# Improving the Snowmelt Simulation of Hydrological Model in Amur River Basin Based on Remote Sensing Data

# Improving the Snowmelt Simulation of Hydrological Model in Amur River Basin Based on Remote Sensing Data

(リモートセンシングデータに基づいたアムール川流域における融雪シミュレーションの水文学モデル改善)

## 2015

**The United School of Agriculture Science, Gifu University**

**Science of Biological Environment**

**(Gifu University)**

**YANG YAN**

# Improving the Snowmelt Simulation of Hydrological Model in Amur River Basin Based on Remote Sensing Data

(リモートセンシングデータに基づいたアムール川流域における融雪シミュレーションの水文学モデル改善)

**YANG YAN**

# Preface

The work outlined in this dissertation was carried out. This dissertation is a part of the fulfillment of the requirement for the degree of Doctor of Philosophy (Ph.D) undertaken by the writer through under the supervisor of Associated Prof. Dr. Takeo Onishi, from the United Graduate School of Agricultural Sciences, Gifu University Japan.

The results included in this dissertation have not been submitted for a degree or diploma or any other qualification at any other university. Furthermore, no part of this dissertation has already been or is currently submitted for any such degree, diploma or other qualification.

# Summary

In most of the middle and high latitude regions, snow accumulation and subsequent snowmelt are considered as the most important hydrological processes, because the stream hydrograph is dominated by spring snowmelt. In addition, nutrient transport from land to sea is significantly influenced by spring flood processes. Hence, knowledge of the spring snowmelt process is essential not only for water resource management, but also for further study of nutrient dynamics and transports. Distributed hydrological models have been proven useful and applicable to investigate stream flow and nutrient transport in snowmelt-dominated basins.

The temperature index method has been widely used despite its simplicity, for the following reasons: (1) wide availability of air temperature data, (2) relative ease of air temperature interpolation and forecasting, and (3) computational simplicity. Because temperature index method is based on an assumption that the relationship between ablation and air temperature is usually expressed in the form of positive temperature sums, thus, the air temperatures are obviously one of the most variables for this method. One widely used hydrologic model is the Soil Water Assessment Tool (SWAT), which was developed to predict the impact of land management practices on water quantity, sediments, and non-point source pollution in large complex watersheds. A simple but improved temperature index model was incorporated into SWAT to simulate snowmelt processes. Although the snowmelt component of SWAT has been successfully applied to different study areas, model applicability is always limited by lower data availability, especially in data-sparse regions.

Spatial interpolation is a common method for increasing data density and creating

accurate air temperature data in data-sparse regions. Furthermore, with development of earth observations, Moderate Resolution Imaging Spectroradiometer (MODIS) remotely sensed land surface temperature (LST) data have proven powerful for creating air temperature data. One advantage of remote sensing data is high spatial and temporal resolution, which provides continuous monitoring and coverage for large spatial scales. With the improvement of sensors and analysis methods, various remote sensing data have been successfully used for air temperature estimation. The linear relationship between LST data and air temperature data has been demonstrated in different study regions.

There are only 87 air temperature stations available in the Amur River basin, the area of Amur River basin is approximate 2,000,000 km$^2$. Moreover, one critical disadvantage of using LST data is that the period of newly created air temperature data is limited by the operational period of the satellite. However, we frequently need historical data, especially for long-term hydrological simulations. Thus, it is necessary to find an easy and effective method to create spatially dense and temporally long-term air temperature data.

Motivated by these unsolved problems, this study set the objectives as follows: 1) to estimate the datasets using different methods and verified them at observation stations. 2) to evaluate the estimated datasets in the test basins for driving the SWAT model to simulate the snowmelt processes. 3) to analyze the factors that influence the accuracy of snowmelt simulation.

Three test basins (Apkoroshi basin, Malinovka basin, Gari basin) were selected for model testing, which are located in the lower, middle and up stream of the Amur River Basin. SWAT model is employed as the test model. In order to generate input data for

SWAT model, digital elevation model (DEM), soil data, land cover (LULC) data, and weather data are prepared. In addition, discharge data are also prepared for calibration of hydrological simulations. In this study, three air temperature datasets are estimated based on different methods for the test basin. These datasets are: 1) one estimated by an inverse distance weighting (IDW) method; 2) one estimated by an improved IDW method considering the elevation influence on temperature (IDWEle); and 3) one estimated by combined use of linear regression and the MODIS LST data. The datasets are verified at observation stations and applied to driving a snowmelt hydrologic model using the SWAT model. The simulation results are compared with observed discharge data and the effects of estimated datasets on the snowmelt simulation results are also discussed.

The different methods were tested at observation stations. The results indicate that the IDW method can obtain better results compared with other two methods according to the low elevation of interpolation stations and estimation stations. Though the linear regression method based on LST data might extend errors for air temperature estimation, this simple linear regression approach can create air temperature data with limited errors range over long periods.

The snowmelt simulation results indicate that the newly estimated air temperature data based on the MODIS LST data can obtain better simulation results than other datasets in all test basins. The performances of IDW and IDWEle method are unstable in different basins with different topographic characters. Analysis of estimated error and simulation errors indicate that the missing consideration of topographic effects in air temperature estimation finally can include errors in the snowmelt simulation.

Totally, the simple linear regression using MODIS LST was generally successful and applicable in our study area. The research showed that using the newly estimated air temperature data based on MODIS LST data to improve the temperature index-based hydrological model (SWAT) is feasible, and the results of this new approach suggest that it could be a powerful means for extending the applicability of the temperature index method to areas with sparse air temperature data.

# Contents

# List of Tables

# List of Figures

# 1 Introduction

Snowmelt is one the most important hydrologic processes in mid- and high-latitude regions (Adam *et al.,* 2009; Edwards *et al.,* 2007). Snowmelt is recognized as one of the main sources for water supply river channels during spring (Brooks *et al.,* 1998). In the Asia, the snowmelt water is extremely important in the Indus basin and important for the Brahmaputra basin, but plays only a modest role for the Ganges River, Yangtze River, and Yellow River (Bai *et al.*, 2011; Immerzeel *et al.,* 2010; Xu *et al.,* 2009). In addition, snowmelt was also proved to have significant influence on nutrient transport (Stewart *et al.,* 2004). For example, Pierson *et al.* (2013) indicated that the annual total dissolved phosphorus load occurs during winter, and nutrient loads during the snowmelt period may account for an average of 46% (18%–73%) of the annual load the New York City water supply region. Corriveau *et al.* (2011) showed that snowmelt plays a key role in nutrient export to prairie aquatic ecosystems and this may have serious impacts on downstream ecosystems in the watersheds of Canadian Prairies.

The Amur River is ranked the 10th by the length area among the world biggest rivers. The total area of this basin is 2,040,700 km$^2$ and its hydrological regime is greatly influence on the water supply of different countries such as Russia, China and Mongolia (Makhinov, 2004). In addition, the Amur River is also considered as an important source for the nutrients (N, P, Fe) to the Okhotsk sea and finally have great influence on the oceanic ecology and fish productivity of the North West Pacific Ocean (Nishioka *et al.,* 2007). During March to May, ice breaking and higher discharge due to snowmelt are distinctive hydrological events in this region (Shibata *et al.*, 2007; Nagao

*et al.,* 2005). The different hydrological regimes of snowmelt period could create different pattern of iron and chloride export to the river compared to that observed during other periods (Shibata *et al*., 2007).

In order to investigate the water resources, nutrient transport, water pollutions and human effects on the water cycles, the hydrological models were developed and applied in different regions (Anord *et al.,* 1998; Bracmort *et al.,* 2006; Liang *et al.,* 1994).The hydrological models were already proved to be useful tools for conducting hydrological and nutrient transport research in this study area (Danilov-Danilyan *et al*., 2014; Onishi *et al*., 2012). Recently, though the physically-based energy balance method has been demonstrated to be accurate and powerful for calculating snowmelt processes (Liston *et al.,* 2006; Marks *et al.,* 1999; Stieglitz *et al.,* 2001), the requirements for accurate and variable input data and complex parameterization still limits the applicability of the physically-based energy balance method (Boone and Etchevers, 2001; Essery *et al.,* 2012). Conversely, the temperature index (henceforth, T−I) method has been widely used despite its simplicity, for the following reasons (Hock, 2003): (1) wide availability of air temperature data, (2) relative ease of air temperature interpolation and forecasting, and (3) computational simplicity. Thus, hydrological models such as the Soil Water Assessment Tool (SWAT, Arnold *et al.,* 1998), Hydrological Simulation Program Fortran (HSPF, Al-Abed *et al.,* 2002), MIKE (Bøggild *et al.,* 1999), and Snowmelt Runoff Model (SRM, Martinec and Rango*,* 1986) have adopted the T−I method to simulate snow accumulation and the snowmelt process. Because T−I method is based on an assumption that the relationship between ablation and air temperature is usually expressed in the form of positive temperature sums (Hock, 2003), thus, the air temperatures are obviously one of the most variables for this method.

Moreover, meteorological data are the most indispensable data sources for extending the applicability of hydrological model, water quality model and ecological models, especially in the data sparse regions (Jasper and Kaufmann 2003; Piper and Stewart 1996). Air temperature is a commonly observed element at meteorological stations. However, according to the data collection of world meteorological organization (WMO, Global Historical Climate Data-Daily data), there are only 87 observation stations available in the Amur River basin. In such data-sparse basins, the density of air temperature observed data cannot meet the requirements of accurate applications of the T-I model. Thus, to create high-density and accurate air temperature data, many data resources and methods have been developed. For example, Saha *et al* (2010) successfully used the global observation weather stations to create large cover NCEP-CFSR data for global applications. Uppla *et al* (2005) also supplied a global cover and long period from 1957 to 2002: the ERA-Interim reanalysis datasets of the European Centre for Medium Range Weather Forecasting. Another way to create gridded climate data covering large area is statistical interpolation of observed data; Yatagai *et al* (2012) successfully used the intense monitoring stations in the Asia to create a continental precipitation datasets for the water resources management application in Asia. These datasets are extending the applicability of hydrological model in the data sparse areas (Vu *et al*., 2012, Yang *et al.,* 2014). Seneviratne *et al.* (2005) once successfully applied the ERA-40 reanalysis data of the European Centre for Medium-Range Weather Forecasts to estimate monthly terrestrial water-storage variations from water-balance computations. Lauri *et al.* (2014) applied TRMM remote sensing data, APHRODITE, NCEP-CFSR and ERA-interim datasets as precipitation sources to drive hydrological model in large basins of Monsoon Asia (Mekong River),

3

which only has sparse surface observation networks of hydrometeorological parameters.

Spatial interpolation is a common method for increasing data density and creating accurate air temperature data in data-sparse regions (Monestiez *et al.*, 2001; DeGaetano and Belcher 2007). A popular spatial interpolation method is inverse distance weighted (IDW) method (Dodson and Marks 1997; Gemmer *et al.,* 2004; Hubbard and You 2005). Perry and Hollis (2005) used the geographical information system capabilities to combine multiple regressions with inverse distance weighted interpolation for analyzing the monthly 5 km x 5 km gridded datasets covering the UK for 36 climatic parameters, including precipitation and air temperature. Shen *et al*. (2001) used the inverse distance weighted method to generate input data for Soil quality models developed for ecodistrict polygons and the polygons of the soil landscapes of Canada to monitor the concentration of soil organic matter. Ninyerola *et al*. (2000) used an interpolation method that combines statistical global analysis with a local interpolation (splines and inverse distance weighting) to generate the monthly mean air temperature map over the Iberian Peninsula. For the spatial interpolations method, the data density can greatly influence the interpolation results (Stahl *et al.,* 2006). In fact, most previous accurate temperature interpolation results were achieved based on high-density, input air temperature data (Gemmer *et al.,* 2004; Hong *et al.,* 2005). For example, Courault & Monestiez (1999) used high-density, air temperature monitoring data (station density is $1/1,250$ km$^2$) and a spatial interpolation method in southern France. Dodson & Marks (1997) used an improved interpolation method and 907 monitoring stations across an 820,000 km$^2$ area (station density is $1/1,000$ km$^2$) in the USA to estimate air temperature in a mountain area. For the Amur River basin, based on the relatively sparse observed data, the applicability of spatial interpolation methods is still need

further validation.

In addition, with development of earth observations, the Moderate Resolution Imaging Spectroradiometer (MODIS) remotely sensed land surface temperature (LST) data have proven powerful for creating air temperature data. For example, Kloog *et al.* (2012) successfully applied the spatial smoothing method to evaluate daily air temperature data using MODIS LST data in Massachusetts, United States. Zhu *et al.* (2013) also used MODIS LST to evaluate daily and sub-daily maximum and minimum air temperature on the northern Tibetan Plateau. Zakšek and Schroedter-Homscheidt (2009) reviewed that there are three different methods commonly applied for estimating the air temperature based on the LST data: 1) the statistical methods 2) the temperature–vegetation index methods (TVX) 3) energy-balance methods. They reported that the statistical methods generally perform well, within the spatial and time frame they were derived, but require large amounts of data to train the algorithms (Benali *et al.*, 2012). The TVX method is based on the assumption that for an infinitely thick canopy, the top-of-canopy temperature is the same as within the canopy (Zhu *et al.*, 2013) and uses the Normalized Difference Vegetation Index (NDVI) as a key input variable. However, the assumption of linear and negative slope between LST and NDVI is not always applicable and is influenced by seasonality, ecosystem type and soil moisture variability (Benali *et al.*, 2012; Vancutsemet *et al.*, 2010), and the period of created data is limited by the periods of both LST and NDVI data. Although the energy-balance methods are physically based, the major disadvantage of this method is the requirement of information that cannot provided by remote sensing (Benali *et al.*, 2012). The linear relationship between MODIS LST and air temperature data has been demonstrated in different study regions (Cresswell *et al.,* 1999, Jones *et al.,* 2004; Mostovoy *et al.*, 2006;

Sun *et al*., 2005).Thus, the linear regression method is a common choice for air temperature data estimation using MODIS LST data. Colombi *et al.* (2007) used the linear regression method and MODIS LST data to generate average daily temperature in Italian alpine areas, and they proved that the result of the linear regression method was superior to that of the spatial interpolation method. Shen and Leptoukh (2011) also used the linear regression relationship between air temperature and MODIS LST data to estimate new daily air temperature data in northern China and central Russia. However, for the air temperature estimation which is based on the remote sensing LST data, one critical problem is the estimated air temperature data are frequently limited by the operation period of the satellite, we frequently need a long-term historical data for driving the snowmelt model.

Generally, the accuracy and reliability of the interpolation data always are need verifications. The direct way for verifying the data is comparing the data with the data from the observation station. For example, Chen *et al.* (2014) evaluated four recent gridded weather data: 55-year Japanese Reanalysis (JRA-55), ERA-Interim data, NCEP-CFSR data, and NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) to clarify their quality in representing the diurnal cycle over East Asia. Bosilovich (2013) also evaluated the performances of gridded precipitation data from MERRA dataset and ERA-Interim dataset in the United States based on the gauged precipitation data. However, considering the spatial distribution of the weather data and its areal effects, the applicability of created weather data for driving the hydrological model is also another way to verify their qualities (Duethmann *et al*., 2013; Fuka *et al*., 2013; Vu *et al*., 2012). For example, Elsner *et al.* (2014) once evaluated differences of four commonly used historical meteorological datasets and their impacts

on stream flow simulations by using the Variable Infiltration Capacity (VIC) model in the Rocky mountain regions of United States. Eum *et al.* (2014) also used the VIC model to evaluate the spatial and temporal differences in precipitation and temperature fields among three high-resolution climate data sets available in Canada, namely, the North American Regional Reanalysis, the Canadian Precipitation Analysis and the thin-plate smoothing splines (ANUSPLIN).

One widely used hydrological model is the Soil Water Assessment Tool (SWAT), which was developed to predict the impact of land management practices on water, sediments, and non-point source pollution in large complex watersheds (Arnold & Fohrer, 2005). A simple but improved temperature index model was incorporated into SWAT to simulate snowmelt processes (Fontaine *et al*., 2002). Although the snowmelt component of SWAT has been successfully applied to different study areas (Zhang *et al.,* 2009; Pradhanang *et al.,* 2011), in Amur River, the applicability of the hydrological model is still limited by the input data. Moreover, the applicability of different air temperature estimated data for driving the temperature index model in data-sparse regions are still not frequently addressed.

Motivated by these unsolved problems, this study set the objectives as follows:

1) to estimate the datasets using the spatial interpolation methods and linear regression method combined with MODIS LST data.

2) to verify the estimated data at observation stations.

3) to evaluate the estimated datasets in the test basins for driving the SWAT model to simulate the snowmelt processes.

4) to analyze the effects of the various air temperature datasets on the air temperature estimation and snowmelt simulations.

# 2 Study area and data

## 2.1. Test Basins

In this study, three basins were selected for model testing. These basins are located in the upper (Gari), middle (Malinovka) and lower (Apkoroshi) stream of the Amur River Basin. Basic geographic characteristics of the basins are shown in Figure 2.1 and Table 2.1. The Amur River is the tenth longest in the world, and is recognized as an important dissolved iron source for the Sea of Okhotsk (Nishioka *et al.,* 2007). There are four distinct phases in the Amur water regime: spring floods, summer low water, summer and autumn floods, and winter low water. The main water source is rainfall, supplying 70–80% of total water and snowmelt during spring floods adds 10–20% (Makhinov, 2004).

In the upper stream (Gari), the annual temperature is -2.4 ºC and the annual precipitation is 494mm, and in the lower stream (Apkoroshi), the annual temperature is -0.1 ºC and the annual precipitation is 641mm. In the middle stream (Malinovka), the mean annual temperature is 1.1 ºC and the annual precipitation is 593mm (Yu *et al.*, 2013).

Table 2.1 Basic geographic information of test basins

| Name | | Apkoroshi | Malinovka | Gari |
|---|---|---|---|---|
| **Area (km²)** | | 4,105 | 5,006 | 3,315 |
| **Highest elevation (m)** | | 2,242 | 1,417 | 776 |
| **Lowest elevation (m)** | | 72 | 82 | 229 |
| **Slope (degree)** | | 18.3 | 12.1 | 2.8 |
| **Land-cover type composition (%)** | Forest | 65 | 90 | 30 |
| | Wetland | 7 | 0 | 25 |
| | Shrub | 7 | 6 | 45 |
| | Farmland | 0 | 4 | 0 |
| | Pasture | 21 | 0 | 0 |

## 2.2. Data for air temperature estimation

### 2.2.1. Observed air temperature data

Daily maximum and minimum air temperature data were obtained from the Global Historical Climatic Network-Daily (GHCN–Daily, Menne *et al.,* 2012) of the National Climate Data Center (NCDC). This data was developed to meet the needs of climate analysis and monitoring studies that require data at a daily time resolution. The dataset contains records from over 75000 stations in 179 countries and territories (Menne *et al.,* 2012). Numerous daily variables are provided, including maximum and minimum temperature, total daily precipitation, snowfall, and snow depth. However, there are only two third of stations applied only the rainfall data (Menne *et al.,* 2012).

