitative information on soil heat transfer which, in turn, depends on the soil thermal properties.

Soil is a complex, dynamic and living system and consists of three phases; solid, liquid and gas. The solid phase includes complex solid materials (i.e. minerals and/or organic matter) randomly distributed and spaced by pores. The pores may be filled with liquid (e.g. water), solid (e.g. ice) and gas or combinations of them. In general, the volumetric fraction of minerals and organic matter, as the main solid phase, evolves slowly over time (e.g. a decade) provided that the physical and chemical properties of soil are not perturbed, for instance because of agricultural activities. Contrariwise, the volumetric fraction of the liquid and gas phases can vary considerably mainly due to changes in the soil water content in response to evaporation and precipitation. The soil thermal properties are influenced by both the volumetric fraction and the type of the soil phases. The spatial and temporal variations of these phases for a given soil profile (and in different soil types) cause changes in the soil thermal properties from regional to global scales.
There are different methods to measure soil thermal properties in situ (Birch, 1950; Lister, 1979; Mottaghy et al., 2008; Somerton, 1992; Waite et al., 2002). Ground soil thermal properties measurements are often time-consuming, costly and impractical for large areas. A number of studies have been conducted to estimate soil thermal properties empirically using most influential factors affecting soil thermal properties such as volumetric fraction of soil constituents (e.g. sand, clay and organic matter), soil water content, and porosity (Chung and Horton, 1987; De Vries, 1963; Farouki, 1981; Usowicz et al., 2008).

These empirical methods often need a large number of parameters to establish the relationship required to estimate soil thermal properties (De Vries, 1952; Johansen, 1975; Kersten, 1949; Van Rooyen and Winterkom, 1957).

We propose an approach to retrieve soil thermal properties by inverse modelling of the Spectral Soil Thermal Admittance (SSTA) estimated from time series of satellite observations of LST and soil heat flux ($G_0$) (for details about estimation of SSTA see Chapter 5). Assuming periodic behaviour of time series of LST and $G_0$, there is an analytical relationship between the soil thermal properties, LST and $G_0$ through soil thermal admittance (Menenti, 1984) (Figure 6.1 and Eq. 6.11). The approach only requires time series data of LST and $G_0$, and it is not necessary to have detailed data on soil constituents and water content. The approach is preferable to in situ and empirical methods, since it can be easily applied over large areas.

Figure 6.1: Soil vertical profile of a three layered soil: a top layer of thickness $z_1$ and a middle layer of thickness $z_2$ overlay a homogeneous semi-infinite soil. $T$, temperature; $G$, heat flux; $y$, thermal admittance.
The key concept of this study is that observations of surface soil thermal admittance at different frequencies provide information about soil vertical structure (e.g. soil thermal properties at different depths). This implies that in principal the soil thermal properties can be retrieved by inverting the forward model (e.g. a physical model) describing the relationship between the SSTA and soil thermal properties against the SSTA estimated with satellite data. The SSTA was calculated as the ratio of the amplitude of $G_0$ to the LST amplitude at significant frequencies identified using Fourier series analysis (see Chapter 5).

It is important to identify which parameters (e.g. soil thermal conductivity) can be retrieved with a set of given observations (i.e. the SSTA at multiple frequencies). The number of retrievable parameters depends on the dimensionality of the observations, i.e. on the number of independent observations. If the number of free parameters is larger than the number of independent observations, the inversion becomes ill-posed. This implies that the solution does not exist, or it is not unique, or it is not stable over the solution space. For instance, the forward model (i.e. Eq. 6.11) may generate identical or very similar SSTA using different combinations of soil thermal properties at each soil depth. To alleviate and overcome the ill-posed-ness of the retrieval problem, it is generally possible to regularize the retrieval problem by narrowing the solution space or using prior information about the solution (Mousivand et al., 2015).

A sensitivity analysis was performed to quantify the influence of model inputs (i.e. soil thermal properties) on the model output. The sensitivity analysis can then help to identify which parameters can be retrieved and which ones can be set constant without significant effect on the final results (Mousivand et al., 2014). The sensitivity analysis results were used as prior information in an inversion model. An inversion technique was applied to retrieve the soil thermal properties of different soil layers. The technique was evaluated by applying the inversion method to synthetic SSTA data. The synthetic SSTA data set was created by using synthetic data on soil thermal properties as inputs to the forward model (i.e. Eq. 6.11) assuming a three layered soil. The synthetic soil thermal properties were calculated using two different semi-empirical methods assuming different volumetric fractions of soil constituents and water content.

The main objective of this work is to investigate the estimation of soil thermal properties by inverse modelling of the SSTA determined with time series data of LST and $G_0$ covering the entire Tibetan plateau.
This chapter is organized as follows. First, in Section 6.2, we illustrate the soil thermal properties and the physical method to estimate them, and then the relationship between the apparent soil thermal admittance at the surface and the soil thermal properties of each layer will be addressed. In Section 6.3, the sensitivity analysis and the inversion methods are discussed. Section 6.4 presents the results obtained from sensitivity analysis and parameter estimation. Finally, Section 6.5 will summarize the conclusions of this study.

6.2. Soil thermal properties

6.2.1. Soil thermal conductivity

Recalling Chapter 5, the steady-state heat flow \( G \) through a soil layer at depth \( z \) is described by the Fourier’ law of heat conduction (Eq. 5.3) and can be calculated if the vertical temperature gradient \( \delta T/\delta Z \) and soil thermal conductivity \( \lambda \) are known at that depth. Soil thermal conductivity is a function of the volumetric fraction of soil constituents, moisture content, and temperature (Chang, 1958). Table 6.1 shows thermal conductivity values of some substances. Soils with a higher amount of quartz have higher thermal conductivity than soils containing more organic matter because of the higher thermal conductivity of quartz than organic matter (Table 6.1). Soil thermal conductivity increases with increasing moisture content since water replaces air in the soil pore space and water has higher thermal conductivity than air. Frozen soil has higher thermal conductivity than unfrozen one. Homogeneous soil has relatively constant thermal conductivity. Soil thermal conductivity usually varies between 0.15 (Wm\(^{-1}\)K\(^{-1}\)) for dry sand up to 4 (Wm\(^{-1}\)K\(^{-1}\)) for saturated soil (Sellers, 1965).

<table>
<thead>
<tr>
<th>substance</th>
<th>( \rho c ) (Jm(^{-3})k(^{-1}))</th>
<th>( \lambda ) (Wm(^{-1})K(^{-1}))</th>
<th>( \alpha ) (m(^2)s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartz</td>
<td>1.9246E+6</td>
<td>8.8</td>
<td>4.6E-6</td>
</tr>
<tr>
<td>Soil minerals</td>
<td>1.9246E+6</td>
<td>2.93</td>
<td>1.5E-6</td>
</tr>
<tr>
<td>Organic matter</td>
<td>2.5104E+6</td>
<td>0.25</td>
<td>1.0E-6</td>
</tr>
<tr>
<td>Water</td>
<td>4.184E+6</td>
<td>0.573</td>
<td>1.42E-6</td>
</tr>
<tr>
<td>Ice</td>
<td>1.726E+6</td>
<td>2.177</td>
<td>1.26E-6</td>
</tr>
<tr>
<td>Air</td>
<td>1213.4</td>
<td>0.025</td>
<td>0.21E-6</td>
</tr>
</tbody>
</table>

Table 6.1: The thermal properties of some soil substances (De Vries, 1963).
There are several methods to estimate soil thermal properties based on different parameterization schemes of the properties and interactions of soil constituents. In this study, two methods were used to calculate soil thermal properties: De Vries model (De Vries, 1952, 1963) and Yang model (Yang et al., 2005).

### 6.2.1.1. De Vries model

Soil thermal conductivity can be estimated by the procedure developed by De Vries (1952, 1963). In this model, the soil thermal conductivity is calculated using the volumetric fraction of the soil constituents and thermal conductivity of each component and take the weighted average of the conductivities as follows:

\[
\lambda = \sum_{i=1}^{n} k_i \frac{\lambda_i x_i}{\sum_{i=1}^{n} k_i x_i}
\]

(6.1)

where \(n\) is the number of components, \(\lambda_i\) (Wm\(^{-1}\)K\(^{-1}\)) is the thermal conductivity of each component (Table 6.1), and \(x_i\) is the volume fraction of each component. The weighting factors \(k_i\) depend on the shape and orientation of the granular particles of the soil and the ratio between thermal conductivities of each component and surrounding medium and are calculated from:

\[
k_i = \frac{1}{3} \sum_{i=1}^{n} \left[ 1 + \left( \frac{\lambda_i}{\lambda_0} - 1 \right) g_i \right]^{-1}
\]

(6.2)

where \(g_i\) shows the shape factors for \(i^{th}\) components and \(\lambda_0\) (Wm\(^{-1}\)K\(^{-1}\)) is the thermal conductivity of air when the soil is dry and water when the soil is moist. The term \(\lambda_i / \lambda_0\) accounts for the difference between the conditions where soil grains are in direct contact with air or with the water film.

### 6.2.1.2. Yang model

The soil thermal conductivity can be estimated using the model proposed by Yang et al. (2005) as follows:

\[
\lambda = k_e (\lambda_{sat} - \lambda_{dry}) + \lambda_{dry}
\]

(6.3)

where \(\lambda_{sat}\) and \(\lambda_{dry}\) are the saturated and dry bulk soil thermal conductivity (Wm\(^{-1}\)K\(^{-1}\)), respectively and \(k_e\) is the Kersten number (Kersten, 1949). The dry soil thermal conductivity is given by the following empirical formula (Johansen, 1975):
\[ \lambda_{\text{dry}} = \frac{0.135 \rho_m + 65.7}{2700 - 0.947 \rho_m} \]  

(6.4)

where \( \rho_m = (1 - \theta_{\text{sat}}) \times 2700 \) is the bulk density of the soil mineral phase (kg m\(^{-3}\)) and \( \theta_{\text{sat}} \) is volumetric saturated water content, equal to the soil porosity. The following relationship is used to calculate the saturated soil thermal conductivity:

\[ \lambda_{\text{sat}} = \lambda_a^{1-\theta_w} \lambda_{\text{sat}}^{-\theta} \lambda_w^{\theta} \]  

(6.5)

where \( \lambda_a = 0.025 \text{ (Wm}^{-1}\text{K}^{-1}) \) is the thermal conductivity of air, \( \lambda_w = 0.6 \text{ (Wm}^{-1}\text{K}^{-1}) \) is the thermal conductivity of water, \( \theta \) is the soil water content and \( \lambda_s \text{ (Wm}^{-1}\text{K}^{-1}) \) is the soil solid phase thermal conductivity. The \( \lambda_a \) is different from \( \lambda_{\text{dry}} \), since the former is just the thermal conductivity of the soil solid phase while the latter is the bulk dry soil thermal conductivity in which the air in the pores is also considered. The \( \lambda_s \) depends on the amount of quartz in soil and can be calculated as follows:

\[ \lambda_s = \lambda_q^{1-q} \lambda_m^{q-1} \]  

(6.6)

where \( \lambda_q = 7.7 \text{ (Wm}^{-1}\text{K}^{-1}) \) is the thermal conductivity of quartz and \( \lambda_m = 2 \text{ (Wm}^{-1}\text{K}^{-1}) \) is thermal conductivity of other minerals. The amount of quartz \( (q) \) is not easy to measure and can be estimated as \( q = \% \text{sand} / 2 \). The Kersten number (Kersten, 1949) is the normalized thermal conductivity that shows the dependency of soil thermal conductivity on soil moisture. It is calculated as follows:

\[ k_s = \exp \left( k_s \left( 1 - \frac{1}{S_r} \right) \right) \]  

(6.7)

where the \( k_T = 0.36 \) is an empirical coefficient and \( S_r = \theta / \theta_{\text{sat}} \) is the saturation degree.