There are totally 87 observed air temperature stations available in Amur River basin. The available data period was from early 1950s to 31 December 2010. According to the location of test basins, this study used the searching radius at 200km to found suitable stations for each subbasin, the locations of all selected air temperature stations are shown in Figure 2.1 and Table 2.2. These stations are used for validation of

9

estimated air temperature, and also used for the spatial interpolation to create high intense air temperature data in different test basins. In addition, part of these stations will be directly used as the input data for SWAT model as "sparse-data" input scenario to simulate the snowmelt processes.
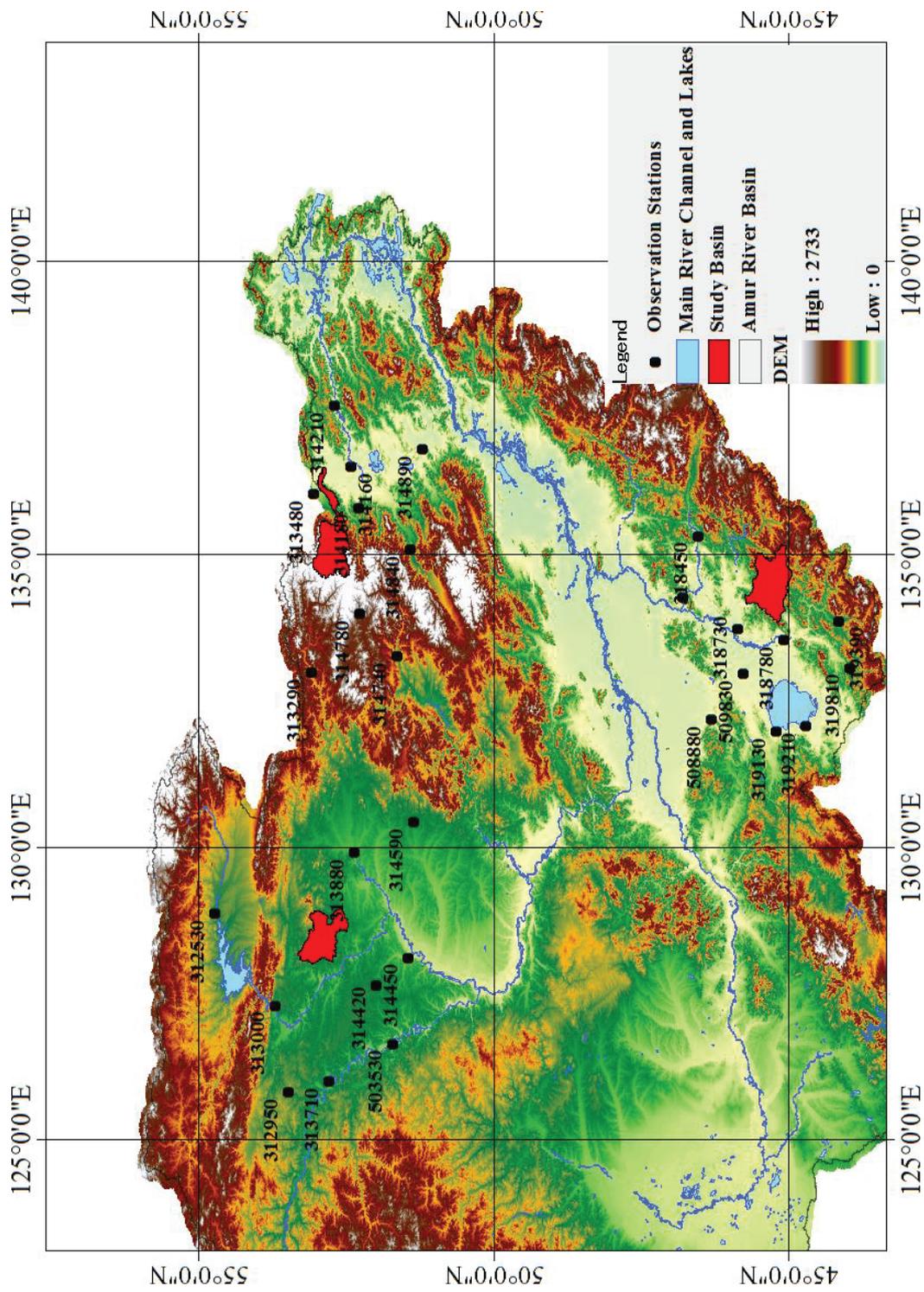
Figure 2.1 The geo-information of the test basin and the observed air temperature stations.

11

Table 2.2 Basic information of the observation air temperature stations

| | Apkroshi | | | | Malinovka | | | | Gari | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| WMOID | Latitude | Longitude | Elevation(m) | WMOID | Latitude | Longitude | Elevation(m) | WMOID | Latitude | Longitude | Elevation(m) |
| 313290 | 53.083 | 132.983 | 542 | 318320 | 46.800 | 134.267 | 68 | 312530 | 54.717 | 128.866 | 364 |
| 313480 | 53.050 | 136.033 | 153 | 318450 | 46.533 | 135.317 | 130 | 312950 | 53.467 | 125.817 | 370 |
| 314160 | 52.417 | 136.500 | 73 | 318730 | 45.867 | 133.733 | 101 | 313000 | 53.700 | 127.300 | 229 |
| 314180 | 52.283 | 135.800 | 201 | 318780 | 45.083 | 133.533 | 98 | 313710 | 52.783 | 126.000 | 210 |
| 314210 | 52.700 | 137.533 | 60 | 319130 | 45.217 | 131.983 | 88 | 313880 | 52.350 | 129.917 | 208 |
| 314740 | 51.633 | 133.267 | 384 | 319210 | 44.717 | 132.067 | 78 | 314420 | 51.983 | 127.650 | 281 |
| 314780 | 52.267 | 133.983 | 902 | 319390 | 44.150 | 133.867 | 259 | 314450 | 51.450 | 128.116 | 197 |
| 314840 | 51.417 | 135.083 | 269 | 319810 | 43.967 | 133.067 | 188 | 314590 | 51.350 | 130.433 | 261 |
| 314890 | 51.200 | 136.800 | 92 | 508880 | 46.317 | 132.183 | 83 | 503530 | 51.717 | 126.650 | 179 |
| | | | | 509830 | 45.767 | 132.967 | 103 | | | | |

### 2.2.2. MODIS land surface temperature data

In this study, the AQUA/MODIS daily LST data at 1-km spatial resolution of both daytime and nighttime are acquired from mid-2002 to 2010. (https://lpdaac.usgs.gov/products/modis_products_table/myd11a1). MODIS LST data are derived from thermal infrared bands 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm). Atmospheric effects are corrected by a generalized split-window algorithm (Wan and Dozier, 1996). The latest LST data are version 005 (MYD11A1, V5); The V5 products are projected in a Sinusoidal grid by mapping the level-2 LST product (MYD11_L2) on 1-kilometer (precisely 1km) grids.

These datasets have error less than 1°C within the range −10 to 50°C, assuming that surface emissivity is known (Wan *et al.,* 2008; Benali *et al*., 2012). In addition, ground-based validation has shown that errors were less than 1°C at homogeneous surfaces such as water, crop, and grassland (Wan *et al.,* 2008).

This study used the ArcGIS Geo-Processing Tool to extract the LST value for the estimated point. The latitude and the longitude of the point are converted into ArcGIS *.shp* file, and the python script language was used to operate the data extraction processing automatically. The LST data was converted from Kelvin degree to Celsius degree, to match the unit of observed air temperature.

## 2.3. Input data of SWAT model

### 2.3.1. Spatial input data

#### 2.3.1.1. DEM data

The SRTM 90m DEM's have a resolution of 90m at the equator, and are provided in mosaiced 5 degree x 5 degree tiles for easy download and use (Jarvis *et al*., 2008). All are produced from a seamless dataset to allow easy mosaicing. These are available in both ArcGIS ASCII and GeoTiff format to facilitate their ease of use in a variety of image processing and GIS applications (Jarvis *et al*., 2008). Digital elevation models (DEM) for the entire globe, covering all of the countries of the world, are available from the website (http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1).

In our study, we downloaded the DEM for the test subbasin and keep the spatial resolution as 90m, the DEM data can cover all the test basins. The elevation of different test basins are shown in Tabel 2.1 and Figure 2.2.The lowest elevation is 72m and highest is 2242m in Apkorshi basin, these values are 82m, 1417m in Malinovka basin, and 242m, 776m in Gari Basin. The slope of Gari basin is 2.8 degree, and in Apkoroshi is 18.3 degree and in Malinovka is 12.1 degree (Table 2.1).  According to the same stretch symbology, it is clear that the Apkoroshi basin has the most significant topographic variations, followed by the Malinovka basin and Gari basin.
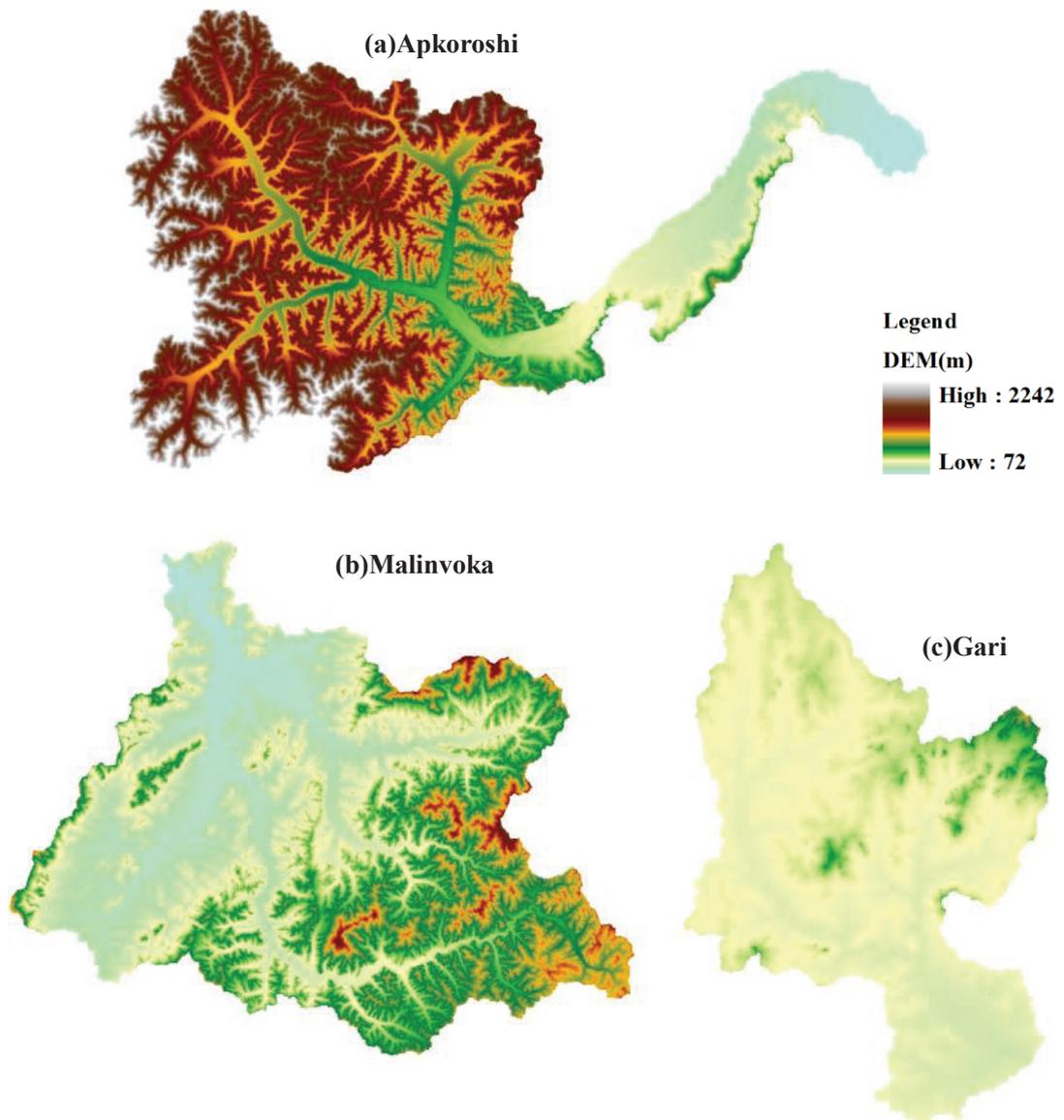
Figure 2.2 The elevations information of all test basins

### 2.3.1.2. Land cover types

The land use/land cover map was constructed by combined use of vegetation maps of China, Mongolia, Russia, and satellite images (Yermoshin *et al.,* 2007), and supplied by the Amu-Okhotsuk Project (http://www.chikyu.ac.jp/AMORE/). This data also supplied 19 classified land cover types. In order to match the application of SWAT model, we combined these data to forest, pasture, range-brush, wetland, and agricultural

land for model constructions. The spatial resolution of the land cover data is 1km, the land cover types of different test basins are shown in Table 2.1 and Figure 2.3. The major land uses in the Gari basin are range brush (45%), forests (30%) and wetlands (25%) in the Apkoroshi basin are forests (65%), pasture (21%), shrub (7%) and wetlands (7%). In the Malinovka basin are forests (90%), shrub (6%) and farmlands (4%).

### 2.3.1.3. Soil data

The soil input data is an essential input data for SWAT model, which will greatly influence its calculations of erosion, ET, water storage (Anorld *et al*., 1998). To define the hydrologic soil water effects requires estimating soil water characteristics for water potential and hydraulic conductivity using soil variables such as texture, organic matter (OM), and structure (Saxton and Rawls, 2006).

The Food and Agriculture Organization of the United Nations (FAO) and the International Institute for Applied Systems Analysis (IIASA) took the initiative of combining the recently collected vast volumes of regional and national updates of soil information with the information already contained within the 1:5,000,000 scale FAO-UNESCO Digital Soil Map of the World, into a new comprehensive Harmonized World Soil Database (HWSD, Nachtergaele *et al.,* 2012). The soil data in our study are obtained from the International Institute for Applied Systems Analysis (IIASA, http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/). The spatial resolution of the soil data is 1km.
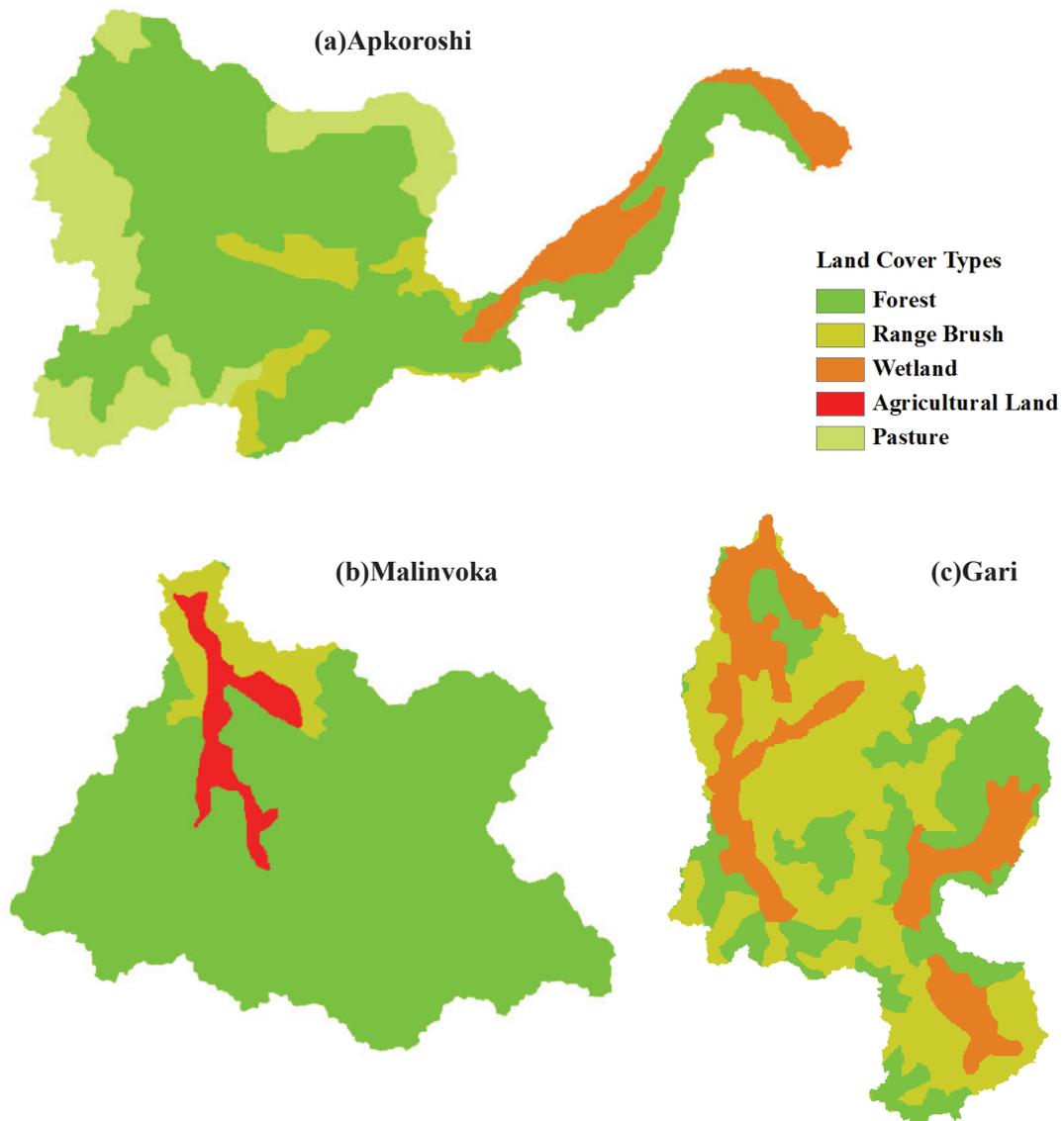
Figure 2.3 The land cover types of the test basins.

Field or laboratory measurements are difficult, costly, and often impractical for many hydrologic analyses (Saxton and Rawls, 2006). Saxton and Rawls (2006) proved that statistical correlations between soil texture, soil water potential, and hydraulic conductivity can provide estimates sufficiently accurate for many analyses and decisions. They developed new soil water characteristic equations from the currently available USDA soil database using (Soil-Plant-Agriculture-Water, SPAW) only the readily available variables of soil texture and OM (Saxton and Rawls, 2006). This

software formed a comprehensive predictive system of soil water characteristics for agricultural water management and hydrologic analyses (Saxton and Rawls, 2006). It is available at the website of Agricultural Research Service, United States Department of Agriculture (http://hydrolab.arsusda.gov/soilwater/Index.htm).

In this study, the physical characters of different soil data are calculated by using the SPAW software and changed it into the database for SWAT model.

### 2.3.2. Weather input data

#### 2.3.2.1. Precipitation data

Precipitation is one of the most basic meteorological elements and it directly and indirectly affects human life (Yatagai *et al.,* 2012). Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE; Yatagai *et al.,* 2012) was used for daily precipitation data in this research.