### 6.2.2. Soil volumetric heat capacity

Soil heat capacity \( (\rho c) \text{ (Jm}^{-3}\text{K}^{-1}) \) has an important role in controlling the soil temperature regime, since it determines its ability to store or release heat. Soil heat capacity \( (\rho c) \text{ (Jm}^{-3}\text{K}^{-1}) \) can be calculated from the volumetric fraction and heat capacity of each constituent according to De Vries (1963) as follows:

\[ \rho c = x_s \rho c_s + x_o \rho c_o + x_m \rho c_m \]  

(6.8)

\[ \rho c_s = \frac{x_s \rho c_m + x_o \rho c_o}{x_s} \]  

(6.9)
\( x_s, x_w, x_{air}, x_m, \) and \( x_o \) are the volume fraction of solid matter, water, air, minerals, and organic matter in the soil, and \( \rho c_s, \rho c_w, \rho c_{air}, \rho c_m \) and \( \rho c_o \) are corresponding heat capacities, respectively.

The ratio between soil thermal conductivity and soil heat capacity is the soil thermal diffusivity \( (m^2 \text{ s}^{-1}) \):

\[
\alpha = \frac{\lambda}{\rho c}
\]

(6.10)

Soil thermal diffusivity characterizes the rate of temperature change. For instance, under the same conditions, a change in temperature in a frozen soil is faster than in an unfrozen one because of its higher thermal diffusivity.

### 6.2.3. The physical model of soil thermal admittance assuming a three-layer soil

In Chapter 5, we described how to estimate SSTA using time series satellite data of LST and \( G_0 \). The final product was a series of maps each representing the spatial pattern of soil thermal admittance at the significant frequencies covering the entire study area (e.g. Figure 5.10) and the frequency-dependent values of soil thermal admittance (in each pixel) give the SSTA which will be used as observation in this study. The observed SSTA contains information about the soil thermal properties in different soil layers which can be retrieved against the SSTA simulated by a physical forward model.

Soil heat flow and the relationship between \( G_0 \) and LST have been described in Chapter 5. The analytical relationship between the soil thermal properties and the apparent soil thermal admittance at the surface (assuming a three-layers soil, Figure 6.1) was introduced by Menenti (1984) as follows (we will call it forward (physical) model in this study):

\[
y_0 = \left[ \frac{\tanh \left( y_1z_1 \right) \frac{y_1}{y_2} + \frac{y_1}{y_1} \tanh \left( y_1z_2 \right) + \frac{y_1}{y_1} \tanh \left( y_1z_1 \right)}{\tanh \left( y_1z_1 \right) \frac{y_1}{y_1} + \frac{y_1}{y_1} \tanh \left( y_1z_1 \right) + 1} \right] y_1
\]

(6.11)

where \( y_0 \) is the apparent surface thermal admittance of the three layers soil system at a specific frequency; \( y_1 \) and \( y_2 \) are the thermal admittances of the first and second layers respectively, and \( y_3 \) is the thermal admittances of the last semi-infinite soil (Figure 6.1), \( z_1 \) and \( z_2 \) are the soil thicknesses of the first and second layers and \( \gamma_1 \) and \( \gamma_2 \) are calculated by using the soil thermal properties of the first and second layers through Eq.
5.14 (see Chapter 5). For each layer, we need to have soil thermal conductivity and heat capacity to calculate the related soil thermal admittance (e.g. $y_1$ for the first layer). Therefore, assuming three soil layers, we need six soil thermal properties and two soil thicknesses, i.e. eight parameters in total to determine the apparent soil thermal admittance at the surface at each frequency.

6.3. Methodology

6.3.1. Smoothing spectral soil thermal admittance

As noted in Section 6.2.3, SSTA can be simulated at any frequency using the soil thermal properties of different soil layers (e.g. assuming three layers as done in this study). Figure 5.11 (see Chapter 5) illustrates, for instance, the observed spectral soil thermal admittance (using satellite data) of a few pixels in the study area. As shown in that figure, there exist some peaks in the sampled pixels, especially at longer periods. We cannot outline any particular reason for such peaks, however they may be either due to specific processes taking place in the underlying soil layers (e.g. melting or freezing of soil water), or due to the estimation error related to satellite based estimates of $G_0$ or $G_0$ amplitude. We have designed three different experiments to find out the reason of the observed peaks in the calculated SSTA from satellite data. The first experiment was to investigate the percentage of the peaks in the data, assuming that a high percentage guarantees that these peaks are valid observations, otherwise they can be considered as outliers. The second experiment regards the capability of the forward model to simulate SSTA from a set of input soil properties. This is to document under what circumstances a peak can be observed in the simulated data. The last experiment is to calculate SSTA using continuous ground truth measurements of LST and $G_0$ as a reliable source of data.

1st Experiment. We plotted the histogram of the observed Soil Thermal Admittance (STA) at each frequency covering the entire study area (see Appendix B). As an example, the histograms of STA for periods of 3288 and 3757 hours are shown in Figure 6.2: the left (right) histogram implies that the 95% (80%) of total STA values of period 3288 hours (3757 hours) are less than 6 Wm$^{-2}$k$^{-1}$ (5 Wm$^{-2}$k$^{-1}$). In most cases the mode in Table 6.2 is small and the percentage of the peaks (Table 6.3) is less than 10% of the data. A few, very large values of thermal admittance appear at all frequencies, on the other hand.
Figure 6.2: The histogram of STA: period = 3288 hours (left), and 3757 hours (right).

Table 6.2: The basic statistics of STA (Wm\(^{-2}\)k\(^{-1}\)) for each period and the entire study area.

<table>
<thead>
<tr>
<th>Periods (hours)</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>mode</th>
<th>Stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>2.89E-04</td>
<td>16.84</td>
<td>5.34</td>
<td>3.16</td>
<td>2.40</td>
</tr>
<tr>
<td>1252</td>
<td>1.00E-06</td>
<td>25.23</td>
<td>3.52</td>
<td>1.3</td>
<td>3.71</td>
</tr>
<tr>
<td>2190</td>
<td>2.00E-06</td>
<td>17.15</td>
<td>1.75</td>
<td>0.04</td>
<td>1.84</td>
</tr>
<tr>
<td>2630</td>
<td>1.00E-06</td>
<td>18.14</td>
<td>2.81</td>
<td>1.1</td>
<td>2.79</td>
</tr>
<tr>
<td>2920</td>
<td>2.20E-05</td>
<td>25.01</td>
<td>4.98</td>
<td>1.04</td>
<td>4.75</td>
</tr>
<tr>
<td>3288</td>
<td>3.00E-06</td>
<td>11.61</td>
<td>2.14</td>
<td>1.46</td>
<td>1.91</td>
</tr>
<tr>
<td>3757</td>
<td>1.20E-05</td>
<td>26.61</td>
<td>3.76</td>
<td>1.63</td>
<td>4.22</td>
</tr>
<tr>
<td>4384</td>
<td>5.70E-05</td>
<td>35.34</td>
<td>4.34</td>
<td>2.86</td>
<td>3.92</td>
</tr>
<tr>
<td>5260</td>
<td>2.00E-06</td>
<td>36.26</td>
<td>5.11</td>
<td>2.6</td>
<td>5.69</td>
</tr>
<tr>
<td>6576</td>
<td>1.00E-06</td>
<td>21.84</td>
<td>2.35</td>
<td>0.78</td>
<td>2.93</td>
</tr>
<tr>
<td>8768</td>
<td>8.00E-06</td>
<td>16.55</td>
<td>1.81</td>
<td>0.72</td>
<td>1.94</td>
</tr>
</tbody>
</table>

Table 6.3: The total percentage of STA values at each frequency lower than the threshold indicated in the bottom row.

<table>
<thead>
<tr>
<th>Periods (hours)</th>
<th>Total %</th>
<th>STA (Wm(^{-2})k(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>99%</td>
<td>13</td>
</tr>
<tr>
<td>1252</td>
<td>95%</td>
<td>11</td>
</tr>
<tr>
<td>2190</td>
<td>99%</td>
<td>9</td>
</tr>
<tr>
<td>2630</td>
<td>95%</td>
<td>8</td>
</tr>
<tr>
<td>2920</td>
<td>80%</td>
<td>7</td>
</tr>
<tr>
<td>3288</td>
<td>95%</td>
<td>5</td>
</tr>
<tr>
<td>3757</td>
<td>80%</td>
<td>5</td>
</tr>
<tr>
<td>4384</td>
<td>80%</td>
<td>5</td>
</tr>
<tr>
<td>5260</td>
<td>85%</td>
<td>5</td>
</tr>
<tr>
<td>6576</td>
<td>90%</td>
<td>5</td>
</tr>
<tr>
<td>8768</td>
<td>90%</td>
<td>3</td>
</tr>
</tbody>
</table>

2\(^{nd}\) Experiment. After exploring the distribution of the STA values, we visualized the simulated SSTA graphs using the forward model (Eq. 6.11). Since there were no ground measurements of soil thermal properties (i.e. soil thermal conductivity, heat capacity) in the study area, we calculated them for different combinations of soil constituents and three soil layers using the procedure described earlier (see Section 6.2). We call these estimates the synthetic soil thermal properties. For instance, we assumed that the top layer is dry, the middle is wet, and the last semi – infinite layer water –
saturated. Figure 6.3 represents the simulated spectral soil thermal admittance using the forward model (Eq. 6.11) and the calculated soil thermal properties in Table 6.4. Figure 6.3 shows there is no peak in this graph and the thermal admittance decreases with increasing period. The results of the other numerical experiments (not shown) also reveal that the simulated SSTA’s show no peaks.