APHRODITE project created continuous daily gridded precipitation data for 1951–2007. These data cover monsoon Asia, the Middle East, and Russia (Yatagai *et al.,* 2012). The calculation framework of APHRODITE data is similar to that of Xie *et al*. (2007), the fields of daily climatology are then adjusted by the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) monthly precipitation climatology to correct the bias caused by orographic effects (Yatagai *et al.,* 2012). In addition, the most important driving data of distributed hydrological models are of accurate precipitation. It has been shown that APHRODITE can give good performance in the Amur River Basin (Gillies *et al.,* 2012; Onishi *et al.,* 2012).

The Version V1101R1 data of monsoon Asia of APHRODITE data (http://www.chikyu.ac.jp/precip/) was used in this research, with spatial resolution of

0.25 degrees (Yatagai *et al.,* 2012). The available period was 1 January 1979 to 31 December 1989.

### 2.3.2.2. Wind speed data

The wind speed data are also obtained from the GHCN–Daily (Menne *et al.,* 2012, https://www.ncdc.noaa.gov/oa/climate/ghcn-daily/) of the National Climate Data Center (NCDC). The available period was 1 January 1979 to 31 December 1989.

### 2.3.2.3. Relative humidity data and solar radiation data

National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (R1) was created by a complex system of programs, libraries, scripts, and datasets involving many steps including decoding, reformatting, quality control, analysis, prediction, post processing, and archiving. NCEP-DOE Reanalysis II (R2) is an improved version of the NCEP-NCAR-R1 (Kanamitsu *et al.,* 2002). The improvements include an updated model with better physical parameterizations, assorted data assimilation errors were fixed and additional data were included (Kanamitsu *et al.,* 2002). The goal of the NCEP-DOE R2 data is to generate new atmospheric analyses historical data.

In our study, relative humidity and solar radiation data were from the NCEP–DOE Reanalysis 2 dataset (Kanamitsu *et al.,* 2002, http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html) on the website of the NOAA Earth System Research Laboratory. The period was 1 January 1979 to 31 December 1989.

**2.3.2.4.   Observed snowmelt runoff data**

Runoff data were from the Russian Federal Service for Hydrometeorology and Environmental Monitoring (Roshydromet). The locations of the observed runoff stations are shown in Figure 2.3 for all three test basins. For each basin, there is one observed station located in the outlet of the test basin. The observed data from 1983 to 1987 were selected for basins Apkorohsi and Malinovka, and because lack of the observed data, the data of 1983, 1984, 1986 were used for Gari basin.

## 2.4.  Subbasin delineation of the test basins

All the spatial data are converted into the same projected coordinate system (*Albers-Fareast*, Yermoshin *et al.,* 2007). Based on the input spatial data and attribute data, this study conducted the subbasin delineation for each test basin, based on the ArcSWAT (Douglas-Mankin *et al*., 2010). The results are shown in Figure 2.3. In the Apkoroshi basin, there are 13 subbasins. In Malrinvok basin, there are 13 subbasins, and in the Gari basin there are 11 subbasins. According to the automatic selection processes of SWAT model, the nearest observation stations of the each subbasin is selected as the only input station for weather data.
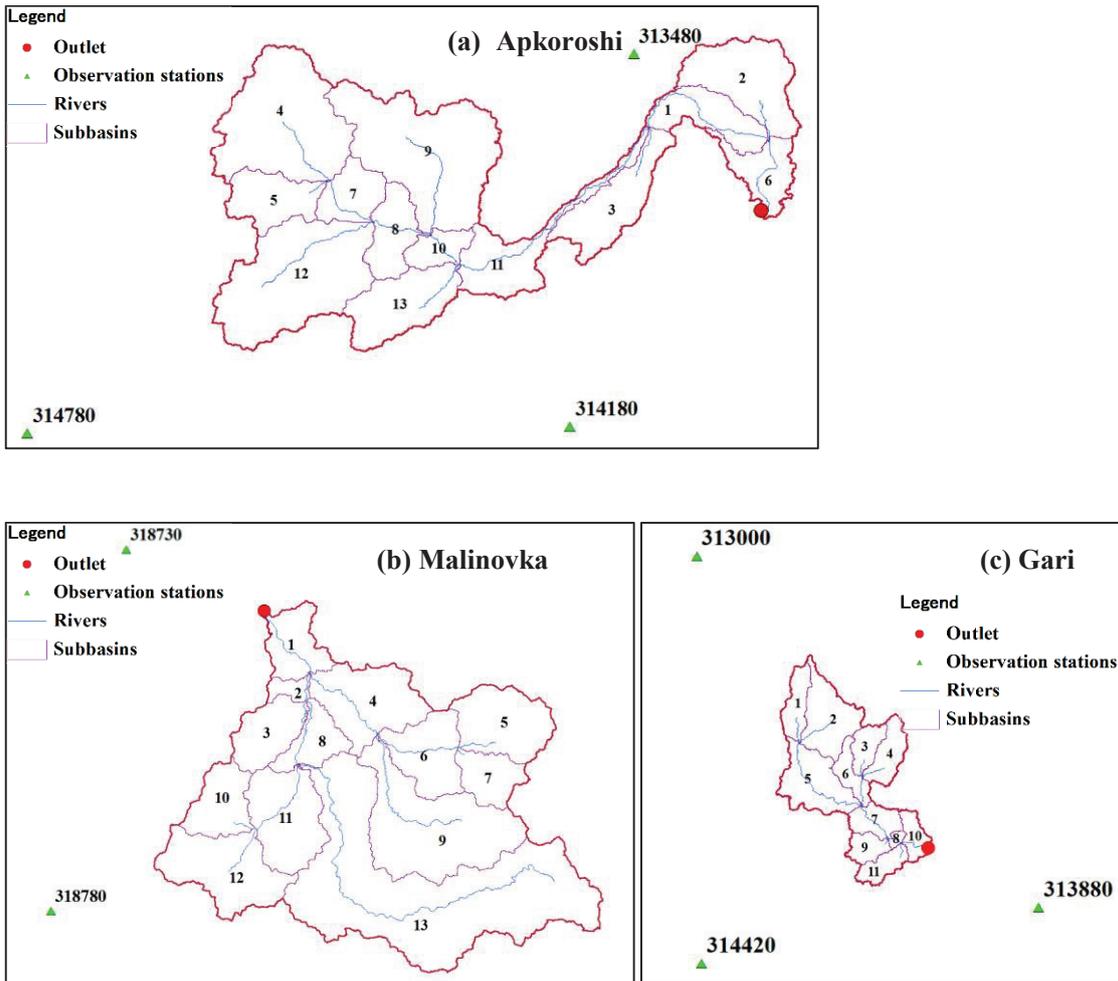
Figure 2.4 The subbasin delineation results of the test basins, (a) Apkoroshi basin, (b) Malinovka basin, (c) Gari basin.

## 2.5. Snowmelt period of the test basin

For discussion the snowmelt simulation, the snowmelt period is need to be confirmed firstly. According to the previous study of, the Amur River the peak flow during the spring always occurred in the May which is induced by the snowmelt, and the mixed rainfall and we snow storms (Makhinov, 2004). For a specific basin the period of snowmelt is influenced by various factors and it also can be changed year by year (Hock, 2003).

In this study we using both the hydrograph of the observation runoff stations and

21

the observed air temperature data that to identify the snowmelt period.

This study used the river runoff data from 1983 to 1987 (March, April, May and June) to find the lowest value of the runoff as the starting point for snowmelt. In addition we also employed the daily maximum air temperature data as another evaluation index to confirm the selection of the starting point, usually the snowmelt occurred when the daily maximum air temperature is over 0.0°C. And then we used the daily average temperature data to confirm the event of precipitation is rainfall or snow fall during the spring flood period, the threshold is 5.0°C (Fontaine *et al*., 2002).If the daily air temperature of basin is always larger than 5.0°C from a specific date, then this data will be selected as the end date of the simulation.
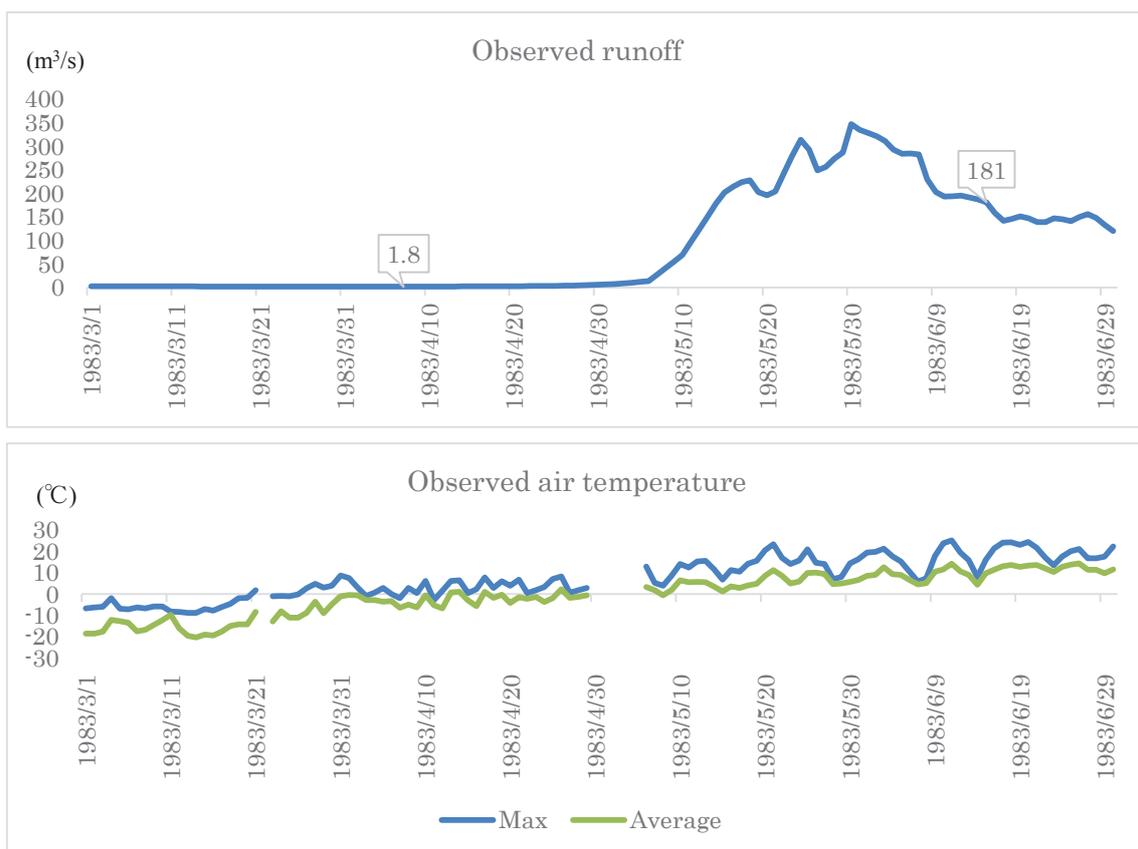


Figure 2.5 The hydrograph and air temperature data from March to June of year 1983 in the Apkoroshi basin.

Here, this study takes the runoff and air temperature data in 1983 of Apkoroshi

basin as an example, the daily air temperature data and the hydrograph of year 1983 from 1983.03.01 to 1983.06.30 is shown in Figure 2.5. The lowest runoff value is observed in 1983.04.07 (1.8 m$^3$/s). The runoff data and daily maximum air temperature data of sub-period from 1983.04.01 to 1983.04.20 are shown in Figure 2.6. It is clear that start from the melt point the daily maximum air temperature data is over 0.0°C.
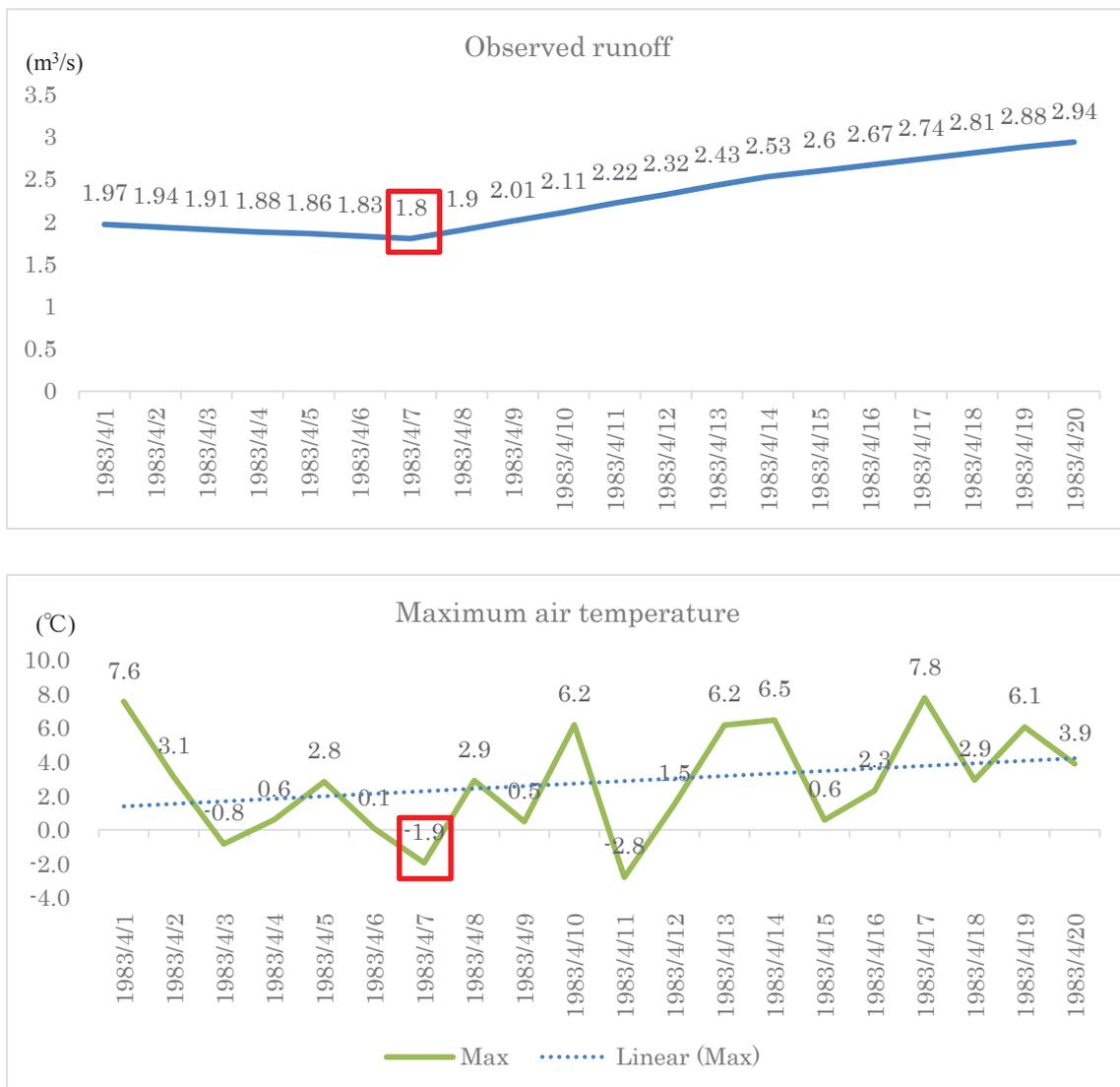


Figure 2.6 The observed runoff and daily maximum air temperature in the Apkoroshi basin of year 1983 during the starting period of snowmelt.

The end point of the snowmelt is confirmed by using the daily average temperature, the data 1987.06.15 (daily average air temperature is 9.9°C) is confirmed as the end

point, because from this date the air temperature will always over 5.0°C until the starting of autumn.
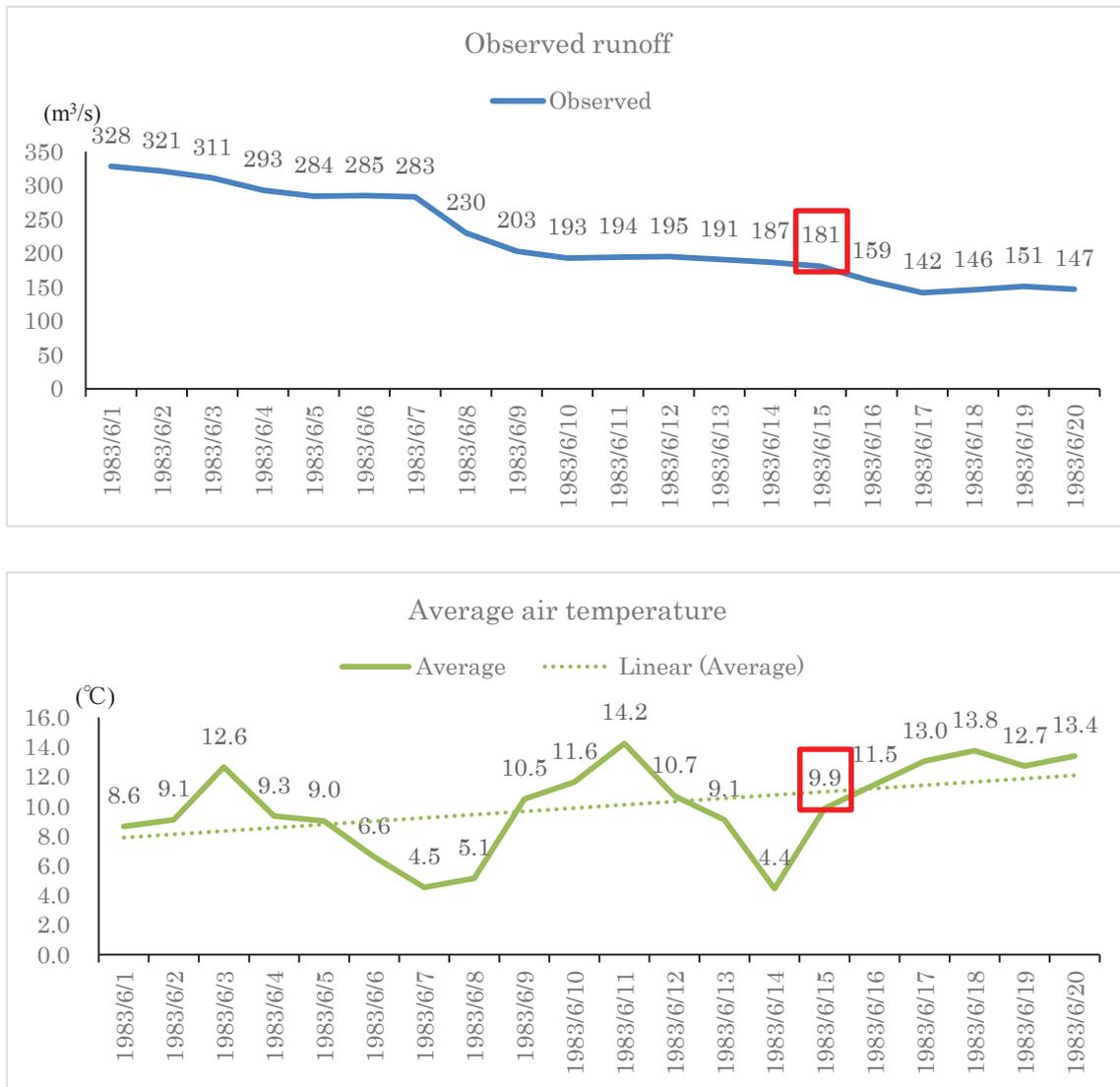


Figure 2.7 The observed runoff and daily average air temperature in the Apkoroshi basin of year 1983 at the end of snow melting period.

By using the same method, this study confirmed the calibration period for snowmelt in all test basins, and the results are listed in the Table 2.3. Moreover, based on the calculation of the slope of each observed point, we selected the point that divide the hydrograph into two periods, the early melting periods and the main melting periods.
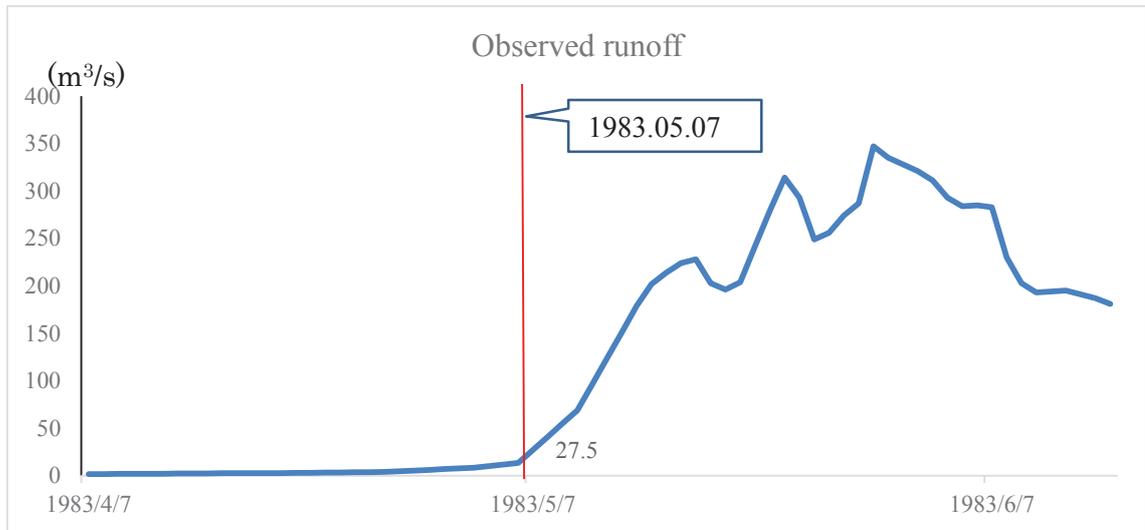
Figure 2.8 The separation of hydrograph for early melting period and main melting period in Apkoroshi basin of year 1983.

During 1983 to 1987, there are 253 observed data for Apkoroshi basin, 185 observed data for Malinovka basin, and 85 observed data for Gari basin.

Table 2.3 The snowmelt simulation period of each test basin.

|  | Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|---|
|  | Start | End | Separation* | Start | End | Separation | Start | End | Separation |
| 1983 | 04.07 | 06.15 | 05.07 | 03.23 | 04.21 | 04.04 | 04.10 | 05.09 | 04.23 |
| 1984 | 04.10 | 05.11 | 04.21 | 03.30 | 04.24 | 04.13 | 04.10 | 05.16 | 04.21 |
| 1985 | 04.05 | 05.29 | 05.27 | 03.21 | 04.13 | 04.07 | No data | No data | No data |
| 1986 | 04.10 | 05.30 | 04.27 | 04.01 | 04.30 | 04.10 | 04.17 | 05.04 | 0.4.27 |
| 1987 | 04.24 | 06.07 | 05.02 | 03.23 | 05.14 | 04.16 | No data | No data | No data |

*Separation is the date that split the hydrograph into the early melting period and the main melting period.

# 3 Methodology

## 3.1. Long-term Air temperature data estimation methods

### 3.1.1. Spatial interpolation methods

The topographic effects were frequently considered when apply the spatial interpolations methods to estimate the air temperature data (Stahl *et al*., 2006).One of the ways to correct the topographic influence on IDW method is to calculate correction term based on the regression between elevation and air temperature (Thornton *et al.,* 1997). Hereafter, we call the improved IDW method as IDWEle. We used both IDW and the improved IDW methods to create air temperature data. In the original IDW, the weight is calculated as

$$w_i = \frac{1}{\sqrt[2]{(x_i - x_p)^2 + (y_i - y_p)^2}} \tag{1}$$

Where, $w_i$ is the weight ($m^{-1}$), $x_i$ and $y_i$ represent the location of monitoring points (m), and $x_p$ and $y_p$ locate the estimation location (m). To use the IDW method, a search radius is fixed to select stations used for estimation, in this study, the searching radius is 200km, using the weight, the estimated value $T_{p,d}$ is calculated as

$$T_{p,d} = \frac{\sum_{i=1}^{n} w_i \times T_{i,d}}{\sum_{i=1}^{n} w_i} \tag{2}$$

Where $T_{i,d}$ symbolizes daily observed air temperature data and *d* is a given day.

In the IDWEle method, to consider elevation effects on the original weight, an

additional regression analysis is performed using air temperature and elevation values of interpolated stations (Thornton *et al.,* 1997). The equation is

$$T_{1,d} - T_{2,d} = b_{0,d} + b_{1,d}(h_1 - h_2)$$
(3)

Where, subscripts 1 and 2 refer to the two stations in a unique pair, $T_{,d}$ (°C) is air temperature data measured in a given day of a given station on day ($d$), and $b_{0,d}$ and $b_{1,d}$ are coefficients of the linear regression equation on the same day. Using the daily observation data and Eq. (3), the final temperature is calculated as

$$T_{p,d} = \frac{\sum_{i=1}^{n} w_i(T_{i,d} + b_{0,d} + b_{1,d}(h_p - h_i))}{\sum_{i=1}^{n} w_i}$$
(4)

Where, subscript $p$ refers to the prediction points and $i$ to the interpolated stations.

### 3.1.2.  Linear regression method combined using the MODIS LST data

The simplest method for estimate the air temperature is to create a linear regression equation between air temperature at points A and B (Figure 3.1) using observed air temperature data. In this case, the linear regression equation can be written as follow:

$$T_{a,B} = a_1 \times T_{a,A} + b_1$$
(5)

Here, $a_1$ and $b_1$ are coefficients of the linear regression equation, and subscripts $A$ and $B$ indicate the points.

We call this the $T_a-T_a$ method (Yang *et al*., 2014; Yang *et al*., 2015). If monitoring data at both point A and B are available, we can use the method. However, if we have no data at a point we need to know, the method is not applicable. Here, we use the observed air temperature stations (Table 2.2) and their nearest stations to evaluate the linear relationships between two different places (Figure 3.1, process ①).
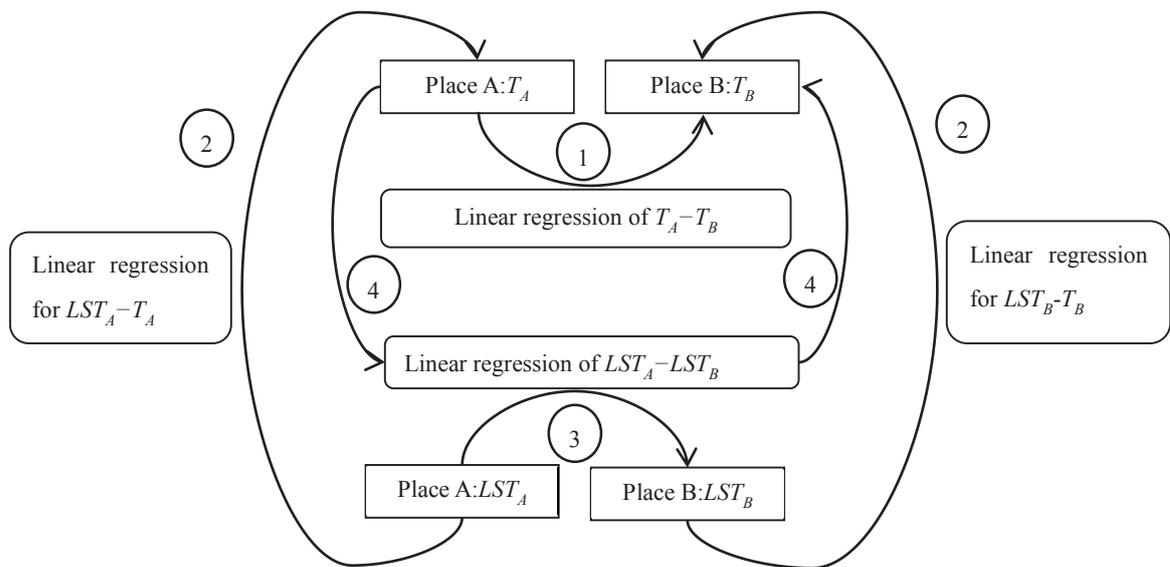


Figure 3.1 Diagram of air temperature estimation methods based on observed air temperature and LST data at observed station

processes ① stand for the linear regression between air temperature at two points, process②stand for the linear regression between LST at each point, process③stand for the linear regression between LST at two points, process④stand for the estimation of linear regression for point B based on the regression equation of LST at two points.
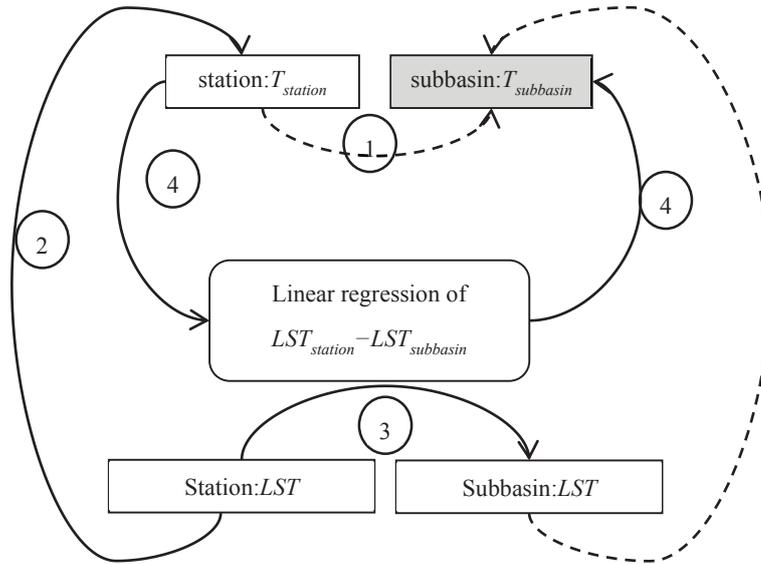
Figure 3.2 Diagram of air temperature estimation methods based on observed air temperature and LST data at subbasins

processes ①, ②, ③, ④ have the same meaning as in Figure 3.1.

Further, the study of Sun *et al*. (2005) presented a theoretical derivation of linear regression relationships between air temperature and LST and proved that the air temperature can be mainly explained by the LST in the linear regression equation; and they also showed the errors of created air temperature based on the linear regression method is limited in a reasonable range, in the North China Plain. In addition, Mostovoy *et al.* (2006) also proposed the similar method to estimate air temperature at any point from LST at that point. By constructing linear regression equations between observed air temperature and LST at 161 observed station points in the state of Mississippi, they also found a common relationship between LST and air temperature, irrespective of location. Furthermore, both of these researches indicate that the first order coefficient of linear equation between air temperature and LST is equal to 1. Here, we use the observed air temperature data and LST data to estimate the air temperature between them; we call this the LST−$T_a$ method (Figure 3.1, process ②). In this case, the linear regression equation can be written as:

$$T_a = LST + const \qquad (6)$$

More, a disadvantage of this method is that because LST data are necessary for this method, it can only be applied to periods after MODIS was launched. As already addressed, we frequently need historical air temperature data to execute hydrological models that include snow accumulation and snowmelt processes. Many watersheds have very sparse observed air temperature data. Because both the $T_a$–$T_a$ and LST–$T_a$ methods are unsuitable for such common cases, we developed a new method as follows.

In the first step, a linear regression equation of LST between two points is created as follows:

$$LST_B' = a_2 \times LST_A + b_2 \qquad (7)$$

Here $a_2$ and $b_2$ are coefficients of the linear regression equation, and subscripts $A$ and $B$ indicate the points (Figure 3.1, process ③). $LST_B'$ is the result of the predicted LST value of point B from the Eq. (7).

In addition, based on the linear analysis results of $T_a$-LST of station pairs, using the $T_a$–LST method, we can estimate $T_{a,A}$ and $T_{a,B}$ at the same time:

$$T_{a,A} \cong LST_A + const_A \qquad (8)$$

$$T_{a,B} \cong LST_B + const_B \qquad (9)$$

30

Here we neglect the error between the $LST_B{}'$ and $LST_B$ and by substituting $LST_B$ in Eq. (9) by $LST_B{}'$, $T_{a,B}$ can be expressed by $LST_A$ as

$$T_{a,B} = a_2 \times LST_A + b_2 + const_B \tag{10}$$

Then, combining with Eq. (8), we can attain

$$T_{a,B} = a_2 \times T_{a,A} + b_2 + const_B - a_2 \times const_A \tag{11}$$

If we can ignore the item $const_B - a_2 \times const_A$ of Eq. (11), and then the new equation is:

$$T_{a,B} = a_2 \times T_{a,A} + b_2 \tag{12}$$

This means that once we acquire coefficients $a_2$ and $b_2$ from linear regression analysis of LST, we can estimate $T_{a,B}$ using a known $T_{a,A}$ (Figure 3.1, process ④). We call this the LST−LST method. We performed a linear regression analysis for creation of $T_a$ based on both daily maximum and minimum LST data.

However, according to the limited observed period of MODIS LST data, the linear regression equations are formulated only in the entire period.

Because in the study basins, there is no observed data in the subbaisn $T_a$-$T_a$ relationship cannot be acquired (Figure 3.2, process ①). Moreover, according to lack of observed data, the $T_a$-LST relationship also can only be acquired in the point which monitors both LST and air temperature data (Figure 3.2, process ②), thus, it means we can't obtain $const_B$ because missing of observed air temperature in the Place B. The

only feasible way is to use the LST-LST method (Figure 3.2, process ③) and the data of the observed station to estimate the air temperature in the subbasin (Figure 3.2, process ④)

It is noteworthy that the linear regression relationship should be firstly valid in the observed stations and it is also that the item $const_B - a_2 \times const_A$ in Eq(11) will greatly influence the accurate of estimate method. Thus, this study used the 28 stations listed in Table 2.1 as the validation stations, and used their nearest stations to conduct the linear relationships analysis based on $T_a$-$T_a$ , $T_a$-LST, and LST-LST methods. In addition, this study also took the $const_B - a_2 \times const_A$ as an entirety, by using the regression coefficients of the different method to valid whether this item can be ignored. The results of the linear regression analysis between air temperature and LST are shown in section 4.1.

### 3.1.3. Evaluation of estimated data at observation stations

In order to compare the accurate of different air temperature estimated method, this research used 28 stations (Table 2.1) for validations to test the performances of IDW, IDWEle and LST-LST methods for estimating air temperature data at these stations. The daily maximum and minimum air temperature are estimated and compared with those values of the validation stations. Two indices, coefficient of determination ($R^2$) and root mean square error (*RMSE*) were used to evaluate the estimation results. Equations for these indices are

$$R^2 = \left\{ \frac{\sum_{i=1}^{n}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\left[\sum_{i=1}^{n}(y_i-\bar{y})^2\right]\ \left[\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2\right]} \right\} \tag{13}$$

$$RMSE = \left(\frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|^2}{n}\right)^{\frac{1}{2}} \tag{14}$$

Where $y_i$ is observed air temperature on day i, $\hat{y}_i$ is the estimated air temperature on that day from each method, $\bar{y}$ is the average observed air temperature, $\bar{\hat{y}}$ is the average of estimated air temperature, and n is the total number of data.

## 3.2. Estimating dense air temperature data for SWAT model

The SWAT model selects one air temperature station one subbasin based on the distance between the geometrical center of subbasin and all candidate stations. The data from the nearest station are chosen for input.

Based on the observed data, and the IDW, IDWEle, LST-Linear methods, we created four different input data for SWAT model. The estimated location of all methods is in the geometrical center of each subbasin, the basic geo-information of each estimated location (geometrical center of subbasin) is listed in Table 3.1 and Figure 3.3. Totally, there are 11 estimated points for Apkoroshi basin, 13 estimated points for Malinovka basin and 13 estimated points for Gari basin. The number of estimated points is as same as that of subbasins in each test basin. The four datasets are described as follows:

(1) T-Observed scenario, the data based on the observed air temperature data. In view of SWAT model automatically assigns the value of nearest station to the subbasin, this study use this data as a default input.

(2) T-IDW scenario, the data estimated from the IDW method, and the interpolation data is based on the candidate stations listed in Table 2.1.

(3) T-IDWEle scenario, the data estimated from the IDWEle method, the candidate stations are as same as the IDW method.

(4) T-LST scenario, the data based on the LST-LST method using MODIS AQUA LST data. The daily maximum and minimum LST data were extracted for both estimated point and their nearest observed stations; the period is from year 2002 to 2010.