![Figure 6.3: Spectral soil thermal admittance using synthetic data.](image)

Table 6.4: Synthetic soil thermal properties assuming a three-layered soil.

<table>
<thead>
<tr>
<th>Soil layers</th>
<th>Volume of mineral ($x_m$)</th>
<th>Volume of organic matter ($x_o$)</th>
<th>Volume of water ($x_w$)</th>
<th>Volume of air ($x_a$)</th>
<th>Heat capacity ($\rho c$)</th>
<th>Thermal diffusivity ($a$)</th>
<th>Thermal conductivity ($\lambda$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>0.768E+6</td>
<td>0.937E-6</td>
<td>0.72</td>
</tr>
<tr>
<td>Middle</td>
<td>0.4</td>
<td>0.01</td>
<td>0.3</td>
<td>0.29</td>
<td>2.05E+6</td>
<td>0.46E-6</td>
<td>0.945</td>
</tr>
<tr>
<td>Infinite</td>
<td>0.4</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>3.314E+6</td>
<td>0.35E-6</td>
<td>1.16</td>
</tr>
</tbody>
</table>

3rd Experiment. To support the numerical experiment explained above, we also used the ground measurements of LST and $G_0$ at the BJ station (see Chapter 2) to generate the spectral soil thermal admittance using the procedure described in Chapter 5 (Figure 6.4, left). This graph again shows no peak.

To conclude, the peaks in the observed graph of SSTA are likely related to the error in the estimation of soil heat flux data (e.g. assuming a constant ratio of soil heat flux to net radiation) and have to be adjusted to get a smoothed graph as predicted by theory and by the BJ ground measurements. The procedure to smooth the SSTA graphs will be described in the next paragraph.

The result of FFT applied to hourly LST time series (Section 5.3.1) shows that the daily and yearly components are the most significant among the components which
account for almost 82% of variance in the LST time series. Therefore, the thermal admittance values at those two frequencies are more reliable than others. We keep the values of thermal admittance at those frequencies unchanged and smooth out the remaining values of soil thermal admittance.

To find out the outliers in SSTA graphs, the difference between thermal admittance value at 24-hour and yearly (i.e. 8768 hours) was calculated and then the values which have higher values than 25% of that difference are evaluated as potential outliers. We got then the mean of each identified outlier and its two adjacent values to get the new value. The iteration continues up to the time that all STA values recognized as outliers are less than 25% of the difference between 24-hour and yearly values. For instance, Figure 6.4 (right) represents the observed SSTA graph with outliers and after smoothing out the outliers. The dependence on the period is similar to the dependence found with both synthetic data and ground measurements.

![Figure 6.4: SSTA calculated using ground measurements at BJ (left) and using satellite data (right).](image)

### 6.3.2. Sensitivity analysis

The SSTA graph observed at the surface is determined by LST and \( G_0 \) amplitudes at different frequencies. The dependence of SSTA on the frequency of the periodic components of LST(t) represents the complex interactions of soil constituents at different depths in response to absorbed radiative energy at the surface (Menenti, 1984). Assuming a soil vertical profile with three different layers, the spectral soil thermal admittance can be calculated using Eq. 6.11. This equation requires eight input variables: three soil thermal conductivities (\( \lambda_1, \lambda_2 \) and \( \lambda_3 \)), three soil heat capacities (\( \rho c_1, \rho c_2 \) and \( \rho c_3 \)) and two soil thicknesses (\( z_1 \) and \( z_2 \)) (Figure 6.1). Table 6.5 gives the range
of variability of the input parameters necessary to simulate the spectral soil thermal admittance.

Given the SSTA observations, it is possible in principle to retrieve soil thermal properties by inverting the forward model of SSTA (i.e. Eq. 6.11). The retrievable parameters are those to which the observations are most sensitive. In a parameter retrieval procedure based on model inversion, if the number of unknown parameters exceeds the number of independent observations the problem becomes ill-posed. Here, in this study, there are eleven soil thermal admittance values at significant frequencies, identified in Chapter 5, (i.e. observations) versus eight input variables (i.e. unknowns); however, the number of independent observations is much lower due to the correlations between different observations. To avoid the problem becoming ill-posed, a limited number of variables have to be taken as free variables during inversion. Therefore, it is necessary to identify the most influential (and non-influential) parameters having higher (or marginal) influence on the output SSTA. This can be achieved through Sensitivity Analysis (SA) by systematic exploration of a broad range of input data variations on the output values. The sensitivity analysis quantifies the contribution of each model input to the output of a model (Mousivand et al., 2014; Saltelli et al., 2008; Saltelli et al., 2004).

The sensitivity of the simulated SSTA to the soil thermal properties of different layers was assessed by applying the approach of Mousivand et al. (2014) to carry out a variance-based sensitivity analysis (Saltelli et al., 2010; Saltelli et al., 2008). The method delivers accurate sensitivity measures because it scans the entire input space and decomposes the output variance into fractions, which quantify the contribution of each model input to the output variance. The decomposition of the output variance into partial variances of increasing dimensionality is:

\[
V(f) = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<k} V_{ijk} + \ldots + V_{12\ldots k}
\]

where \(V(f)\) is the total variance of the output function \(f\); \(V_i = V(E_{x\sim i} (f | x_i))\) represents the contribution of input \(x_i\) (first order variance contribution) to the total output variance; \(V_{ij} = V(E_{x\sim ij} (f | x_{ij}))\) is the contribution of the interaction between \(x_i\) and \(x_j\) to the total output variance; and so up to \(V_{12\ldots k}\), which is the contribution of the interactions of all parameters to the output variance. \(E_{x\sim i} (f | x_i)\) represents the expectation of \(f\) given \(x_i\), \(x\sim i\) means all the input parameters except \(x_i\) and \(V\) denotes variance over all the possible values of \(x_i\).
An extensive sensitivity analysis would require the calculation of a number of sensitivity indices. These indices can be used to interpret the final results of sensitivity analysis, and they are the first-order, second-order, higher order and total-order indices, however, it is in general sufficient to only use the total sensitivity (Saltelli et al., 2010). The first-order index represents the fraction of the total output variance related to a single parameter and it indicates the relative importance of that parameter on the model output. The second-order index indicates the mutual interactions effect between parameters and the total-order index is the first-order effect plus its interactions with all other parameters. Such indices can be evaluated as the ratio between the partial variances and the total variance.

\[
S_i = \frac{V_i}{V(f)} \quad (6.13)
\]

\[
S_{ij} = \frac{V_{ij}}{V(f)} \quad (6.14)
\]

\[
S_{ti} = 1 - \frac{V_{ti}}{V(f)} \quad (6.15)
\]

\(S_i, S_{ij}, \text{ and } S_{ti}\) are the first-order, second-order and total-order indices respectively. In this study, the total-order sensitivity index was used as a measure of the sensitivity of each individual variable since it includes the influence of that variable plus all the interaction effects with the other parameters (Mousivand et al., 2014). A number of 60000 synthetic SSTA values were generated to calculate the sensitivity of SSTA to the given parameters (see Table 6.5). The results and interpretation of the sensitivity analysis will be addressed in Section 6.4.6.

Table 6.5: Input parameters of the three-layered soil thermal admittance model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\lambda) (W/m K)</td>
<td>Soil thermal conductivity</td>
<td>0.15</td>
<td>4</td>
</tr>
<tr>
<td>(\rho c) (Jm(^{-3})k(^{-1}))</td>
<td>Volumetric soil heat capacity</td>
<td>1E+6</td>
<td>3.5E+6</td>
</tr>
<tr>
<td>(z) (m)</td>
<td>Soil depth</td>
<td>0.01</td>
<td>1</td>
</tr>
</tbody>
</table>
6.3.3. Parameter retrieval using the inversion method

The physical model in Eq. 6.11 relates soil thermal properties of a three-layered soil to the soil thermal admittance at the surface. This equation has to be inverted against the observed SSTA (from satellite data) in order to retrieve soil thermal properties. In this case, we may define the cost function as follows:

$$\min (y_{obs} - y_{mod})^T C_y^{-1} (y_{obs} - y_{mod})$$  \hspace{1cm} (6.16)

where \(y_{obs}\) denotes the observed SSTA, \(y_{mod}\) is the simulated SSTA using Eq. 6.11 and \(C_y\) is the observation covariance matrix. Since there were no measurements of the covariance between STA at different frequencies, we assumed the covariance matrix as an identity matrix in this case. The minimization of the cost function minimizes the difference between the measured and the simulated SSTA values. Since the physical forward model (Eq. 6.11) is highly nonlinear, we have to apply numerical optimization techniques that are capable of handling nonlinear problems. A numerical optimization technique starts with an initial guess and searches for the optimum parameter set that creates signatures similar to the observed values (minimizing the residuals) through an iterative process. The optimization technique has to be robust and reliable to find a global minimum of the cost function at a reasonable computational cost. We have used the Levenberg-Marquardt method which is a modified Gauss–Newton method with a damping factor. The Levenberg–Marquardt method is an iterative method that computes a new estimate (solution) at each iteration using a forward–model linearization at the current estimate (solution) (Mousivand et al., 2015). The update in each iteration is given by:

$$\Delta p = \left[ J^T C_y^{-1} J + \psi I \right]^{-1} \left[ J^T C_y^{-1} (y_{obs} - y_{mod}) \right]$$ \hspace{1cm} (6.17)

where \(\Delta p\) is the change in parameters required to reduce the difference \((y_{obs} - y_{mod})\) between the current model estimate and the measured one, \(J\) is the Jacobian matrix which is the partial derivatives of the model output with respect to parameters changes, \(I\) is identity matrix and \(\psi\) is the damping factor in the Levenberg-Marquardt algorithm that is adjusted during the course of the algorithm to ensure convergence. The iteration stops when either the convergence criterion is met or the user-defined number of iterations is reached (60000 iterations in this study). The convergence criterion is a threshold that causes the iterations to stop when the change \((\Delta p)\) is below this threshold (here is 1E-7).
It is important to evaluate the capability of the model inversion to minimize the difference between the observed and the simulated SSTA. We used the SSTA calculated from the ground measurements (from BJ station) of LST and $G_0$ during 2008-2010 for this purpose. We implemented the same method, as described in Chapter 5, to calculate the SSTA using these data (i.e. observations). The model inversion requires a forward model, a set of observations and a cost function that needs to be minimized plus an initial guess of the free parameters. In this case, Eq. 6.11 was used as the forward model, the calculated SSTA from the ground measurements are the observations and the median of the parameter bounds in Table 6.5 was used as the initial guess. Then, the inversion method was applied to the observed SSTA to investigate whether the simulated SSTA by the forward model was similar to the observed SSTA from ground measurements. Figure 6.5 shows the workflow of this experiment. Figure 6.6 shows the observed SSTA from ground measurements and simulated SSTA by forward model. The initial and final values of SSTA were in very good agreement with RMSE = 0.04 Wm$^{-2}$k$^{-1}$ and $R^2 = 0.9971$. The result demonstrates that the inversion of the forward model (Eq. 6.11) does not introduce additional errors of estimate. The small mismatch between observed and simulated SSTA is probably due to random errors in ground measurements.