Table 3.1 Basic information of estimated point in each test basin

| | Apkroshi | | | | Malinovka | | | | Gari | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Subbasin | Latitude | Longitude | Elevation(m) | Subbasin | Latitude | Longitude | Elevation(m) | Subbasin | Latitude | Longitude | Elevation(m) |
| 1 | 52.92 | 136.28 | 109 | 1 | 45.67 | 134.24 | 94 | 1 | 53.12 | 128.11 | 325 |
| 2 | 52.74 | 135.95 | 232 | 2 | 45.57 | 134.23 | 146 | 2 | 53.06 | 128.37 | 324 |
| 3 | 52.91 | 134.84 | 814 | 3 | 45.47 | 134.16 | 167 | 3 | 52.98 | 128.64 | 312 |
| 4 | 52.76 | 134.80 | 1491 | 4 | 45.54 | 134.46 | 172 | 4 | 52.91 | 128.77 | 305 |
| 5 | 52.76 | 135.07 | 1227 | 5 | 45.50 | 134.86 | 507 | 5 | 52.81 | 128.32 | 324 |
| 6 | 52.68 | 135.21 | 718 | 6 | 45.43 | 134.64 | 194 | 6 | 52.87 | 128.53 | 341 |
| 7 | 52.85 | 135.31 | 890 | 7 | 45.38 | 134.82 | 647 | 7 | 52.66 | 128.74 | 260 |
| 8 | 52.65 | 135.36 | 452 | 8 | 45.46 | 134.30 | 145 | 8 | 52.59 | 128.90 | 245 |
| 9 | 52.65 | 135.64 | 346 | 9 | 45.28 | 134.66 | 354 | 9 | 52.55 | 128.71 | 272 |
| 10 | 52.59 | 134.88 | 1412 | 10 | 45.32 | 134.03 | 194 | | | | |
| 11 | 52.54 | 135.24 | 730 | 11 | 45.29 | 134.22 | 146 | | | | |
| | | | | 12 | 45.18 | 134.04 | 221 | | | | |
| | | | | 13 | 45.15 | 134.64 | 597 | | | | |
| average | | | 766 | | | | 276 | | | | 301 |
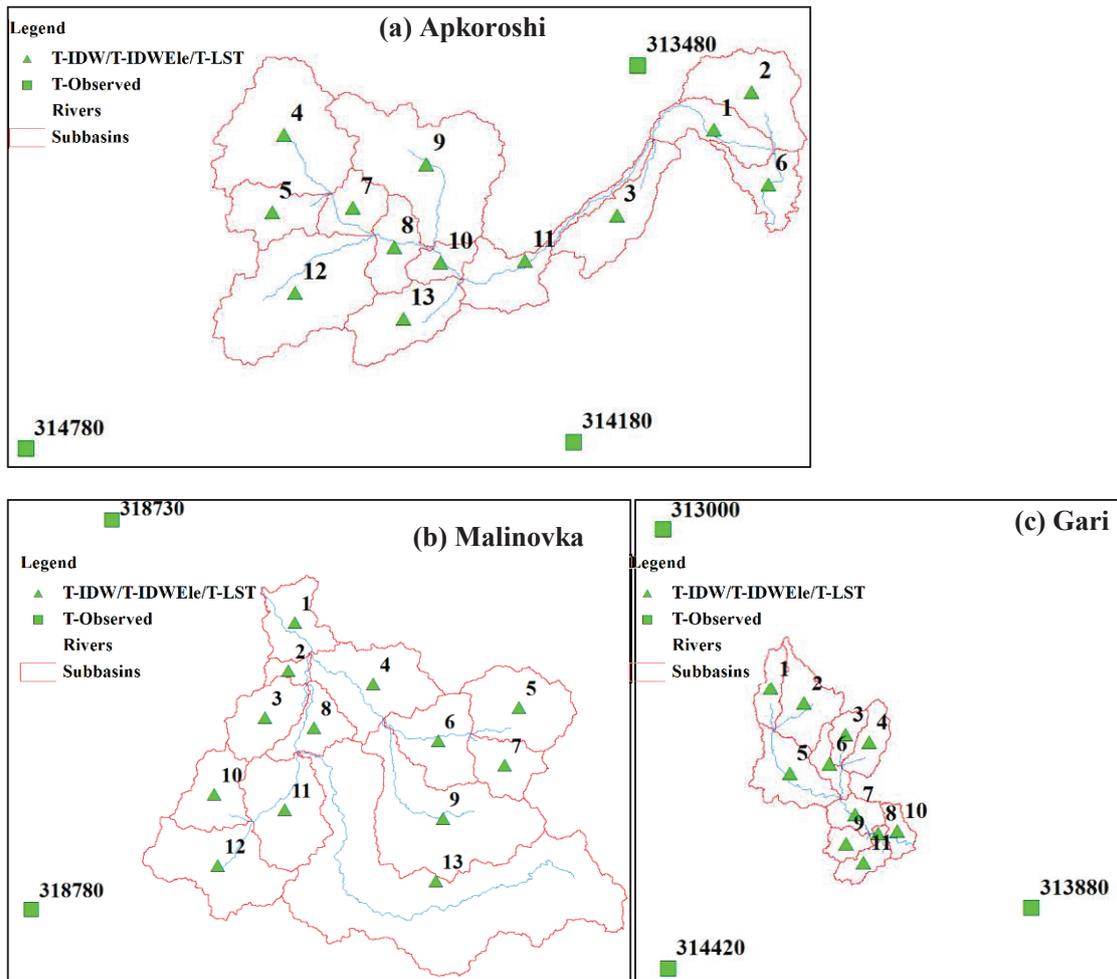
Figure 3.3 The locations of different created air temperature for SWAT model, (a) Apkoroshi basin, (b) Malinovka Basin, (c) Gari Basin.

## 3.3. Introductions of SWAT model

### 3.3.1. Basic theories of SWAT model

SWAT is, a river basin, or watershed, scale model developed by Dr. Jeff Arnold for the USDA Agricultural Research Service (ARS) (Neitsh *et al.*, 2009). SWAT was developed to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time (Neitsh *et al.*, 2009; Zhang *et al.*,

36

2009). The SWAT model is a semi-distributed model, and the construction of SWAT model is based on relatively available input datasets (Anorld *et al*., 1998). Previous studies already proved that the computationally spent of SWAT model is relative efficient (Anorld *et al*., 1998; Yang *et al*., 2008).

For modeling purposes, the SWAT model delineated the test basin into a number of subbasins. The use of subbasins in a simulation is particularly beneficial when different areas of the watershed are dominated by land uses or soils dissimilar enough in properties to impact hydrology (Neitsh *et al*., 2009). Hydrologic response units (HRUs) are further divided based on the unique land cover, soil, and management combinations in one subbasin (Anord *et al.,* 1998; Neitsh *et al*., 2009). The HRU is the smallest and basic simulation unit of the SWAT model.

Simulation of the hydrology in SWAT model includes two major components. The first component is the land phase of the hydrologic cycle, and the second is routing phase of the hydrologic cycle (Anorld *et al*., 1998; Neitsh *et al*., 2009).

The hydrologic cycle as simulated by SWAT is based on the water balance equation (Neitsh *et al*., 2009):

$$SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{gw}) \tag{20}$$

Where $SW_t$ is the final soil water content (mm), $SW_0$ is the initial soil water content on day $i$ (mm), $t$ is the time (days), $R_{day}$ is the amount of precipitation on day $i$ [mm], $Q_{surf}$ is the amount of surface runoff on day $i$ (mm), $E_a$ is the amount of evapotranspiration on day $I$ (mm), $w_{seep}$ is the amount of water entering the vadose zone from the soil profile on day $i$ (mm), and $Q_{gw}$ is the amount of return flow on day $i$

(mm).

The modified SCS curve number method is used to simulate the river runoff in the SWAT model (Anorld and Fohrer, 2005), the Evapotranspiration (ET) is calculated by using the Penman-Monteith method (Douglas-Mankin *et al.*, 2010). In addition, the river channel routing of SWAT model is calculated by using a variable storage method (Anorld and Fohrer, 2005).

### 3.3.2. Snowmelt simulation in the SWAT Model

In the SWAT model, the snow coverage is allowed to decline non-linearly based on an aerial depletion curve (Fontaine *et al.*, 2002). The depletion curve function is:

$$\text{SNO}_{COV} = \frac{\text{SNO}}{\text{SNO}_{100}} \times \left[ \frac{\text{SNO}}{\text{SNO}_{100}} + \exp\left( \text{cov}_1 - \text{cov}_2 \times \frac{\text{SNO}}{\text{SNO}_{100}} \right) \right] \tag{21}$$

Where $SNO_{COV}$ is the fraction of the hydrological response unit (HRU) area covered by snow, $SNO$ is the water content of the snow pack on a given day (mm), $SNO_{100}$ is the threshold depth of snow at 100% coverage (*mm*), $cov_1$ and $cov_2$ are coefficients that define the shape of the curve.

The snow pack temperature is a function of the mean daily temperature during the preceding days and varies as a dampened function of air temperature.

$$T_{snow(d_n)} = T_{snow(d_n-1)} \times TIMP + \left( \frac{T_{maximum} + T_{minimum}}{2} \right) \times (1 - TIMP) \tag{22}$$

Where snow $T_{snow(d_n)}$ is the snow pack temperature on a given day (℃), snow

$T_{snow(d_n-1)}$ is the snow pack temperature on the previous day (°C), $TIMP$ is the snow temperature lag, and $T_{maximum}$ is the maximum air temperature on a given day (°C), $T_{minimum}$ is the minimum air temperature on a given day (°C).

The snow melt simulation of the SWAT model can be described as:

$$SNO_{mlt} = b_{mlt} \times SNO_{cov} \times \left( \frac{T_{snow}+T_{maximum}}{2} - T_{mlt} \right) \tag{23}$$

Where $SNO_{mlt}$ is the amount of snow melt on a given day (mm), $b_{mlt}$ is the melt factor for the day (mm day$^{-1}$ °C$^{-1}$), $SNO_{cov}$ is the fraction of the HRU area covered by snow, $T_{snow}$ is the snow pack temperature on a given day (°C) is a function of the mean daily temperature during the preceding days and varies as a dampened function of air temperature (Anderson, 1976), and $T_{mlt}$ is the base temperature above which snow melt is allowed (°C). The melt factor is calculated as:

$$b_{mlt} = \frac{(SMFMX+SMFMN)}{2} + \frac{(SMFMX-SMFMN)}{2} \times sin\left( \frac{2\pi}{365} \times (d_n - 81) \right) \tag{24}$$

Where $b_{mlt}$ is the melt factor for the day (mm day$^{-1}$ °C$^{-1}$), $SMFMX$ is the melt factor for June *21th* (mm day$^{-1}$ °C$^{-1}$), $SMFMN$ is the melt factor for December *21th* (mm day$^{-1}$ °C$^{-1}$), and $d_n$ is the day number of the year.

### 3.3.3. Evaluation of snowmelt simulation

All scenarios were during the period 1979–1987. The period 1979–1982 was used as the warming period. Table 3.2 lists explanations of 16 parameters needed for model

operation. To test the effects of the various air temperature datasets on snowmelt, all other factors that could influence the simulation results should be fixed. This means subbasin delineation and HRU creation for all scenarios were identical. Furthermore, all the air temperature scenarios shared the same parameter sets, generated by the Latin hypercube method of SWAT-CUP, with sample number 2,000. The workflow of the model testing for all different estimated data is shown in Figure 3.4.

This research uses SWAT Calibration and Uncertainty Program (SWAT-CUP) 4.3.7 (Abbaspour *et al.*, 2007), which integrates Sequential Uncertainty Fitting version 2 (SUFI-2) for model calibration, parameter sensitivity, and uncertainty analysis.

Nash-Sutcliffe Efficiency (*NSE*) and $R^2$ were used for evaluating the snowmelt simulations. *NSE* is calculated by

$$NSE = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \qquad (25)$$

Where $O_i$ is observed runoff on day $i$, $\bar{O}$ is the average observed runoff, $P_i$ is simulated runoff on day $i$, and $\bar{P}$ is the average simulated runoff. Daily absolute error and mean absolute error (*MAE*) were used for evaluating the snowmelt simulations. In addition, the analysis of variance (*ANOVA*, $p<0.05$) and the Steel–Dwass test ($p<0.05$) were used for the nonparametric multiple comparisons of different input data.

Moreover, simulation outputs of hydrologic models depend on their structures, quality of input data, and parameterizations (Muleta and Nicklow, 2005). If we leave model structure and parameter sets unchanged, based on hydrologic model uncertainties we can evaluate uncertainties caused by the various input data. It is clear that air temperature is one of the most important inputs to temperature index-based snowmelt

models. Thus, it is instructive to evaluate uncertainties caused by the various input temperature data via the evaluation of simulation results.

The *p-factor* and *r-factor* indices are used for evaluation of uncertainties. Uncertainty is quantified by the 95% prediction uncertainty band (95PPU) calculated by the 2.5% and 97.5% percentiles of the cumulative distribution function of every simulated discharge point (Abbaspour *et al.*, 2007, Yang *et al.,* 2008). The *p-factor* is the percentage of observed data bracketed by the 95PPU band. The *r-factor* is calculated as

$$r - factor = \frac{\frac{1}{n}\sum_{t_i=1}^{n}(y_{t_{i,97.5\%}} - y_{t_{i,2.5\%}})}{\sigma_{obs}} \tag{26}$$

Where $y_{t_{i,97.5\%}}$ and $y_{ti,2.5\%}$ represent the upper and lower boundaries of 95PPU for a given observation point $t_i$, and $\sigma_{obs}$ symbolizes the standard deviation of the measured data (Yang *et al.,* 2008). The *r-factor* is an index of the relative degree of the uncertainty band against observations.

In the SUFI-2 method, a threshold is needed to distinguish the behavioral and non-behavioral simulations. If behavioral simulations exist, then the *p-factor* and *r-factor* for these solutions are also calculated (Abbaspour *et al.*, 2007). The threshold for distinguishing the behavioral and non-behavioral simulation was 0.5 (*NSE*) in this study. The optimum calibration and parameter uncertainty for both methods is measured on the basis of proximity of the *p-factor* to 100% and *r-factor* to 1 (Yang *et al.,* 2008). According to the guidelines of Moriasi *et al* (2007), runoff simulation can be judged as "satisfactory" if NSE > 0.50 for long-term runoff simulation under a monthly time step.

In addition, Engel *et al* (2007) pointed out that model simulation results will become poorer with a decrease in the simulation time step. Since this study adopted a daily time step for model operations, we used 0.5 as threshold to identify behavioral and non-behavioral simulations.

Table 3.2 Calibration parameters of SWAT model for snowmelt in the test basins

| Parameters | Definition | Method | Range/Percent |
|---|---|---|---|
| ALPHA_BF | Base-flow alpha factor (days) | V* | 0-1 |
| CN2 | Initial SCS CN II value | R** | ±50% |
| ESCO | Soil evaporation compensation factor | V | 0.25–0.75 |
| GW_DELAY | Groundwater delay (days) | V | 0–30 |
| LAT_TTIME | Lateral flow travel time (days) | V | 0–180 |
| OV_N | Manning's "n" value for overland flow | V | 0–0.8 |
| SFTMP | Snowfall temperature (°C) | V | −5 to 5 |
| SMTMP | Snowmelt base temperature (°C) | V | −5 to 5 |
| SMFMN | Melt factor for snow on December 21 (mm°Cday) | V | 0–3 |
| SMFMX | Melt factor for snow on June 21 (mm $H_2O$°Cday) | V | 3–9 |
| SNO50COV | Fraction of snow volume represented by SNOCOVMX corresponding to 50% snow cover | V | 0.01–0.99 |
| SNOCOVMX | Minimum snow water content corresponding to 100% snow cover (mm) | V | 0–500 |
| SOL_AWC | Average available water | R | ±50% |
| SOL_K | Saturated conductivity | R | ±50% |
| SURLAG | Surface runoff lag time (days) | V | 1–24 |
| TIMP | Snowpack temperature lag factor (°C) | V | 0.01–1 |

* V stands for actual value of variation of each parameter.

** R stands for relative range of variation of each parameter.
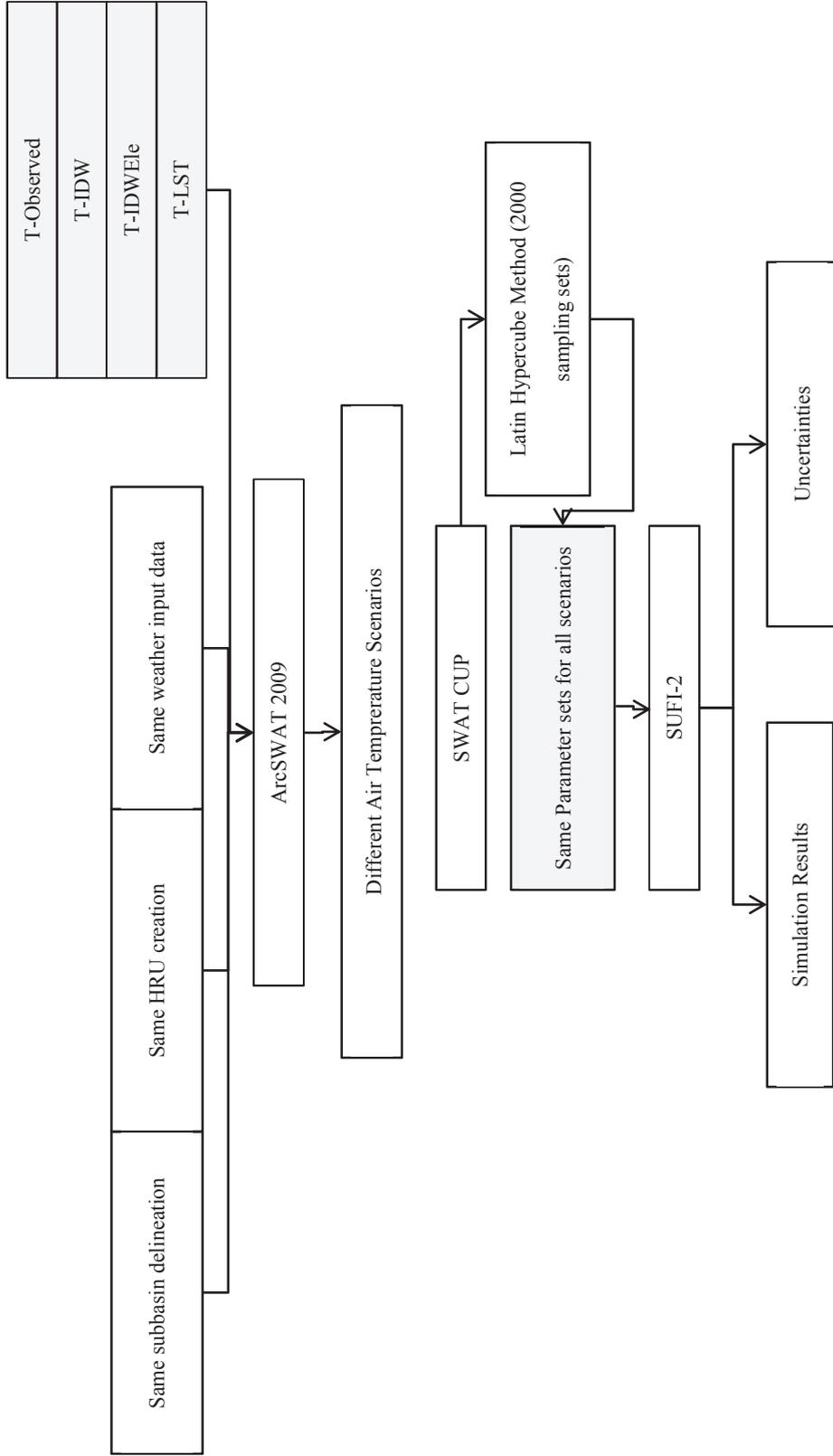
Figure 3.4 The schematic diagram for workflow of model verification for different estimated air temperature data.

43

# 4 Results and discussions

## 4.1. Validation of LST-LST method at observation stations

### 4.1.1. Results of $T_a$-$T_a$ regression analysis

According to the introduction of section 3.1.2, the $T_a$-$T_a$ linear regression relationship between the estimated station and its nearest station should be proved firstly, thus, this study used the observation stations listed in Table 2.2 to conduct the linear regression analyses.