![Diagram](image)

Figure 6.5: The workflow showing the procedure to evaluate the capability of the inversion method to minimize the residuals between the observed and simulated SSTA.
6.4. Results and discussions

6.4.1. Forward model validation

The spectral soil thermal admittance (SSTA) can be estimated either from the ratio between the amplitudes of \( G_0 \) and LST at the same frequency (see Chapter 5, Eq. 5.15) or using the forward model (Eq. 6.11). We hereafter shall refer to the former model as model 1 and the latter as model 2. Model 1 needs the amplitudes of \( G_0 \) and LST at significant frequencies measured (in situ) or estimated (from satellite observations) at the soil surface. Model 2 needs the soil thermal properties (i.e. \( \lambda \) and \( \rho_c \)) of different soil layers and their thickness (\( z \)). The soil thermal properties can be estimated either from the volumetric fractions of soil constituents (i.e. soil texture), water content, porosity, and organic carbon (Section 6.2) or directly measured in the field. Since model 1 used the data obtained at the soil surface and model 2 below the soil surface, we call them as a top – down and bottom – up procedure, respectively.

Here we want to evaluate the capability of the forward model (model 2) to simulate SSTA by comparing it with the model 1 as a reference model. To do so, the SSTA calculated with model 1 using the ground measurements of \( G_0(t) \) and LST\((t) \) (measured at BJ station) was compared to SSTAs simulated with model 2. The input data of model 2 (i.e. soil thermal properties) was either measured directly at BJ station or estimated using semi-empirical models (see Section 6.2). The soil constituents measured at BJ station were used as input parameters in semi-empirical models. The procedure of
applying model 1 to estimate SSTA has been described in detail in Chapter 5 and the results are presented in Figure 6.4 (left). The preparation of the input data for model 2 (i.e. the soil thermal properties) will be explained in the next section. The workflow of this procedure was illustrated in Figure 6.7.

6.4.2. Data preparation for forward model

As we mentioned before, the forward model requires the soil thermal properties of different soil layers (here we assumed three layers). The ground measurements of soil properties (i.e. volume of minerals, organic carbon, water, etc), needed to estimate soil thermal properties using semi-empirical methods (see Section 6.2.1 and 6.2.2) at distinct depths, have been obtained from the work of Chen et al. (2012) who measured them at different stations in the Tibetan Plateau. Specifically, we used data collected at the BJ station where we have already other data like LST, $G_0$ and soil water content (Table 6.6).
Table 6.6: The measured soil organic carbon mass content, porosity, soil gravel mass, bulk density, and soil texture at BJ station (Chen et al. 2012).

<table>
<thead>
<tr>
<th>Depth (cm)</th>
<th>( m_{soc} ) (%)</th>
<th>( \theta_{sat} ) (%)</th>
<th>( m_g ) (%)</th>
<th>( \rho_b ) (kg m(^{-3}))</th>
<th>% sand</th>
<th>% clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 10</td>
<td>2.1</td>
<td>44</td>
<td>2.1</td>
<td>1411</td>
<td>87.3</td>
<td>0.71</td>
</tr>
<tr>
<td>10 - 20</td>
<td>1.3</td>
<td>33</td>
<td>2.3</td>
<td>1712</td>
<td>94.18</td>
<td>0.01</td>
</tr>
<tr>
<td>20 - 30</td>
<td>1.1</td>
<td>37</td>
<td>7.9</td>
<td>1526</td>
<td>92.32</td>
<td>0.34</td>
</tr>
<tr>
<td>30 - 40</td>
<td>1.5</td>
<td>37</td>
<td>3.2</td>
<td>1560</td>
<td>84.05</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Table 6.6 contains the soil organic carbon mass content \( (m_{soc}) \), soil porosity \( (\theta_{sat}) \), gravel mass proportion \( (m_g) \), soil bulk density \( (\rho_b) \) and soil texture \( (% \text{ sand and } % \text{ clay}) \). The volumetric content of soil organic carbon can be estimated using the following formula:

\[
x_{oc} = \frac{\rho_p(1-\theta_m)m_{soc}}{\rho_p(1-m_{soc})+\rho_p(1-\theta_m)m_{soc}+(1-\theta_m)(\frac{\rho_m m_g}{1-m_g})} \tag{6.18}
\]

where \( \rho_p = 2700 \) (kgm\(^{-3}\)) is the mineral particle density, \( \rho_o = 130 \) (kgm\(^{-3}\)) is the bulk density of peat, \( \theta_m \) is the porosity of mineral soils without considering the organic matter which can be estimated as follows (Farouki, 1981):

\[
\theta_m = 0.489 - 0.00126 \times (\% \text{ sand}) \tag{6.19}
\]

The volumetric soil water content \( (\theta\%) \) at the BJ station has been measured every 30 minutes from 2008 to 2010 at two different depths of 4 (cm) and 20 (cm). The yearly averages of soil water content \( (\overline{\theta} \%) \) at these depths are listed in Table 6.7. We assumed the soil layers thickness as \( z_1 = 4 \) cm, \( z_2 \approx 20 \) cm and the semi-infinite layer as having a water content of 5%. This stratification was also confirmed by (Yang et al., 2005) who clearly observed the distinct layers as defined above.

Table 6.7: The yearly average of volumetric soil water content at BJ station during 2008-2010.

<table>
<thead>
<tr>
<th>Depth</th>
<th>( \overline{\theta} % ) (2008)</th>
<th>( \overline{\theta} % ) (2009)</th>
<th>( \overline{\theta} % ) (2010)</th>
<th>( x_w = \text{Mean}(\overline{\theta} %) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 4 (cm)</td>
<td>16.23</td>
<td>14.3</td>
<td>15.15</td>
<td>15.22</td>
</tr>
<tr>
<td>4 – 20 (cm)</td>
<td>5.9</td>
<td>5.54</td>
<td>6.4</td>
<td>5.96</td>
</tr>
<tr>
<td>Semi-infinite</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>
6.4.3. Estimation of soil thermal properties

The soil parameters needed to calculate soil thermal properties are listed in Table 6.8 as measured or estimated at the BJ station. The soil heat capacity ($\rho c$) was estimated using the Eq. 6.8 and Eq. 6.9 with the data in Table 6.1 and Table 6.8. We used both the De Vries and the Yang model (see Section 6.2) to calculate the soil thermal conductivity ($\lambda$). The soil thermal conductivity was also measured (using the KD2 Thermal Properties Analyzer instrument) at BJ station for four different layers at the same depths as in Table 6.6 by Yang et al. (2005). They presented a graph showing the measured soil thermal conductivity against soil water content. We used that graph knowing the water content and got the related soil thermal conductivity of different layers (Table 6.8). So we have three sources of data for soil thermal conductivity for the same location at BJ coming from three different methods namely are as follows:

1) the measured soil thermal conductivity at the BJ station (method 1).
2) the estimated soil thermal properties using the De Vries model (method 2).
3) the estimated soil thermal properties using the Yang et. al model (method 3).

<table>
<thead>
<tr>
<th>$z$ (cm)</th>
<th>$\alpha$</th>
<th>$\alpha_o$</th>
<th>$\alpha_w$</th>
<th>$\varphi_o$</th>
<th>$\rho c$ (Jm$^{-3}$k$^{-1}$)</th>
<th>$\alpha$ (m$^2$s$^{-1}$)</th>
<th>$\lambda$ (Wm$^{-1}$K$^{-1}$) (method 1)</th>
<th>$\lambda$ (Wm$^{-1}$K$^{-1}$) (method 2)</th>
<th>$\lambda$ (Wm$^{-1}$K$^{-1}$) (method 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.44</td>
<td>0.13</td>
<td>0.15</td>
<td>0.28</td>
<td>1.95E+6</td>
<td>0.56E-6</td>
<td>0.87</td>
<td>0.68</td>
<td>0.44</td>
</tr>
<tr>
<td>20</td>
<td>0.57</td>
<td>0.1</td>
<td>0.06</td>
<td>0.27</td>
<td>1.82E+6</td>
<td>0.49E-6</td>
<td>0.93</td>
<td>0.69</td>
<td>0.42</td>
</tr>
<tr>
<td>Infinite</td>
<td>0.54</td>
<td>0.09</td>
<td>0.05</td>
<td>0.32</td>
<td>1.68E+6</td>
<td>0.12E-6</td>
<td>0.76</td>
<td>0.36</td>
<td>0.31</td>
</tr>
</tbody>
</table>

6.4.4. The results of validation

We evaluated the SSTA calculated (using ground measurements) as the ratio between the amplitudes of $G_0$ and LST (model 1) as a reference for the three SSTAs estimated from the forward model (model 2) and the thermal properties as detailed above. Figure 6.8 shows the SSTA variation using all the methods mentioned above. The results indicate that the SSTA calculated from model 1 and SSTA from model 2 using all three methods of estimating soil thermal properties are in good agreement with each other at all periods except for short periods (12- and 24-hour).

The goodness of fit between the estimated SSTA applying model 1 and model 2 (forward model) using three methods (method 1, 2 and 3) was calculated based on
RMSE, $R^2$, bias and MAE (Table 6.9). The results of calculated bias show that model 2 overestimates the SSTA especially for the short periods (12- and 24- hour) using all three methods (method 1,2 and 3). This might be due to the use of the yearly averaged soil water contents, which can be representative of the soil water content over the longer periods but less so for the daily and 12-hour periods. The $R^2 > 0.99$ using all three methods shows that the forward model and model 1 are completely correlated, and the sign of changes is the same. The results also show that using method 3 gives lower RMSE = 0.46 (Wm$^{-2}$K$^{-1}$) and MAE = 0.32 (Wm$^{-2}$K$^{-1}$) compared to method 2 and method 1.

To explain the deviation between SSTA calculated from model 1 and model 2, we calculated the STA for periods 12- and 24-hour using ground measurements of daily water content applying method 3 (as it has better results than method 1 and method 2). The mean and error bar (1 standard deviation) of calculated STAs for those periods using daily measurements of $\theta$ (1096 days over three years) are shown in Figure 6.8 (mean daily STA for 12- and 24-hour period). The conclusion is that the estimated SSTA from data on the soil surface using model 1 can be reproduced with acceptable accuracy using model 2 (forward model) and the data measured under the soil surface.

![Figure 6.8](image-url)

**Figure 6.8:** The SSTA estimated by model 1 using ground measurements of $G_0$ and LST at BJ station (blue line), by model 2 using the measured soil thermal properties (red line), by model 2 using estimated soil thermal conductivity from De Vries model (green line) and, by model 2 using estimated soil thermal conductivity from Yang model (method 3) (violet line); Mean daily STA for 12- and 24-hour periods.
Table 6.9: The statistical measures of goodness of fit between SSTA estimated by model 1 and model 2.

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE (Wm$^{-2}$K$^{-1}$)</th>
<th>$R^2$</th>
<th>Bias (Wm$^{-2}$K$^{-1}$)</th>
<th>MAE (Wm$^{-2}$K$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method (1)</td>
<td>1.31</td>
<td>0.9967</td>
<td>-0.664</td>
<td>0.7</td>
</tr>
<tr>
<td>Method (2)</td>
<td>2.07</td>
<td>0.9967</td>
<td>-1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>Method (3)</td>
<td>0.46</td>
<td>0.9968</td>
<td>-0.263</td>
<td>0.32</td>
</tr>
</tbody>
</table>

6.4.5. Validation of estimated SSTA from satellite data

The SSTA estimated from time series satellite observations of $G_0$ and LST was compared to the SSTA calculated from ground measurements of $G_0$ and LST applying model 1 at BJ station (Figure 6.9).

The results imply that there are discrepancies between the SSTA graphs especially at 24-hour periods. There are some reasons that may explain this finding:

1- the difference in footprint between the ground measurement and satellite observations is large, since the satellite pixel size is 5x5 (Km) while the ground observation is just a point in this pixel (scale problem).

2- the accuracies of estimated $G_0$ from satellite observations are low. Figure 6.10 represents the scatter plot between measured and estimated LST (left) and $G_0$ (right). The estimated LST values from satellite data are in far better
agreement with the LST measured in situ. i.e. \( R^2 = 0.7 \), than the estimated instantaneous \( G_0 \) values with the measured ones. i.e. \( R^2 = 0.07 \). This is because of using a rather rough approach to estimate soil heat flux and assuming clear sky conditions throughout the day (Faivre, 2014). The estimated soil heat flux was used to estimate the sensible (H) and latent (\( \lambda E \)) heat flux by Faivre (2014). The evaluation of estimated heat flux densities revealed that estimated sensible and latent soil heat flux were in good agreement with ground measurements with RMSE = 33.88 (Wm\(^{-2}\)) and 35.66 (Wm\(^{-2}\)), respectively (Faivre, 2014).

![Figure 6.10: The estimated and measured LST (left) and \( G_0 \) (right).]

3- the daily amplitudes of estimated \( G_0 \) at 24-hour period are underestimated while the daily amplitudes of estimated LST at 24-hour period are overestimated. To describe the divergences between estimated and measured soil thermal admittance at the 24-hour period, we estimated the amplitudes of estimated and measured LST and \( G_0 \) on each day (over 3 years) and the scatter plots were created (Figure 6.11). The soil thermal admittance at any specific period is the ratio between the amplitude of soil heat flux and LST, so, when the estimated amplitude of \( G_0 \) is underestimated (bias = 28.53) and the LST overestimated (bias = -4.85), the values of estimated soil thermal admittance would be less than calculated ones (Figure 6.12). The two errors appear to compensate each other since the correlation of satellite vs. ground daily STA is better than either LST and \( G_0 \) amplitude and the RMSE, MAE values are acceptable.
6.4.6. Sensitivity analysis of SSTA response to soil thermal properties

The variance-based global sensitivity analysis of SSTA for different input variables at different frequencies has been applied using the procedure introduced by (Mousivand et al., 2014). The input parameters and their range of variation are shown in Table 6.5. Figure 6.13 displays the total-order sensitivity of SSTA to the soil thermal properties of three different soil layers at the predefined frequencies. The parameters with low total-order variance (i.e. namely non-influential parameters) can be kept constant as they have little or no influence on the total variance of the model output. The results indicate that for instance the soil heat capacity and soil thermal conductivity of the third layer ($\rho c_3$ and $\lambda_3$) as well as the soil thickness of the second layer ($z_2$) have a
small effect on the output SSTA at almost all frequencies, while the soil thermal conductivity of the first layer ($\lambda_1$) has a large effect on the total variance (between ~65% to ~45%). This is explained by the fact that a small change in soil water content in the top soil layer has greater influence on soil thermal conductivity than on the soil heat capacity. This is because the heat capacity has a linear dependency with water content while soil thermal conductivity depends non-linearly on water content. As expected, this confirms that the soil thermal conductivity in the top soil is the most influential parameter. The soil thickness of the second layer ($z_2$) has less impact on the total variance of the model output at short periods (i.e. 24-hour period) than long period (i.e. yearly period), however its overall impact on total variance across all frequencies is less than other parameters, therefore, we can set $z_2$ constant (i.e. $z_2 = 1$ m).

![Sensitivity analysis of thermal admittance model assuming three-layers soil depth](image)

**Figure 6.13:** Sensitivity of spectral soil thermal admittance assuming three-layer soil across all frequencies.

### 6.4.7. Parameter retrieval

The inversion method was applied to estimate the soil thermal properties assuming a three layers soil from satellite estimates of SSTA for the whole Tibetan plateau. We used Eq. 6.11 as a forward model. The consistency and reliability of this equation for simulating the SSTA will be tested prior to parameter retrieval. Then, the parameter estimation model will be validated using synthetic data. Finally, the soil thermal properties maps for different soil layers in the study area will be retrieved. In the next sections, we will focus on the above-mentioned challenges.
6.4.8. Validation of model inversion

As no ground measurements of all required soil thermal properties were available, the validation of the parameter estimation approach was performed on synthetic data. The synthetic SSTA data set was first simulated by the forward model assuming a three layers soil using synthetic soil thermal properties (Table 6.4) calculated by De Vries model. We used De Vries model because it is generally the most well-known semi-empirical model and more widely used in literature. The synthetic SSTA observations obtained with the forward model was then inverted to retrieve the soil thermal properties. Figure 6.14 illustrates the above – mentioned procedure.

Table 6.10 lists the results of parameters estimation assuming three soil layers. The accuracy of estimated soil thermal properties from inversion model against the ones calculated with the semi-empirical model was shown in Figure 6.15 with $R^2 = 0.94$. We carried out several numerical experiments with different parameterizations (not shown) and we found out that the final accuracy of results was clearly related to the results of the sensitivity analysis. For example, the soil thermal conductivity of first layers ($\lambda_1$) is the most influential parameter on SSTA and it can be retrieved with higher accuracy (Table 6.10). This is confirmed by the results in Table 6.10 (0.72 Wm$^{-1}$K$^{-1}$ against 0.718 Wm$^{-1}$K$^{-1}$). The results on the first layer soil heat capacity ($\rho c_1$) and the soil thickness ($z_1$) were similar, in the sense that these parameters were the 2$^{nd}$ and 3$^{rd}$ influential parameters respectively and the agreement was good. So these numerical experiments show that it is possible to retrieve the soil thermal properties with $R^2 = 0.7$ to 0.9.

![Figure 6.14: The workflow used to validate the inversion method.](image-url)
Table 6.10: The calculated and estimated soil thermal properties for a three layers soil.

<table>
<thead>
<tr>
<th>Soil thermal properties</th>
<th>$\rho c_1$ (Jm$^{-3}$k$^{-1}$)</th>
<th>$\rho c_2$ (Jm$^{-3}$k$^{-1}$)</th>
<th>$\rho c_3$ (Jm$^{-3}$k$^{-1}$)</th>
<th>$a_1$ (m$^2$s$^{-1}$)</th>
<th>$a_2$ (m$^2$s$^{-1}$)</th>
<th>$a_3$ (m$^2$s$^{-1}$)</th>
<th>$\lambda_1$ (Wm$^{-1}$K$^{-1}$)</th>
<th>$\lambda_2$ (Wm$^{-1}$K$^{-1}$)</th>
<th>$\lambda_3$ (Wm$^{-1}$K$^{-1}$)</th>
<th>$z_1$ (m)</th>
<th>$z_2$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculated</td>
<td>$0.768E+6$</td>
<td>$2.05E+6$</td>
<td>$3.31E+6$</td>
<td>$0.937E-6$</td>
<td>$0.46E-6$</td>
<td>$0.35E-6$</td>
<td>$0.72$</td>
<td>$0.945$</td>
<td>$1.16$</td>
<td>$0.04$</td>
<td>$0.1$</td>
</tr>
<tr>
<td>Estimated</td>
<td>$0.769E+6$</td>
<td>$1.76E+6$</td>
<td>$2.36E+6$</td>
<td>$0.933E-6$</td>
<td>$0.64E-6$</td>
<td>$0.69E-6$</td>
<td>$0.718$</td>
<td>$1.125$</td>
<td>$1.628$</td>
<td>$0.039$</td>
<td>$0.117$</td>
</tr>
</tbody>
</table>

Figure 6.15: Scatter plot of the retrieved and calculated soil thermal properties.

To support experiments implemented in Section 6.3.1 to smooth the observed SSTA calculated using satellite observations, here, we compared the retrieved soil thermal properties applying the parameter retrieval method before and after smoothing. As an example, the histograms of the retrieved soil thermal conductivity of the first layer ($\lambda_1$) and soil heat capacity of the first layer ($\rho c_1$) before and after smoothing are shown in Figure 6.16 and Figure 6.17 respectively. The results show that histograms are reasonably similar. From all experiments in Section 6.3.1 and these results, we concluded that smoothing the peaks (outliers) from observed SSTA had no critical impact on retrieved soil thermal properties.

In the next section, we will apply the parameter retrieval procedure on the calculated SSTA from satellite data to retrieve the soil thermal properties in the whole Tibetan Plateau. The results will be a set of soil thermal properties maps assuming a three-layered soil.
6.4.9. Soil thermal properties maps

We applied the inversion method to the calculated SSTA from satellite data to retrieve pixel by pixel the soil thermal properties assuming a three-layered soil in the whole study area. As we mentioned before, we fixed the soil thickness of the 2nd layer ($z_2$) at 1 m and retrieved all seven remaining parameters, i.e. the soil thermal conductivity ($\lambda$) and heat capacity ($\rho c$) of the three soil layers (Figure 6.18).