The results are showed in Figure 4.1 and Table 4.1. One example of station 313000 in Gari basin is shown in the Figure 4.2. The results indicated that the $R^2$ of both daily maximum and minimum air temperature data are larger than 0.95 in all test basins (Figure 4.1 (a)), and the standard deviations indicate that the linear regression are relatively stable for the observation stations. For daily maximum air temperature, the results of all test basins shows the same $R^2$ values ($R^2$=0.98). For daily minimum air temperature data, the validation stations of Malinvoka basin show the highest values ($R^2$=0.97) compared with the stations of other two basins. The results of daily maximum data are slightly higher than those of daily minimum data.

In addition, the results of the first order coefficient (Figure 4.1, (b)) of all stations are close to 1. The results of $R^2$ and the first order coefficients clearly demonstrate that that based on the long monitoring period and abundant monitoring data; there are strong linear relationships in all station pairs, and the Eq.(5) of section 3.1.2 are applicable in these test basins for further linear analysis with LST data.

**(a)*R²* of T$_a$-T$_a$ linear regression at validation stations**
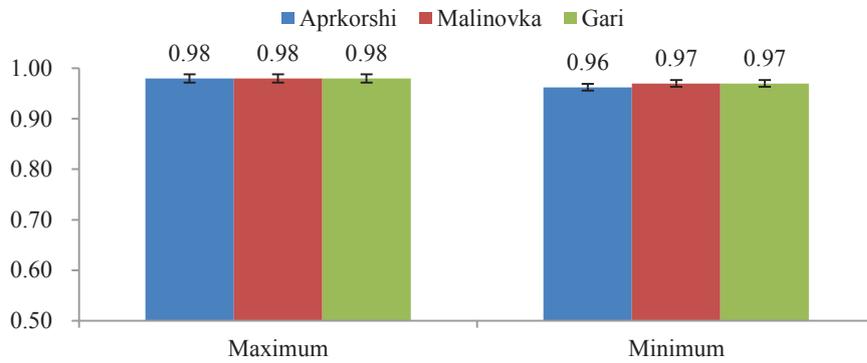


**(b)*first order coefficient* of T$_a$-T$_a$ linear regression at validation stations**
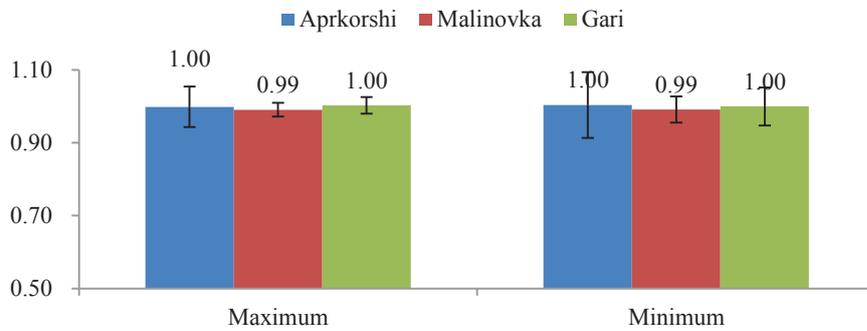


Figure 4.1 The linear regression analysis results of T$_a$-T$_a$ method at observed stations in all test basins.

**Daily Maximum T$_a$-T$_a$**
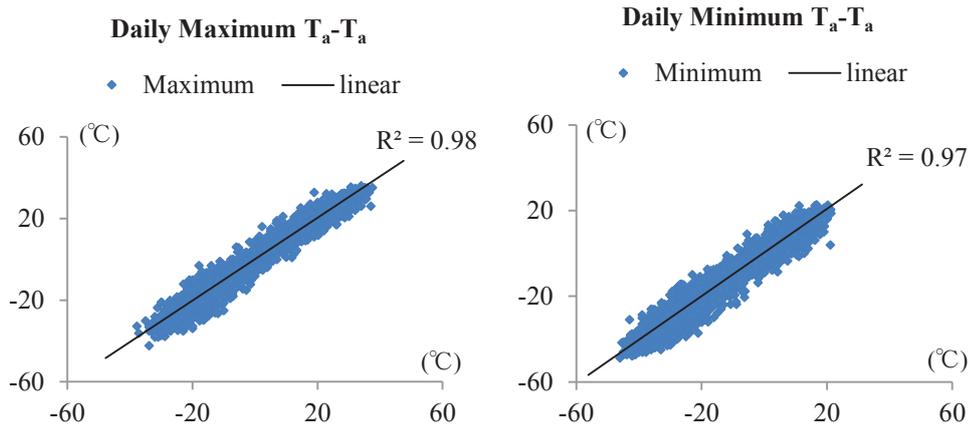
**Daily Minimum T$_a$-T$_a$**



Figure 4.2 The scatter plot of daily maximum and minimum air temperature regression analysis based on the T$_a$−T$_a$ method, taking station 313000 as an example.

Table 4.1 the linear regression analysis results of observed air temperature between the validation station and its nearest stations ($T_a$-$T_a$), (a) daily maximum air temperature, and (b) daily minimum air temperature.

(a) daily maximum air temperature data

| Apkoroshi | | | | Malinovka | | | | Gari | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a_1$* | $b_1$* | WMOID | $R^2$ | $a_1$ | $b_1$ | WMOID | $R^2$ | $a_1$ | $b_1$ |
| 313290 | 0.98 | 1.07 | 0.98 | 318320 | 0.98 | 0.99 | -0.03 | 312530 | 0.98 | 1.02 | -0.46 |
| 313480 | 0.98 | 0.97 | -0.31 | 318450 | 0.98 | 0.99 | 0.21 | 312950 | 0.98 | 0.97 | -0.48 |
| 314160 | 0.98 | 1.02 | -0.15 | 318730 | 0.99 | 1.02 | -0.33 | 313000 | 0.98 | 1.01 | 0.16 |
| 314180 | 0.98 | 0.96 | 0.22 | 318780 | 0.99 | 0.98 | 1.03 | 313710 | 0.98 | 1.01 | 0.57 |
| 314210 | 0.98 | 0.98 | 1.17 | 319130 | 0.97 | 0.99 | 0.01 | 313880 | 0.98 | 1.02 | -1.09 |
| 314740 | 0.98 | 1.05 | 3.40 | 319210 | 0.97 | 0.98 | 0.26 | 314420 | 0.98 | 0.97 | -0.14 |
| 314780 | 0.98 | 0.93 | -3.16 | 319390 | 0.99 | 1.02 | -0.86 | 314450 | 0.98 | 1.02 | 0.23 |
| 314840 | 0.96 | 0.93 | 2.02 | 319810 | 0.99 | 0.96 | 0.98 | 314590 | 0.98 | 0.98 | 0.71 |
| 314890 | 0.97 | 1.08 | -1.55 | 508880 | 0.98 | 1.00 | 0.77 | 503530 | 0.98 | 1.02 | 0.66 |
| | | | | 509830 | 0.99 | 0.97 | 0.43 | | | | |

(b) daily minimum air temperature data

| Apkoroshi | | | | Malinovka | | | | Gari | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a_1$ | $b_1$ | WMOID | $R^2$ | $a_1$ | $b_1$ | WMOID | $R^2$ | $a_1$ | $b_1$ |
| 313290 | 0.97 | 0.99 | 2.40 | 318320 | 0.96 | 0.92 | 2.65 | 312530 | 0.97 | 0.98 | 2.66 |
| 313480 | 0.96 | 0.99 | -4.07 | 318450 | 0.96 | 1.04 | -3.06 | 312950 | 0.97 | 0.89 | 0.15 |
| 314160 | 0.97 | 1.17 | -1.77 | 318730 | 0.98 | 1.00 | -0.99 | 313000 | 0.97 | 1.02 | 0.33 |
| 314180 | 0.97 | 0.83 | 1.28 | 318780 | 0.98 | 1.04 | -0.06 | 313710 | 0.97 | 1.09 | -0.45 |
| 314210 | 0.96 | 0.98 | -1.97 | 319130 | 0.97 | 0.98 | -0.66 | 313880 | 0.97 | 1.01 | 1.61 |
| 314740 | 0.97 | 1.02 | 2.77 | 319210 | 0.97 | 0.99 | 0.60 | 314420 | 0.98 | 0.98 | -0.83 |
| 314780 | 0.97 | 0.95 | -3.10 | 319390 | 0.98 | 1.02 | -1.44 | 314450 | 0.98 | 1.00 | 0.68 |
| 314840 | 0.95 | 1.03 | -1.23 | 319810 | 0.98 | 0.96 | 1.29 | 314590 | 0.96 | 1.03 | -3.86 |
| 314890 | 0.94 | 1.07 | -0.46 | 508880 | 0.98 | 0.98 | 0.10 | 503530 | 0.98 | 1.00 | 1.03 |
| | | | | 509830 | 0.98 | 0.98 | 0.93 | | | | |

$a_1$ and $b_1$ stand for the first order coefficient and constant item of Eq.(5)

### 4.1.2. Results of LST-T$_a$ regression analysis

The results of linear regression analysis between air temperature data and LST data at observed stations are showed in Figure 4.3, Figure 4.4 and the detail results are shown in Table 4.2. One example of station 313000 and its nearest station 312950 are shown in the Figure 4.5. The results indicate that the linear regression between air temperature and LST data can be detected for both validation stations and their nearest stations.

The $R^2$ of the estimated stations (Figure 4.3 (a)) are larger than 0.90, this indicate that the linear relationship between air temperature and LST data are strong at all validation stations. For the results of linear regression between daily maximum LST and air temperature, the validation stations of Apkoroshi basin and Gari basin show the same $R^2$ values ($R^2$=0.95), the $R^2$ value of Malinovka basin is slightly decreased ($R^2$=0.94). For linear results of daily minimum LST and air temperature, the trend is as same as the daily maximum analysis, the validation stations of Apkoroshi basin and Gari basin ($R^2$=0.97) are higher than the result of Malinovka basin ($R^2$=0.96).

The results of the nearest stations (Figure 4.3 (b)) also present the same trend; the values are similar as those of estimated stations, except the $R^2$ values of daily minimum LST-T$_a$ regression are slightly higher than those of daily maximum regression analysis in all test basins.

In addition, this study also analyzed the first order coefficients of linear regression equation (Eq.(7)) to validate whether these values are close to 1. The results of first order coefficient for both validation stations and their nearest stations are shown in Figure 4.4. It is clear that the first order coefficients of the validation stations are close to 1.

For the daily maximum regression analysis, the results of the validation stations are shown in Figure 4.4 (a) and Table 4.2. The largest value is obtained in the Gari basin (1.09), the smallest values is in results of Apkoroshi basin (1.03), the results of the Malinovka basin is 1.05. For the daily minimum LST-$T_a$ regression analysis, the Gari basin obtained the smallest value (1.02) and the values of other two basins (1.03 for both Apkoroshi and Malinovka) are slightly increased.

The results of the nearest stations are shown in Figure 4.4 (b) and Table 4.3, the results show the same trend as the validation stations. The values of daily maximum LST-$T_a$ regression analysis are 1.05, 1.04, 1.11 for Apkoroshi basin, Malinovka basin and Gari basin, respectively. The values of daily minimum LST-$T_a$ regression analysis are 1.04, 1.04, 1.11 for Apkoroshi basin, Malinovka basin and Gari basin, respectively. The values of the nearest stations were increased compared those values of validation stations, especially for the daily maximum data of Gari basin, the values are larger than 1.1 for most of the stations.

Generally, the results clearly demonstrate that there are high linear relationships in all station pairs for the LST data and observed air temperature data, and the first order coefficients of the linear regression equation is close to 1. Thus, these results can give valid support of the assumptions in the Eq.(8) and Eq.(9), which used the linear regression relationships between the air temperature and LST temperature to derive the relationship between $T_a$-Ta and LST-LST.

**(a)$R^2$ of LST-T$_a$ method at validation stations**

**(b)$R^2$ of LST-T$_a$ method at nearest stations**
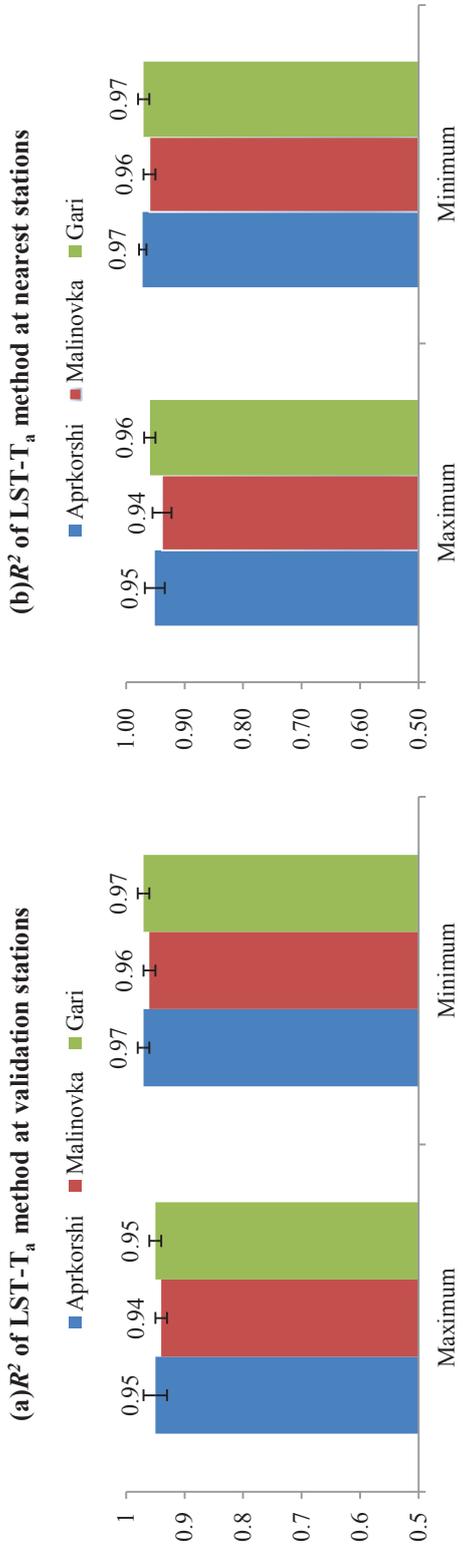
Figure 4.3 The $R^2$ of LST-T$_a$ method at all validation stations and their nearest stations.

**(a) first order coefficient of LST-T$_a$ method at validation stations**

**(b) first order coefficient of LST-T$_a$ method at nearest stations**
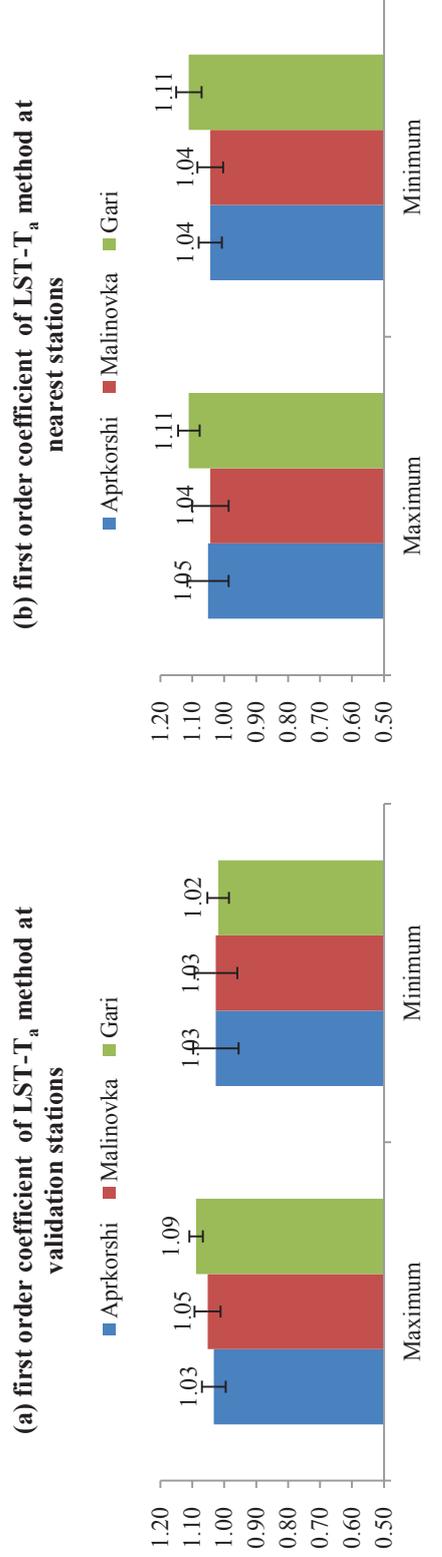
Figure 4.4 The first order coefficient of LST-T$_a$ method at all validation stations and their nearest stations.
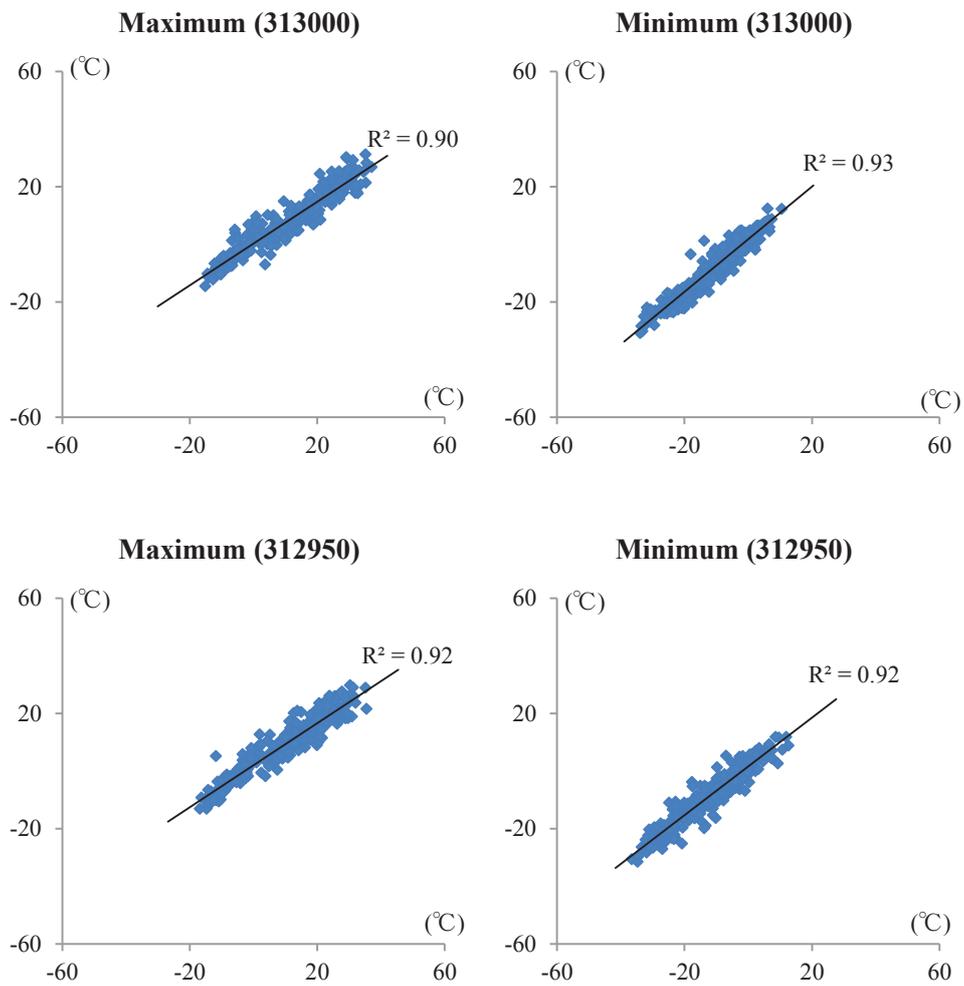
Figure 4.5 The scatter plot of daily maximum and minimum air temperature regression analysis based on the LST−$T_a$ method, taking station 313000 and its nearest station 312950 as an example.