It can be seen that the values of thermal conductivity range from 0.15 to 4 (Wm$^{-1}$K$^{-1}$) over the entire study area and in each layer (Table 6.11). The thermal conductivity of the 1st layer ($\lambda_1$) was mostly between 0.15 to 0.3 (Wm$^{-1}$K$^{-1}$) while in the south west of the study area the $\lambda_1$ values were from 0.4 to 4 (Wm$^{-1}$K$^{-1}$). The second layer had $\lambda_2$ values mostly in the high range of thermal conductivity, i.e. 0.7 to 4 (Wm$^{-1}$K$^{-1}$) and the third layer had a combination of low and high values (two distinct patterns) of thermal conductivity, with low range 0.15 to 0.18 (Wm$^{-1}$K$^{-1}$) and high range 0.7 to 4 (Wm$^{-1}$K$^{-1}$).
The values of soil heat capacity ($\rho c$) range from $1\times10^6$ to $3.5\times10^6$ (Jm$^{-3}$K$^{-1}$) covering the entire range of heat capacity values defined in Table 6.5. Comparing the pattern of $\rho c_1$ and $\lambda_1$, it is evident that there was a clear relationship between the two parameters in which the areas with high values of $\lambda_1$ also had high values of $\rho c_1$ and vice versa. This can be explained by the previous results on thermal admittance maps especially for daily admittance (Figure 5.10). For a given heat input, a large heat capacity values causes small temperature changes (low LST amplitude), which then indicates large thermal admittance and the large values of soil thermal conductivity mean more heat is taken away from the upper layer and so the LST rises less (low LST amplitude) and also leads to large thermal admittance. The same evidence can be seen from our results, in which the areas having high values of $\lambda_1$ and $\rho c_1$ have large values of soil thermal admittance (see daily thermal admittance map in Figure 5.10). Furthermore, we can say that both $\lambda$ and $\rho c$ increase with the soil water content.

Table 6.11: Basic statistics of soil thermal properties maps and soil depth map.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho c_1$ (Jm$^{-3}$K$^{-1}$)</td>
<td>$1.00\times10^6$</td>
<td>$3.50\times10^6$</td>
<td>$1.12\times10^6$</td>
<td>$2.67\times10^5$</td>
</tr>
<tr>
<td>$\rho c_2$ (Jm$^{-3}$K$^{-1}$)</td>
<td>$1.00\times10^6$</td>
<td>$3.50\times10^6$</td>
<td>$2.92\times10^6$</td>
<td>$8.00\times10^5$</td>
</tr>
<tr>
<td>$\rho c_3$ (Jm$^{-3}$K$^{-1}$)</td>
<td>$1.00\times10^6$</td>
<td>$3.50\times10^6$</td>
<td>$2.65\times10^6$</td>
<td>$1.15\times10^6$</td>
</tr>
<tr>
<td>$\lambda_1$ (Wm$^{-1}$k$^{-1}$)</td>
<td>0.15</td>
<td>4</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>$\lambda_2$ (Wm$^{-1}$k$^{-1}$)</td>
<td>0.15</td>
<td>4</td>
<td>3.5</td>
<td>1.12</td>
</tr>
<tr>
<td>$\lambda_3$ (Wm$^{-1}$k$^{-1}$)</td>
<td>0.15</td>
<td>4</td>
<td>2.6</td>
<td>1.78</td>
</tr>
<tr>
<td>$z_1$ (m)</td>
<td>0.01</td>
<td>1</td>
<td>0.12</td>
<td>0.17</td>
</tr>
</tbody>
</table>

The soil thickness of the 1st layer varies from 1 cm to 1 meter with an average of 12 cm (Figure 6.19 and Table 6.11). The areas with large soil thickness are located in the South-West, partly North and West with almost 1 meter depth. The central, East and South parts of the study area have shallower soils with thickness between 1 to 10 cm (Figure 6.19).

In order to validate the capability of the forward model to simulate the SSTA based on a given range of parameters via the applied inversion method, the original 11 soil thermal admittance maps created in Chapter 5 and the soil thermal admittance maps simulated by forward model from the retrieved soil thermal properties were classified by the k-means method assigning 10 classes, and then the resulting maps were compared. The result implies that the calculated and simulated SSTA maps gave rather
similar patterns, thus confirming that the inversion model does not introduce errors (Figure 6.20, left).

The SSTA calculated from satellite data for each pixel was compared to the simulated one. The MAE map shows high and low values between 5.2 and 0.025 (Wm\(^{-2}\)K\(^{-1}\)) respectively. The RMSE values vary between 0.029 and 9.4 (Wm\(^{-2}\)K\(^{-1}\)). The fraction of pixels with RMSE (MAE) > ~1.8 (> ~1.4) Wm\(^{-2}\)K\(^{-1}\) is ~5% of total number of pixels in the study area, which are along the top of Himalaya mountain range where the STA at some frequencies is extremely large because of inconsistency between soil heat flux and LST.

The retrieved soil thermal properties were evaluated against soil thermal properties estimated globally using Pedotransfer Functions (PTFs) established as described by Dai et al. (2013) and Shangguan et al. (2013). The data set on the required soil properties, i.e. soil texture, bulk density and soil organic matter, was generated by Shangguan et al. (2013) using about 9000 soil profiles in China, with a relatively high soil sampling density in Tibet. This data set includes several soil properties for eight soil horizons gridded at 1 km x 1 km.

The evaluation was done extracting first ten samples from our retrieved thermal properties at locations chosen to span the full range of thermal admittance values. Since our data set has a grid size of 5 km x 5 km we averaged 25 samples at the 1 km x 1 km grid size and compared the results (Table 6.12).

<table>
<thead>
<tr>
<th>Latitude (°N)</th>
<th>Longitude (°E)</th>
<th>Estimated λ of dry soil using soil texture (Wm(^{-2})K(^{-1}))</th>
<th>Retrieved λ (Wm(^{-2})K(^{-1}))</th>
<th>Estimated λ of saturated soil using soil texture (Wm(^{-2})K(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.50599884</td>
<td>84.67251363</td>
<td>0.2</td>
<td>0.15</td>
<td>2.7</td>
</tr>
<tr>
<td>37.18723273</td>
<td>104.37948137</td>
<td>0.2</td>
<td>0.16</td>
<td>2.2</td>
</tr>
<tr>
<td>38.30965691</td>
<td>86.41136674</td>
<td>0.2</td>
<td>0.247</td>
<td>2.7</td>
</tr>
<tr>
<td>39.90340317</td>
<td>70.79987479</td>
<td>0.1</td>
<td>0.354</td>
<td>1.6</td>
</tr>
<tr>
<td>26.42203621</td>
<td>74.92625027</td>
<td>0.2</td>
<td>0.471</td>
<td>2</td>
</tr>
<tr>
<td>27.36392331</td>
<td>71.43286005</td>
<td>0.2</td>
<td>0.536</td>
<td>2.1</td>
</tr>
<tr>
<td>26.3366072</td>
<td>89.86390297</td>
<td>0.2</td>
<td>0.616</td>
<td>2.1</td>
</tr>
<tr>
<td>31.7596002</td>
<td>73.25797403</td>
<td>0.2</td>
<td>0.709</td>
<td>2.2</td>
</tr>
<tr>
<td>25.21117319</td>
<td>72.55952078</td>
<td>0.2</td>
<td>0.914</td>
<td>2.1</td>
</tr>
<tr>
<td>26.20843924</td>
<td>85.57632265</td>
<td>0.2</td>
<td>1.664</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Table 6.12 compares the estimated soil thermal conductivity of dry and saturated soil with the retrieved soil thermal conductivity using inversion model for some pixels in the study area. The results shows that the retrieved soil thermal conductivity is correctly within the range of the reference soil thermal conductivity for dry and saturated conditions. The reference data on thermal properties give only the dry and wet limits since soil water content, which changes over time, is not taken into account.

6.5. Conclusion

In Section 6.4.1, we discussed about the accuracy of the forward model to simulate SSTA using a given range of soil thermal properties. There was some discrepancy between observed and simulated SSTA as the RMSE map (Figure 6.20, right) shows. Besides the uncertainty in calculated $G_0$, which is due to the assumption of clear sky and using just instantaneous values, there are some other issues which may cause this mismatch. Also, we have assumed a three – layers soil everywhere, while in reality the soil vertical structure may be different.

The effective soil thermal properties of different soil layers can be estimated by inverse modelling of spectral soil thermal admittance determined with time series of satellite observations of land surface temperature and soil heat flux. The main point in this approach is that the observations of the surface thermal admittance at different frequencies (SSTA) provide information on the vertical structure of the soil.

In this study, we applied an inversion method to estimate the soil thermal properties (i.e. soil thermal conductivity and heat capacity) over the Tibetan Plateau. The results of this study demonstrated a new approach to estimate soil thermal properties using satellite data over the large area. It has been shown that by only having the time series of satellite observations of LST and $G_0$, it is possible to retrieve the soil thermal properties.
Figure 6.18: Soil thermal conductivity map of first ($\lambda_1$), second ($\lambda_2$) and third ($\lambda_3$) layer (left); soil heat capacity map of first ($\rho c_1$), second ($\rho c_2$) and third ($\rho c_3$) layer (right) of a three-layer soil in Tibetan Plateau.
Figure 6.19: The first layer soil thickness map ($z_1$).

Figure 6.20: The calculated and simulated classified map of SSTA (left); The MAE and RMSE between calculated and simulated SSTA profile (right).
Chapter 7

7. Conclusions and recommendations

The main purpose of this chapter is to summarize the findings of the current study in the light of the objectives stated in Chapter 1. In Section 7.1, a brief summary of this study is presented. Section 7.2 lists the conclusions related to each separate research objective. Finally, recommendations for future work are put forward in Section 7.3.

7.1. Brief summary

The soil thermal properties are of great importance for a variety of applications that require quantitative information on soil heat transfer. There are different methods either to measure soil thermal properties in situ or estimate them empirically using measurements of most influential factors, e.g. soil texture, affecting soil thermal properties. However, the use of such methods over large and heterogeneous regions like the Qinghai – Tibet Plateau becomes rather challenging, since these methods are often time-consuming, costly and impractical for large areas. The aim of this thesis was to propose and evaluate an approach to retrieve the soil thermal properties by inverse modeling of the spectral soil thermal admittance estimated using time series of satellite observations of LST and soil heat flux over large areas. The key element of this approach is that the observations of soil thermal admittance at different frequencies provide information about the soil structure (e.g. the soil thermal properties of different layers). But, such investigation needs continuous, consistent and accurate time series data of LST and soil heat flux over long periods of time (e.g. one year) with high temporal resolution (e.g. hourly). Such time series of observations are often contaminated by gaps and outliers. Therefore, the second element of this work was to determine how to identify and remove outliers and then fill the gaps that already existed in the time series plus the ones created by removing the outliers. Gap-filled LST data was used to identify most significant periodic components over three years. The
amplitudes of LST and soil heat flux at those periods (frequencies) were estimated to determine the spectral soil thermal admittance.

7.2. Results versus research objectives

1st objective: Gap-filling and outlier removal of hourly time series satellite observation of LST.