Table 4.2 The linear regression analysis results of air temperature data and LST data ($T_a$-LST) at the all validation stations, (a) daily maximum air temperature data and LST data, and (b) daily minimum air temperature and LST data.

(a) daily maximum air temperature and LST data

| Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a^*$ | WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ |
| 313290 | 0.96 | 1.12 | 318320 | 0.94 | 1.03 | 312530 | 0.95 | 1.10 |
| 313480 | 0.94 | 0.94 | 318450 | 0.93 | 0.99 | 312950 | 0.95 | 1.13 |
| 314160 | 0.97 | 1.08 | 318730 | 0.97 | 1.07 | 313000 | 0.95 | 1.08 |
| 314180 | 0.95 | 0.99 | 318780 | 0.95 | 1.06 | 313710 | 0.97 | 1.10 |
| 314210 | | | 319130 | 0.93 | 0.94 | 313880 | 0.95 | 1.02 |
| 314740 | 0.93 | 0.98 | 319210 | 0.92 | 0.98 | 314420 | 0.95 | 1.08 |
| 314780 | 0.94 | 1.14 | 319390 | 0.93 | 1.08 | 314450 | 0.95 | 1.09 |
| 314840 | 0.92 | 0.99 | 319810 | 0.94 | 1.11 | 314590 | | |
| 314890 | 0.96 | 1.03 | 508880 | 0.95 | 1.15 | 503530 | 0.93 | 1.11 |
| | | | 509830 | 0.93 | 1.11 | | | |

(b) daily minimum air temperature and LST data

| Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ |
| 313290 | 0.98 | 1.01 | 318320 | 0.96 | 1.01 | 312530 | 0.97 | 1.03 |
| 313480 | 0.97 | 0.98 | 318450 | 0.97 | 0.97 | 312950 | 0.97 | 1.05 |
| 314160 | 0.98 | 1.05 | 318730 | 0.96 | 1.09 | 313000 | 0.97 | 1.05 |
| 314180 | 0.97 | 1.10 | 318780 | 0.97 | 1.11 | 313710 | 0.97 | 0.99 |
| 314210 | | | 319130 | 0.95 | 1.02 | 313880 | 0.98 | 1.00 |
| 314740 | 0.98 | 1.01 | 319210 | 0.94 | 1.03 | 314420 | 0.97 | 1.01 |
| 314780 | 0.97 | 1.02 | 319390 | 0.96 | 1.00 | 314450 | 0.97 | 1.01 |
| 314840 | 0.96 | 1.00 | 319810 | 0.96 | 0.99 | 314590 | | |
| 314890 | 0.96 | 1.04 | 508880 | 0.96 | 1.02 | 503530 | 0.98 | 1.02 |
| | | | 509830 | 0.97 | 1.02 | | | |

*$a$ stand for the first order coefficient of Eq.(6)

Table 4.3 The linear regression analysis results of air temperature data and LST data (T$_a$-LST) at the nearest stations of all validation station , (a) daily maximum air temperature data and LST data, and (b) daily minimum air temperature data and LST data

(a) daily maximum air temperature and LST data

| Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a$* | WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ |
| 313290 | 0.94 | 1.14 | 318320 | 0.93 | 0.99 | 312530 | 0.98 | 1.18 |
| 313480 | 0.97 | 1.08 | 318450 | 0.94 | 1.03 | 312950 | 0.97 | 1.10 |
| 314160 | 0.95 | 0.99 | 318730 | 0.93 | 1.11 | 313000 | 0.95 | 1.13 |
| 314180 | 0.97 | 1.08 | 318780 | 0.97 | 1.07 | 313710 | 0.95 | 1.13 |
| 314210 | 0.97 | 1.08 | 319130 | 0.92 | 0.98 | 313880 | | |
| 314740 | 0.94 | 1.14 | 319210 | 0.93 | 0.94 | 314420 | 0.95 | 1.09 |
| 314780 | 0.93 | 0.98 | 319390 | 0.95 | 1.06 | 314450 | 0.95 | 1.08 |
| 314840 | 0.95 | 0.99 | 319810 | 0.93 | 1.08 | 314590 | 0.96 | 1.10 |
| 314890 | 0.92 | 0.99 | 508880 | 0.93 | 1.11 | 503530 | 0.95 | 1.08 |
| | | | 509830 | 0.97 | 1.07 | | | |

(b) daily minimum air temperature and LST data

| Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ | WMOID | $R^2$ | $a$ |
| 313290 | 0.97 | 1.02 | 318320 | 0.97 | 0.97 | 312530 | 0.98 | 0.98 |
| 313480 | 0.98 | 1.05 | 318450 | 0.96 | 1.01 | 312950 | 0.97 | 0.99 |
| 314160 | 0.97 | 1.10 | 318730 | 0.97 | 1.02 | 313000 | 0.97 | 1.05 |
| 314180 | 0.98 | 1.05 | 318780 | 0.96 | 1.09 | 313710 | 0.97 | 1.05 |
| 314210 | 0.98 | 1.05 | 319130 | 0.94 | 1.03 | 313880 | | |
| 314740 | 0.97 | 1.02 | 319210 | 0.95 | 1.02 | 314420 | 0.97 | 1.01 |
| 314780 | 0.98 | 1.01 | 319390 | 0.95 | 1.06 | 314450 | 0.97 | 1.01 |
| 314840 | 0.97 | 1.10 | 319810 | 0.96 | 0.99 | 314590 | 0.96 | 1.10 |
| 314890 | 0.96 | 1.00 | 508880 | 0.97 | 1.02 | 503530 | 0.97 | 1.01 |
| | | | 509830 | 0.96 | 1.09 | | | |

*$a$ stand for the first order coefficient of Eq.(6)

### 4.1.3. Results of LST-LST regression analysis

The results of linear correlation analysis for air temperature at observed stations are showed in Figure 4.6, and the detail information are shown in the Table 4.4. One example of station 313000 is shown in the Figure 4.7.

The results indicate that the $R^2$ of daily maximum and minimum air for LST-LST method are larger than 0.90 in all test basins, and the standard deviations indicate that the results are relatively stable amongst the validation stations. For daily maximum LST-LST regression analysis, the validation stations in Apkoroshi basin shows the smallest value of $R^2$ ($R^2$=0.92), and this value is largest in the Gari basin ($R^2$=0.97), and it is 0.95 in the Malinovka basin. For daily minimum LST, all the test basins shows the similar results, the $R^2$ is 0.98 in the Apkoroshi basin, this value is 0.97 in the Malinvoka basin, and 0.98 in the Gari basin. The values of daily minimum regression analysis are higher than those of daily maximum regression analysis. The results indicate that that based on the LST data; there are high linear correlations in all validation station.

**(a) $R^2$ of LST-LST method of validation stations**



**(b)First order coefficient of LST-LST linear regression at validation stations**
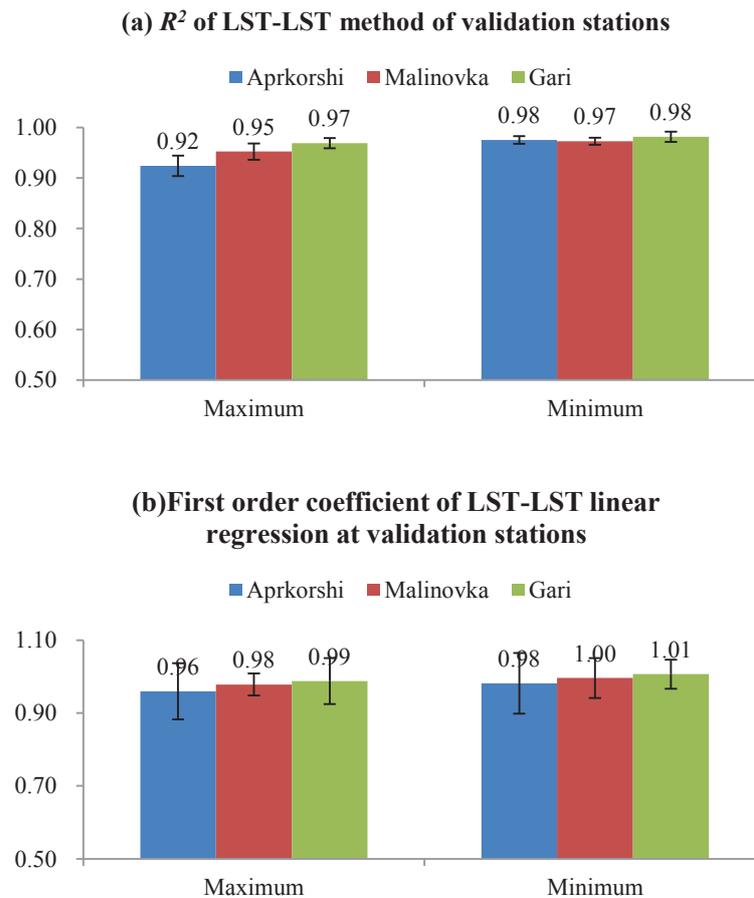


Figure 4.6 The results of LST-LST regression analysis at validation stations.

(a) Daily Maximum LST-LST          (b) Daily MinimumLST-LST
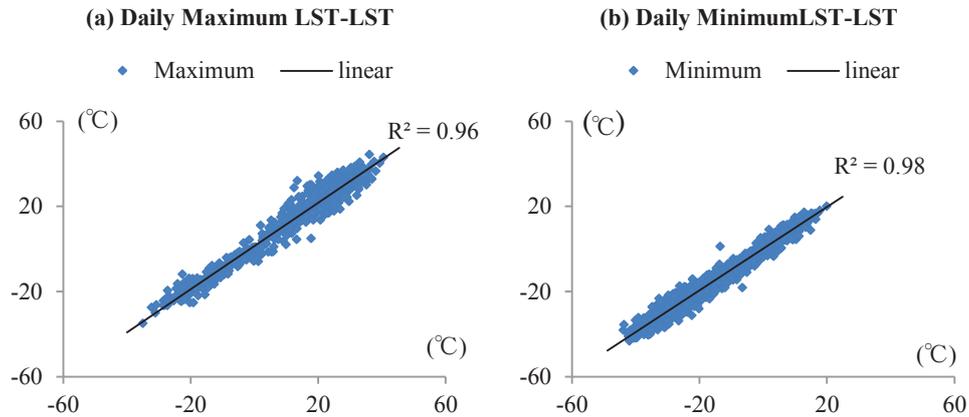
Figure 4.7 The scatter plot of daily maximum and minimum air temperature regression analysis based on the LST−LST method, taking station 313000 as an example.

In addition, the results of first order coefficient for LST-LST are presented in the Figure 4.6 (b). The values of first order coefficients for LST-LST linear regression equation of all test basins are close to 1. The values of first order coefficients for daily maximum LST-LST regression analysis are 0.96, 0.98 and 0.99 for Apkoroshi basin, Malinovka basin, Gari basin, respectively. And, the results for daily minimum LST-LST regression analysis for all three basins are 0.98, 1.00 and 1.01, respectively.

The absolute differences between the first order coefficients of $T_a$-$T_a$ and LST-LST regression analysis are shown in Figure 4.8 (a). The maximum value is 0.04 and 0.02 for daily maximum and minimum temperature. Considering the small difference, the results indicate that the first order coefficient of LST is an approximation of that of $T_a$-$T_a$ method. Moreover, we also listed the results of the differences between the constant item of the linear regression equations of LST-LST and $T_a$-$T_a$ analysis, the results are shown in Figure 4.8 (b), it is clear that the differences between the constant items are around 2℃.

(a) Absolute difference of first order coefficient for $T_a$-$T_a$ and LST-LST linear regression equations

(b) Absolute difference of constant item for $T_a$-$T_a$ and LST-LST linear regression equations
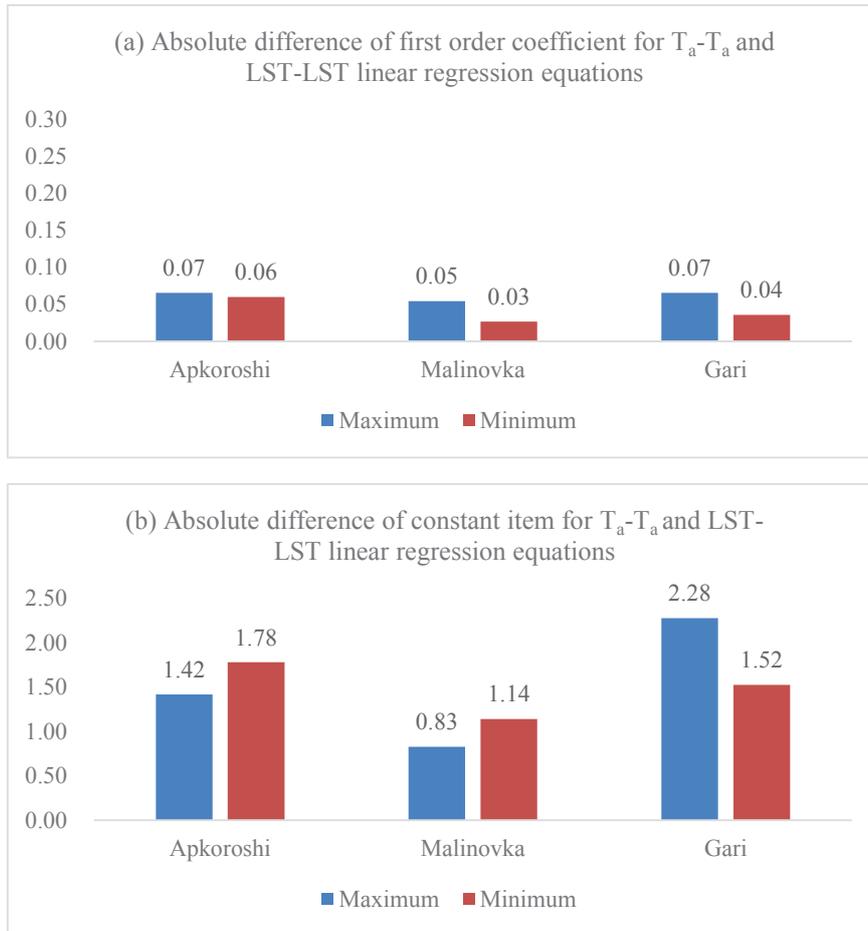
Figure 4.8 The scatter plot of daily maximum and minimum air temperature regression analysis based on the LST−LST method, taking station 313000 as an example.

Totally, these results demonstrate that the LST data of the estimated stations and its nearest station exits strong linear relationships in all test basins, and it also proved the Eq.(7) of section 3.1.2 are applicable in this study area.

Table 4.4 The linear regression analysis results of LST data ($T_a$-LST) at the validation stations and their nearest stations, (a) daily maximum LST data, and (b) daily minimum LST data

(a) daily maximum LST data

| WMOID | $R^2$ | $a_2$* | $b_2$* | WMOID | $R^2$ | $a_2$ | $b_2$ | WMOID | $R^2$ | $a_2$ | $b_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 313290 | 0.96 | 1.02 | 0.64 | 318320 | 0.95 | 0.99 | 1.82 | 312530 | 0.97 | 0.93 | 0.23 |
| 313480 | 0.92 | 0.90 | -0.75 | 318450 | 0.95 | 0.95 | -0.83 | 312950 | 0.98 | 1.01 | 1.64 |
| 314160 | 0.94 | 1.01 | -0.49 | 318730 | 0.93 | 0.93 | -2.85 | 313000 | 0.96 | 0.95 | -0.86 |
| 314180 | 0.94 | 0.93 | 1.41 | 318780 | 0.98 | 0.98 | 0.34 | 313710 | 0.98 | 0.97 | -1.34 |
| 314210 | 0.93 | 0.93 | 0.47 | 319130 | 0.96 | 0.94 | -0.47 | 313880 | 0.97 | 1.01 | -0.94 |
| 314740 | 0.90 | 0.83 | 4.04 | 319210 | 0.96 | 1.02 | 1.09 | 314420 | 0.97 | 0.94 | 0.22 |
| 314780 | 0.90 | 1.09 | -2.80 | 319390 | 0.97 | 0.97 | -0.77 | 314450 | 0.97 | 1.03 | 0.17 |
| 314840 | 0.92 | 0.94 | 2.11 | 319810 | 0.97 | 0.99 | 1.40 | 314590 | 0.98 | 0.93 | -1.00 |
| 314890 | 0.92 | 0.99 | -0.91 | 508880 | 0.95 | 1.01 | 0.71 | 503530 | 0.97 | 1.12 | 0.55 |
| | | | | 509830 | 0.93 | 1.00 | 4.18 | | | | |

(b) daily minimum LST data

| WMOID | $R^2$ | $a_2$ | $b_2$ | WMOID | $R^2$ | $a_2$ | $b_2$ | WMOID | $R^2$ | $a_2$ | $b_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 313290 | 0.98 | 1.01 | 1.94 | 318320 | 0.97 | 0.99 | 1.41 | 312530 | 0.97 | 1.04 | 3.89 |
| 313480 | 0.98 | 0.91 | -1.99 | 318450 | 0.97 | 0.98 | -1.70 | 312950 | 0.98 | 0.92 | -2.96 |
| 314160 | 0.98 | 1.11 | 0.43 | 318730 | 0.97 | 1.03 | -1.69 | 313000 | 0.98 | 1.00 | -0.51 |
| 314180 | 0.98 | 0.88 | -0.69 | 318780 | 0.97 | 1.08 | -0.81 | 313710 | 0.98 | 1.06 | 2.93 |
| 314210 | 0.98 | 0.92 | -1.60 | 319130 | 0.97 | 0.96 | -0.34 | 313880 | 0.98 | 1.03 | 1.96 |
| 314740 | 0.97 | 1.00 | 4.81 | 319210 | 0.97 | 1.01 | 0.15 | 314420 | 0.99 | 0.98 | -1.36 |
| 314780 | 0.97 | 0.97 | -5.25 | 319390 | 0.99 | 1.08 | -2.00 | 314450 | 0.99 | 1.01 | 1.25 |
| 314840 | 0.97 | 0.93 | 0.47 | 319810 | 0.99 | 0.91 | 1.72 | 314590 | 0.98 | 1.00 | -2.85 |
| 314890 | 0.96 | 1.10 | -1.17 | 508880 | 0.97 | 0.97 | -0.53 | 503530 | 0.98 | 1.02 | 0.35 |
| | | | | 509830 | 0.97 | 0.95 | 1.42 | | | | |

*$a_2$, $b_2$ stand for the first order coefficient and the constant item of Eq.(7)

### 4.1.4. Calculation results of the ignorance item in LST-LST method

As presented in section 3.1.2, the constant item $const_B - a_2 \times const_A$ of Eq.(11) may influence the accuracy of LST-LST method. Thus, based on the analysis results of LST-$T_a$ method and LST-LST method, this study calculated the average values this

constant item at validation stations of different test basins, the results are presented in Figure 4.9 and Table 4.5. The values of Apkoroshi basin are -0.44°C and 0.46°C for daily maximum and daily minim LST-LST regression analysis, respectively. These values are -0.54°C and 0.2°C in Malinovka basin, and 0.55°C, -0.01°C in Gari basin.