The HANTS algorithm was used to fill the gaps in the hourly LST time series observed by the FY-2C satellite. Based on the experiments described in Chapter 3, the accuracy of the HANTS algorithm to fill short gaps was evaluated. A short (long) gap is shorter (longer) than the period of time sampled by half of the observation points (Jia et al., 2011; Roerink and Menenti, 2000a). We used both the old and the new version of HANTS in gap-filling of three days and one-month hourly LST time series. The old version allows users to choose significant periodic components independently from each other to reconstruct a given time series and performed better than the new version. One of the difficulties in using HANTS is that there is no objective way to set parameter values and it takes some testing to find the most suitable parameter setting for a specific data set. One way is to use FFT to find the dominant periodic components in a time series. Even though the FFT helps to identify the significant periodic components, to obtain accurate results different combinations of parameters have to be tested. The accuracy of gap-filling was evaluated by using both ground measurements of LST and synthetic data in which the gaps were artificially imposed. The evaluation of gap-filling by HANTS applied to these gappy data with different parameterizations showed that by increasing the number of frequency (NOF) the accuracy was improved. However, it should be noted that when the total number of gaps in gapped data was higher than half of data points and the longest continuous gaps were located at the beginning or at the end of the time series, the reconstructed signal was not reliable. The results of applying HANTS on synthetic data show that when the total number of gaps reach 50-60% of total number of observations, the RMSE will be around 4K, and the same RMSE will be achieved, when the longest continuous gaps are up to 20-30%. The results of this evaluation suggested that a method considering both spatial and temporal correlation to fill the gaps could perform better than HANTS, which considers temporal correlation for each pixel only.

In Chapter 4, the (M)-SSA algorithm that uses both spatial and temporal correlation at the same time to fill the gaps was applied for gap-filling, noise and outlier
removal in the same hourly LST time series. The LST time series had in some cases rather long and continuous gaps (sometimes up to 70-80% of data points) in time. The results obtained by applying (M)-SSA showed the capability and usefulness of this method to fill the gaps with acceptable accuracy even when extended and continuous gaps were present. Different methods have been tested to identify the significant periodic components and window size to process the LST time series using (M)-SSA. Applying (M)-SSA on monthly segments of hourly LST data, we selected window size = 72 hours and number of components = 7 (the significant periodic components with 97.5% significant level were 24, 12 and 8 hours when the window size was 72 hours ). The performance of (M)-SSA was evaluated using both LST ground measurements and synthetic data with different gap number, gap size and location of gaps. The result of applying SSA on LST ground measurements implied that even with total 63% of gaps, the $R^2 = 0.83$ with $\text{MAE} = 2.25$ K. The evaluation results on synthetic data can help users to assess the expected accuracy before applying the M-SSA in a time series with a known gaps distribution. The results of applying M-SSA for gap-filling of synthetic data with 18054 realizations show that when the total number of gaps reach up to 60-70%, the expected RMSE will be 3-4 K and when the longest continuous gaps are up to 50% the RMSE will be around 4 K. The drawback of the M-SSA is the computational cost due to M-SSA executing a very large number of calculations iteratively. For example, applying M-SSA to a monthly multi-channel time series with 18054 pixels, each containing 744 data, takes 2.5 hours (e.g. on a regular computer with 64bit, 16GB RAM and CPU 2.8 Dual Core). The gap-filling procedure was applied on a monthly basis for three years of hourly LST time series covering the entire study area. The comparison between the actual and reconstructed data of the successive months shows that the monthly LST data are connected reasonably well together with acceptable accuracy after gap-filling to create a three-year continuous data set.

2nd objective and 3rd objective: Estimate the amplitude of LST and soil heat flux at significant periodic components to determine the spectral soil thermal admittance.

The feasibility of using gap-filled time series satellite observations of LST and soil heat flux for estimating soil thermal admittance was investigated in **Chapter 5**. The developed method is simple and it needs minimum input data compared to other methodologies. The data needed for estimation of spectral soil thermal admittance are the amplitudes of LST(t) and $G_0(t)$ periodic components, which can be estimated by
means of Fourier series and least square method, assuming the periodicity of both time
series. The most significant periodic components in LST time series were identified by
applying Fast Fourier Analysis (FFT) and power spectrum analysis over three-year LST
data. The results implied that the most significant periods were 365 and 1 days
respectively with total power of 82.5%. We selected 11 significant periods which
contain around 88.4% of total power. They were yearly, daily, 1252 hours (1.74
months), 2190 hours (~3 months), 2630 hours (~3.65 months), 2920 hours (~4 months),
3288 hours (~4.5 months), 3757 hours (~5 months), 4384 hours (~6 months), 5260
hours (~7 months) and 6576 hours (~9 months). After selecting the most significant
periodic components, the corresponding amplitudes at those periods were calculated
pixel by pixel covering entire Tibetan plateau. When the available time series was
incomplete (i.e. time series of $G_0$) and consisted of just the daily mean and one
instantaneous value in each day, it was possible to derive the amplitude of time series
components with good accuracy and estimate the soil thermal admittance. The soil
thermal admittance values at significant frequencies, i.e. the Spectral Soil Thermal
Admittance (SSTA), were calculated over the entire study area having the amplitudes of
LST and soil heat flux at those significant frequencies. SSTA contains valuable
information about the soil thermal properties of different soil layers. The thermal
properties of soil layers can be estimated indirectly using the spectral soil thermal
admittance.

4th objective: Estimate the soil thermal properties at different soil depths by inversion of
the forward model.

The effective soil thermal properties of different soil layers (in this study, we
assumed a three-layer soil) were estimated by inverse modelling of the spectral soil
thermal admittance calculated by time series satellite observations of land surface
temperature and soil heat flux. In Chapter 6, we applied the inversion model to
estimate the soil thermal properties (i.e. soil thermal conductivity and heat capacity) of
the Tibetan Plateau. The accuracy of the forward model (Eq. 6.11) was evaluated by
comparing the SSTA calculated using the amplitudes of ground measurements (at BJ
station) of LST and $G_0$ (as a reference) against the SSTAs simulated by the forward
model in which the input soil thermal properties were measured directly or estimated by
semi-empirical models at the BJ station. The results implied that the forward model can
simulate the SSTA with a rather good accuracy. The RMSE and $R^2$ were 0.46 Wm$^{-2}$k$^{-1}$
and 0.99 when the Yang et. al model (Yang et al., 2005) was used to estimate the soil thermal properties needed as input in the forward model. In order to find out the most influential (and non-influential) parameters having higher (or marginal) influence on the simulated SSTA, a sensitivity analysis was conducted. The results showed that from eight retrievable parameters (three soil thermal conductivities, three soil heat capacities and two soil thicknesses for each separate soil layer assuming a three-layer soil), the soil thermal conductivity of the first layer \( \lambda_1 \) had the largest effect on the total variance (between \( \sim 45\% \) to \( \sim 65\% \)) of simulated SSTA by the forward model across all frequencies, while the soil thickness of second layer \( z_2 \) had the lowest impact on total variance of forward model. Therefore, we assigned \( z_2 \) a constant values of one meter and retrieved all seven remaining parameters. To validate the inversion model, the synthetic soil thermal properties were used. The synthetic soil thermal properties were calculated assuming a three-layer soil with different soil constituents in each layer and applying the De Vries model (De Vries, 1963). Then, the SSTA simulated by the forward model using the synthetic soil thermal properties was estimated by inverse modelling to retrieve the soil thermal properties. The comparison between the calculated and retrieved soil thermal properties showed that the inversion model can retrieve the soil thermal conductivity, soil heat capacity and thickness of the first layer very close to the reference (input) values and the soil thermal properties of the second and third layer with acceptable accuracy (with overall \( R^2 = 0.94 \)). The results of this study documented the feasibility of a new approach to estimate soil thermal properties using satellite data over large areas. It has been shown that by only having the time series satellite observation of LST and \( G_0 \), it is possible to retrieve the soil thermal properties. The agreement between synthetic and estimated parameters was rather satisfactory.

7.3. Recommendations

Throughout the current research several techniques have been introduced to support the main aim of this study to utilize the time series satellite data of LST and soil heat flux to estimate the soil thermal properties. The investigations on the soil thermal properties retrieval in Chapter 6 are a first attempt to use time series satellite observations of LST and soil heat flux over very large areas. However, this approach still suffers from some limitations due to the constraints faced in this research. The following recommendations are proposed for future work:
1. In Chapter 4, the M-(SSA) was applied on monthly segments of hourly LST time series due to computational and memory constraints. To be sure that successive monthly gap-filled LST data were connected correctly, the length of time series should be increased slightly in each month by adding extra observations from previous and next month to create some overlap at the beginning and end of each month after reconstruction. The average of reconstructed values across the overlap will give better results.

2. In Chapter 5, the amplitude of soil heat flux time series has been estimated based on just an instantaneous and daily mean values in each day. To get more reliable results, the soil heat flux observations should have the same temporal resolution as LST data.

3. In Chapter 6, the validation is done by comparing the calculated soil thermal properties using an empirical model and estimated ones by inverting the physical model output. It would be more reliable if the proper ground measurements of soil thermal properties, LST and soil heat flux would be available in different soil types distributed more across the study area.
8. References

Al Nakshabandi, G., & Kohnke, H. (1965). Thermal conductivity and diffusivity of soils as related to moisture tension and other physical properties. *Agricultural Meteorology, 2, 271-279*


Murray, T., & Verhoef, A. (2007). Moving towards a more mechanistic approach in the determination of soil heat flux from remote measurements. a universal approach to calculate thermal inertia. *Agricultural and Forest Meteorology*, 147, 80-87


Pestiaux, P., Mersch, I., Berger, A., & Duplessy, J.C. (1988). Paleoclimatic variability at frequencies ranging from 1 cycle per 10 000 years to 1 cycle per 1000 years: Evidence for nonlinear behaviour of the climate system. *Climatic Change, 12*, 9-37


NOAA/AVHRR operational satellite data. *Advances in Space Research, 11*, 51-54


NDVI in southern Africa and America continent. (pp. 19-24). Wageningen (The Netherlands): Dlo Winand Staring Center


Appendix A

Factorization of a matrix

The process of decomposition of a matrix into the product of simpler matrices is called factorization. One kind of factorization is Gaussian elimination in which a matrix $A$ will be written as a product $LU$, where $L$ and $U$ are lower and upper triangular matrices, respectively. A triangular matrix is one in which all entries above or below the main diagonal are zero. For an example:

$$L = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 4 \\ 0 & 0 & 5 \end{bmatrix}$$

$L$ is an upper triangular matrix or right triangular matrix.

Any matrix $A$ having nonzero pivot (the main diagonal entries) has a unique $LU$ factorization. If $A = LU$ then all entries of main diagonal of $L$ are equal to 1, and it is possible to factorize $U$ as follows:

$$U = DU = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_{n-1} \end{bmatrix} \begin{bmatrix} 1 & u_{12} / d_1 & u_{13} / d_1 & \cdots \\ 0 & 1 & u_{23} / d_2 & \cdots \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

in which the main diagonal values of $U$ are the same as in $D$. Then the triangular decomposition of $A$ is given by:

$$A = LDU'$$

If a matrix $A$ has some zero value in its pivots, then there might be a possibility to change the row ordering in order to have nonzero pivot matrix and then we can factorize it. When no row reordering of $A$ will produce a nonzero pivot, then an $LDU$ factorization is impossible, and we will call the matrix $A$ ‘singular’. In linear algebra a square matrix $A$ ($n \times n$) is called invertible (non-singular), if there is a matrix $B$ ($n \times n$) such that $AB = BA = I$, $I$ is the identity matrix in which all main diagonal elements are one, and the rest are zero, as an example:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The inverse of a matrix $A$ will be shown as $A^{-1}$, so

$$AA^{-1} = A^{-1}A = I$$
A matrix $A$ is invertible when its determinant is not zero, or none of the pivot values of $A$ is zero. So, if a matrix $A$ is invertible, it is non-singular.

The transpose of a matrix $A$ is a matrix $B$ (denoted as $A^T$) in which the columns and rows of $A$ are exchanged. For example, if

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 3 & -1 & 4 \\ 3 & 2 & 1 \end{bmatrix}$$

then,

$$A^T = \begin{bmatrix} 1 & 3 & 3 \\ 2 & -1 & 2 \\ 2 & 4 & 1 \end{bmatrix}$$

Some properties of a transposed matrix are as follows:

$$(A + B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

$$(A^{-1})^T = (A^T)^{-1}$$

**Symmetric and Hankel matrix**

A matrix $A$ is symmetric if there is a property like this:

$$A = A^T$$

So a symmetric matrix has to be a square matrix, and for all $a_{ij} = a_{ji}$,

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = A^T$$

If $A$ is a symmetric matrix and has no zero value in its pivot, then factorization of it to $LDU$ will be easier, and $U$ will be the transposed of $L$ and then we have:

$$A = LDL^T$$

As an example, see the following decomposition of a symmetric matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

A Hankel matrix is a square matrix in which the positive sloping diagonals elements are constant, and $A_{i,j} = A_{i-1,j+1}$. For instance:

$$A = \begin{bmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ d & e & f & g \end{bmatrix}$$
Determinant of a matrix

SSA can be used to find the eigenvalues of a matrix A. The values of $\lambda$ that make the $A-\lambda I$ singular are the eigenvalues of matrix A. The determinant of a matrix A gives us such a test to find the singularity of the matrix. When the determinant of A (denoted as $|A|$) is zero, then A is singular, and if $|A| \neq 0$ then A is invertible. So in order to find the eigenvalues of a matrix A, the following equation should be solved:

$$|A-\lambda I|=0$$  \hspace{1cm} (13)

The determinant of a matrix $2 \times 2$ is:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$  \hspace{1cm} (14)

Eigenvalue and eigenvector

Let A be a square matrix of order n with real or complex entries, the number $\lambda$ is called an eigenvalue of A if there exists a non-null (nonzero) vector $x$ such that:

$$Ax = \lambda x$$  \hspace{1cm} (15)

The vector $x$ is the eigenvector associated with the eigenvalue $\lambda$ and the set of the eigenvalues of A is called the spectrum of A. In a simpler way, when a vector $x$ multiplied by a matrix A it changes both direction and magnitude, but in a special case where this multiplication only change the magnitude of vector, it is named the eigenvector of A and the associated factor by which the vector $x$ scaled is named eigenvalue of that eigenvector. As an example if:

$$A= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Finding the eigenvalues of A will be as follows:

$$|A-\lambda I|=0$$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0$$

$$(2-\lambda)^2-1=0$$

$\lambda_1 = 1$ and $\lambda_2 = 3$ are the eigenvalues of A, and the associated eigenvectors of A are as follow:

$$(A-\lambda I)x = 0 \text{ or } Ax = \lambda x$$
then,

\[
x_1 = \begin{bmatrix} 3 \\ -3 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}
\]

\(x_1\) and \(x_2\) are the eigenvectors of matrix \(A\).

**Trace**

Trace is the sum of the \(n\) diagonal entries of a matrix \(A\) (\(A\) is \(n \times n\) matrix), the trace of \(A\) is equal to the sum of eigenvalues of \(A\).

\[
\lambda_1 + \lambda_2 + \ldots + \lambda_n = a_{11} + a_{22} + \ldots + a_{nn}
\]

(16)

For a diagonal or triangular matrix \(A\), the eigenvalues of \(A\) are equal to the diagonal elements of \(A\).

**Linear dependency**

A set of vectors \(v_1, v_2, \ldots, v_n\) are linearly dependent if and only if there exists a set of scalars \(c_1, c_2, \ldots, c_n\), not all zero, such that:

\[
c_i v_1 + c_2 v_2 + \ldots + c_n v_n = 0 \quad \forall c_i, \quad i = 1, 2, \ldots, n
\]

(17)

As an example:

\[
v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}
\]

\[
c_1 v_1 + c_2 v_2 = 0
\]

\[
\lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -3 \\ 2 \end{bmatrix} = 0
\]

\[
c_1 - 3c_2 = 0 \quad \text{and} \quad c_1 + 2c_2 = 0 \quad \text{then} \quad c_1 = c_2 = 0
\]

So \(v_1\) and \(v_2\) are linearly independent because \(c_1\) and \(c_2\) are both zero. If the determinant of a matrix is zero, then the vectors containing the columns of that matrix are linearly dependent, and the matrix is not invertible and it is singular.

**Diagonal form of a matrix**

If \(A\) is a \(n \times n\) matrix with \(n\) linearly independent eigenvectors \(e_1, e_2, \ldots, e_n\) then:

\[
E^{-1} AE = \Lambda
\]

(18)
Eq. 4.18 shows a diagonal form of the matrix $A$ where, $E$ is a matrix whose columns are the eigenvectors of $A$ and $\Lambda$ is a diagonal matrix whose nonnegative values of entries are eigenvalues of $A$.

**Spectral decomposition**

If $A$ is a real, symmetric matrix ($A = A^T$), then if all eigenvalues of $A$ are distinct then their corresponding eigenvectors are orthogonal. That means if $x_i$’s are the eigenvectors, then $x_i^T x_j = 0$ for all $i \neq j$. Normalizing these eigenvectors results in a set of orthonormal eigenvectors as follows:

$$e_i = \frac{x_i}{\|x_i\|}$$  \hspace{1cm} (19)

where,

$$\|x_i\| = \sqrt{x_{i1}^2 + x_{i2}^2 + \ldots + x_{in}^2}$$  \hspace{1cm} (20)

So, a real, symmetric matrix $A$ can be diagonalized by an orthonormal matrix $Q$ whose columns are the orthonormal eigenvectors of $A$ as follows:

$$Q^{-1}AQ = \Lambda$$  \hspace{1cm} (21)

or ,

$$A = \lambda_1 e_1e_1^T + \lambda_2 e_2e_2^T + \ldots + \lambda_n e_ne_n^T$$  \hspace{1cm} (22)

This is called the spectral decomposition of matrix $A$.

**Embedding procedure**

In the Appendix A. 8, we describe how to decompose a symmetric matrix into a set of simpler matrices. This idea is the starting point for understanding SSA. Now, the question is how from a single time series record, it is possible to create a symmetric matrix? Before that, we begin with some notation about multivariate statistics.

The main concern in employing SSA is the analysis of individual time series and the method used to change a single time series to a multivariate time series is considering a time lagged copy of the time series as additional time series (Variables) (Ruelle 1980). The procedure of making a univariate time series to a multivariate time series is called an embedding (or, method of delay) and the number of lags is called the embedding dimension.
Trajectory matrix

Let $F = (f_1, f_2, \ldots, f_N)$ be a time series of length $N$, and $M$ be an integer called “window length”. If $K = N - M + 1$, and the M-lagged vectors $X_j = (f_j, \ldots, f_{j+M-1})$, $j = 1, 2, \ldots, K$, then the trajectory matrix $(M \times K)$ is created as:

$$X = [X_1, \ldots, X_K]$$  

(23)

Lagged covariance matrix

The lagged covariance matrix is defined as $S = XX^T$ in which $X$ is the trajectory matrix. The lagged covariance matrix is a symmetric matrix. As $S$ is a symmetric matrix, there is a simplified decomposition of $S$ as follows:

$$S = EE^T \lambda$$  

(24)

where $\lambda$ is a diagonal matrix whose non negative values are eigenvalues of $S$ and square roots of $\lambda$ are called singular values of $X$. $E$ is a matrix whose columns are eigenvectors of $S$ and they are orthogonal and also named singular vectors of $X$. As $S = XX^T$ and because columns of $E$ are orthonormal eigenvectors of $S$, then $EE^T = I$, then one can use the definition of $S$ as follows:

$$SE = E \lambda$$

$$XX^T E = E \lambda$$

$$E^T XX^T E = \lambda$$

$$\left( X^T E \right)^T \left( X^T E \right) = \lambda$$

(25)

As $E$ is composed of orthogonal matrices, the components of matrix $X^T$ that align the basis $E$ after transformation $(X^T E)$ are uncorrelated. It is the way we can transfer the data in new coordinate system and calculate the principal components.

Singular value decomposition

Singular value decomposition (SVD) is a method to decompose a matrix into simpler matrices. It has three main capabilities. First, it can be used as a method for transforming correlated variables into sets of uncorrelated variables which can be interpreted for better representation of various relationships between original data. The second capability of SVD is identifying and ordering most informative dimensions in a data set and finally using those informative dimensions to represent the original data.
with a lower dimensionality. It can be done, for instance, by choosing those dimensions describing more than 90% of the variance in data set. We can then ignore variations below that threshold and reduce massively our data without losing the main relationships of interest.

The mathematical description of SVD is as follows; let $X$ be a rectangular matrix with $m \times n$ dimensions, by using SVD, we can break it up to three matrices (an orthogonal matrix $D$, a diagonal matrix $L$ and the transpose of a matrix $E$) as follows:

$$ X_{m \times n} = D_{m \times m} L_{m \times n} E_{n \times n}^T $$

In which the columns of $D$ are orthonormal eigenvectors of $XX^T$, the columns of $E$ are the orthonormal eigenvectors of $X^TX$ and $L$ is a diagonal matrix containing the square roots of eigenvalues of $XX^T$ in descending order.
Appendix B

The histograms of calculated soil thermal admittance maps at significant periods using the satellite observations of LST and $G_0$. 