The results indicate that the values of constant item is relatively higher for daily maximum calculation compared those values in the daily minimum calculation. The results are positive in Gari basin, negative in Apkoroshi basinand Malinovka basin (-0.13) for daily maximum data. Compared with the daily maximum results; the results of daily minimum data present a contrary trend. The values are positive in Apkoroshi basin (0.46) and Malinovka basin (0.20), while the result is negative in Gari.

Considering the relatively low values of the item $const_B - a_2 \times const_A$ in the test basin, the constant item of Eq.(11) were ignored in this study and the Eq.(12) is used to estimate air temperature data in the location of each subbasin.

**Avereage values of the constant item for LST-LST method in all test basins**
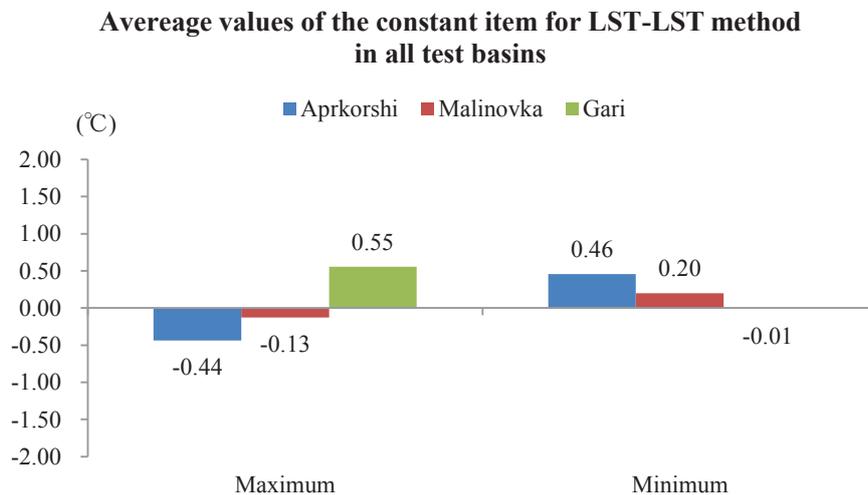


Figure 4.9    Calculation results of constant item ($const_B - a_2 \times const_A$) for LST-LST method at validation stations.

Generally, the analysis results of $T_a$-$T_a$, LST-$T_a$ and LST-LST clearly demonstrated that there are strong linear relationship between the air temperature data and LST data in

both estimated station and its nearest station, the assumptions and approximation of the section 3.1.2 are valid and acceptable.

Table 4.5 The calculation results of ignorance constant item at the validation stations in all test basins, (a) daily maximum data, and (b) daily minimum data

(a) daily maximum data

| Apkoroshi | | Malinovka | | Gari | |
|---|---|---|---|---|---|
| WMOID | CItem*(℃) | WMOID | CItem(℃) | WMOID | CItem(℃) |
| 313290 | -1.58 | 318320 | 1.49 | 312530 | 0.67 |
| 313480 | -0.16 | 318450 | -1.57 | 312950 | 1.96 |
| 314160 | -1.93 | 318730 | -3.37 | 313000 | -1.10 |
| 314180 | 1.71 | 318780 | -0.88 | 313710 | -1.98 |
| 314210 | | 319130 | -0.61 | 313880 | |
| 314740 | -0.48 | 319210 | 0.62 | 314420 | 0.09 |
| 314780 | 0.25 | 319390 | 0.50 | 314450 | -0.12 |
| 314840 | 0.17 | 319810 | 0.06 | 314590 | 1.75 |
| 314890 | -1.47 | 508880 | -1.11 | 503530 | 3.18 |
| | | 509830 | 3.53 | | |

(b) daily minimum data

| Apkoroshi | | Malinovka | | Gari | |
|---|---|---|---|---|---|
| WMOID | CItem(℃) | WMOID | CItem(℃) | WMOID | CItem(℃) |
| 313290 | -1.12 | 318320 | -2.33 | 312530 | 0.31 |
| 313480 | 3.26 | 318450 | 2.37 | 312950 | -2.98 |
| 314160 | 2.57 | 318730 | -0.84 | 313000 | -0.94 |
| 314180 | -2.32 | 318780 | 0.30 | 313710 | 3.19 |
| 314210 | | 319130 | 0.61 | 313880 | 1.54 |
| 314740 | 2.04 | 319210 | -0.67 | 314420 | -0.44 |
| 314780 | -1.99 | 319390 | 2.58 | 314450 | 0.45 |
| 314840 | 2.48 | 319810 | 0.06 | 314590 | |
| 314890 | -1.26 | 508880 | -0.88 | 503530 | -1.13 |
| | | 509830 | 0.77 | | |

*CItem is the constant item of Eq.(11)

## 4.2. Validation of estimated air temperature at observation stations

As presented, the spatial interpolation methods are also frequently used for estimating the spatial dense air temperature data. Thus, this study also used the IDW and IDWEle method to estimate the air temperature data at the observed stations (Table 2.1). The results of the IDW, IDWEle and LST-LST methods are presented and compared. The elevation information of IDW and IDWEle method are listed in Table 4.6. In addition, the number of available training data for linear regression analysis of LST-LST method is also listed in Table 4.6. According to the Eq.(3) , the station amount for linear regression between air temperature and elevation should be larger than three, however, for the station 312530 and station 313290 (Table 2.2), there are less than three stations for IDW and IDWEle method, thus they are excluded from the validations.

Table 4.6 The basic information of validations stations used for different data created methods

| Apkoroshi | | | Malinovka | | | Gari | | |
|---|---|---|---|---|---|---|---|---|
| Station | Ele_Sta(m) | Ele_Inter(m) | Station | Ele_Sta(m) | Ele_Inter(m) | Station | Ele_Sta | Ele_Inter |
| 313480 | 153 | 238 | 318320 | 68 | 111 | 312950 | 370 | 263 |
| 314160 | 73 | 220 | 318450 | 130 | 101 | 313000 | 229 | 312 |
| 314180 | 201 | 233 | 318730 | 101 | 107 | 313710 | 210 | 291 |
| 314210 | 60 | 117 | 318780 | 98 | 132 | 313880 | 208 | 243 |
| 314740 | 384 | 444 | 319130 | 88 | 171 | 314420 | 281 | 216 |
| 314780 | 902 | 294 | 319210 | 78 | 190 | 314450 | 197 | 216 |
| 314840 | 269 | 312 | 319390 | 259 | 131 | 314590 | 261 | 253 |
| 314890 | 92 | 128 | 319810 | 188 | 176 | 503530 | 179 | 218 |
| | | | 508880 | 83 | 114 | | | |
| | | | 509830 | 103 | 119 | | | |

Ele_Sta: the elevation of validation station.

Ele_Inter: the average elevation of the interpolation stations.

### 4.2.1. Performances of different methods at all stations

The validation results of estimated air temperature data at observation stations are shown in Figures 4.10. The scatter plot of different estimated data versus the observed air temperature data is shown Figure 4.10, taking the station 313000 as an example. According to the results of $R^2$ and the *RMSE*, The results indicate that the IDW method can always obtain best results compared with other the IDWEle method and LST-LST method.

The values of $R^2$ are over 0.80 at all validation stations for the daily maximum air temperature estimations. This indicates that all the estimated data has a strong linear relationship with the observed data. For the IDW method, the values of $R^2$ for daily maximum data is 0.91. These values are 0.86 and 0.88 for IDWEle method and LST-LST method. The results of $R^2$ indicate that the strong linear relationship can be detected for all the data creation methods. However, for the daily minimum data estimation, the results are decreased in all the methods compared with the results of daily maximum data. The $R^2$ are 0.77, 0.68, 0.69 respectively for IDW method, IDWEle method and LST-LST method.

For the IDW method, the values of *RMSE* for daily maximum and minimum air temperature estimation are 2.02°C and 2.50°C, respectively. These values are 2.60°C and 3.11°C for IDWEle method, and 2.74°C, 3.04°C of the LST-LST method. As same as the results of $R^2$, the IDW method obtained the best results. However, the differences of the different methods are less than 1°C. And the results also indicate that the estimated errors are larger for the daily minimum air temperature than the daily maximum air temperature.
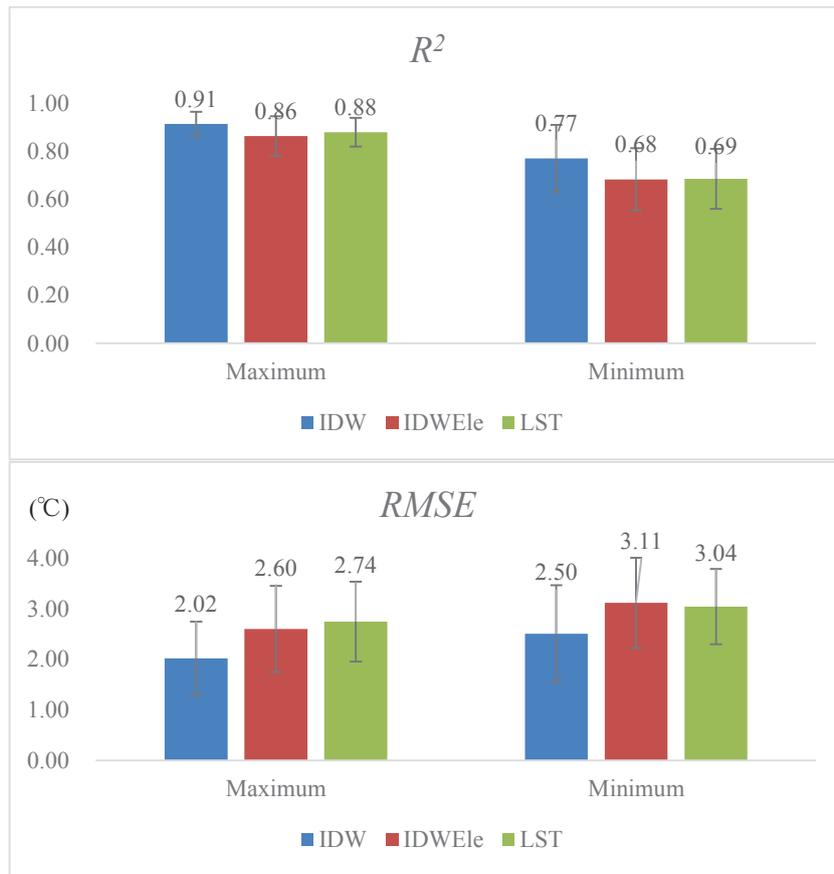
Figure 4.10 The *R²* and *RMSE* of all validation stations.
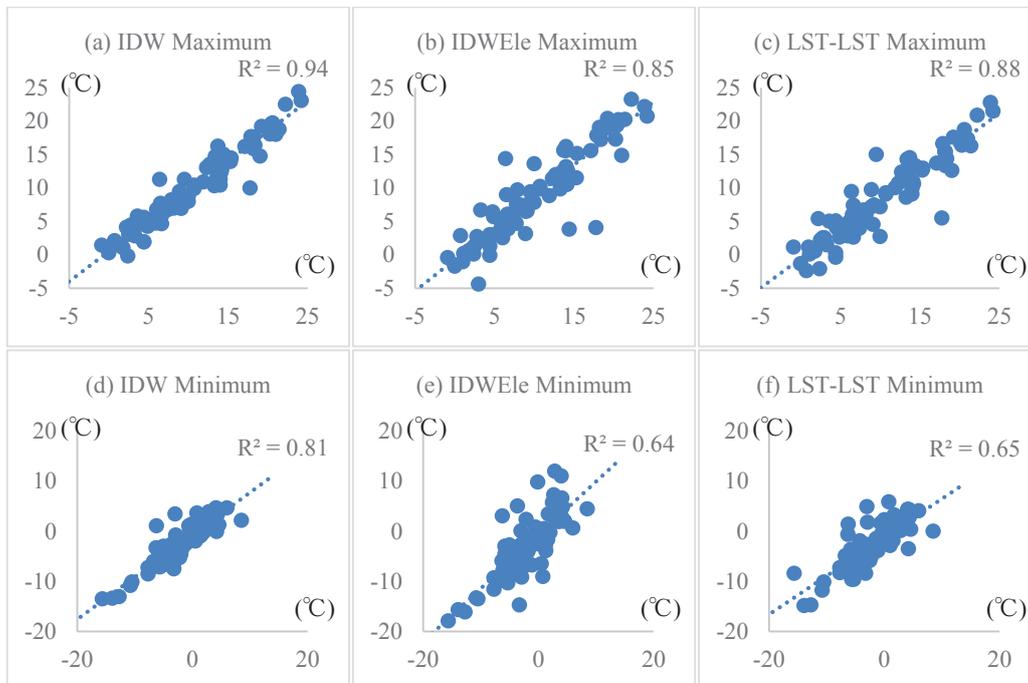


Figure 4.11 The scatter plot of created air temperature during snowmelt period (1983-1987), taking station 313000 station of Gari basin as an example.

### 4.2.2. Validation results in different test basins

There are nine validation stations in Apkoroshi basin, the results of average values of $R^2$ and $RMSE$ are shown in Figure 4.11 (a) and Figure 4.12 (a), and the detail results are listed in Table 4.7 and Table 4.8. All the threes methods obtained relative good results in the Malinovka basin. Judged by the values of $RMSE$, the results of Apkoroshi basin are worse than those results of Gari basin, while according to the results of $R^2$, the performances of Apkoroshi basin is better for daily minimum data and the Gari basin is better for daily maximum data.

The values of $R^2$ show the same trend as the results that are shown in Figure 4.12 (a). The IDW method always obtains the best results following with LST-LST method and IDWEle method. The average values of $R^2$ in Apkoroshi basin are lower than those of all stations. For the $RMSE$ index (Figure 4.12 (a)), the results show that for the daily minimum air temperature estimation, the results of Apkoroshi basins are larger than that of all stations. However, the values of $RMSE$ for daily maximum estimation are improved nay the IDW method and LST-LST method.

The average values of $R^2$ and $RMSE$ are shown of the validation stations in Malinovka basin are shown in Figure 4.12 (b) and Figure 4.13 (b). The $R^2$ and the $RMSE$ values presented the same trends, the best values are obtained in IDW method, follow with IDWEle method and LST method. The $R^2$ and $RMSE$ values are better compared with those values of all stations (Figure 4.10).

The results of Gari basin are shown in Figure 4.12 (c) and Figure 4.13 (c). The $R^2$ and the RMSE values presented the same trends as those of all stations (Figure 4.8); the best values are obtained in IDW method, following with LST-LST method and IDWEle method. Especially, for the IDWEle method, its $R^2$ values of daily minimum air temperature are decreased compared with the results in the Apkoroshi basin and Malinovka Basin. The values of $RMSE$ (Figure 4.13 (c)) in Gari basin are similar with the results of Apkoroshi basin (Figure 4.13 (a)) while larger than the results of Malinovka basin.

Generally, according to the results of $R^2$, all the estimation data can present strong linear relationships with the observed data. However, judged by the values of $RMSE$, it is clear that the estimated methods have different performances in different basins. The

methods can perform well in Malinovka basin with error controlled less than 3°C, followed by Gari basin and Apkoroshi basin. And for the daily maximum air temperature estimation, its results are always better than those of daily minimum air temperature.
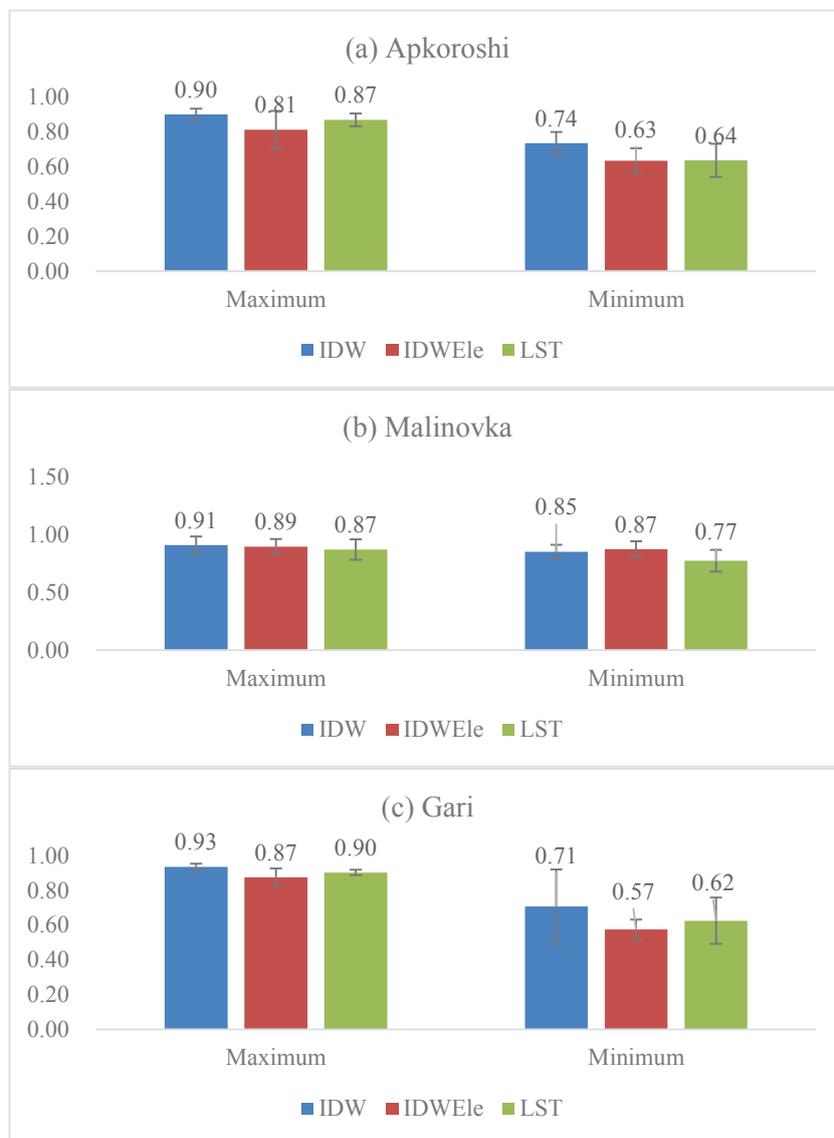


Figure 4.12 The results of $R^2$ at validation stations in different test basins during snowmelt period (1983-1987), (a) Apkoroshi basin, (b) Malinovka basin, (c) Gari basin.
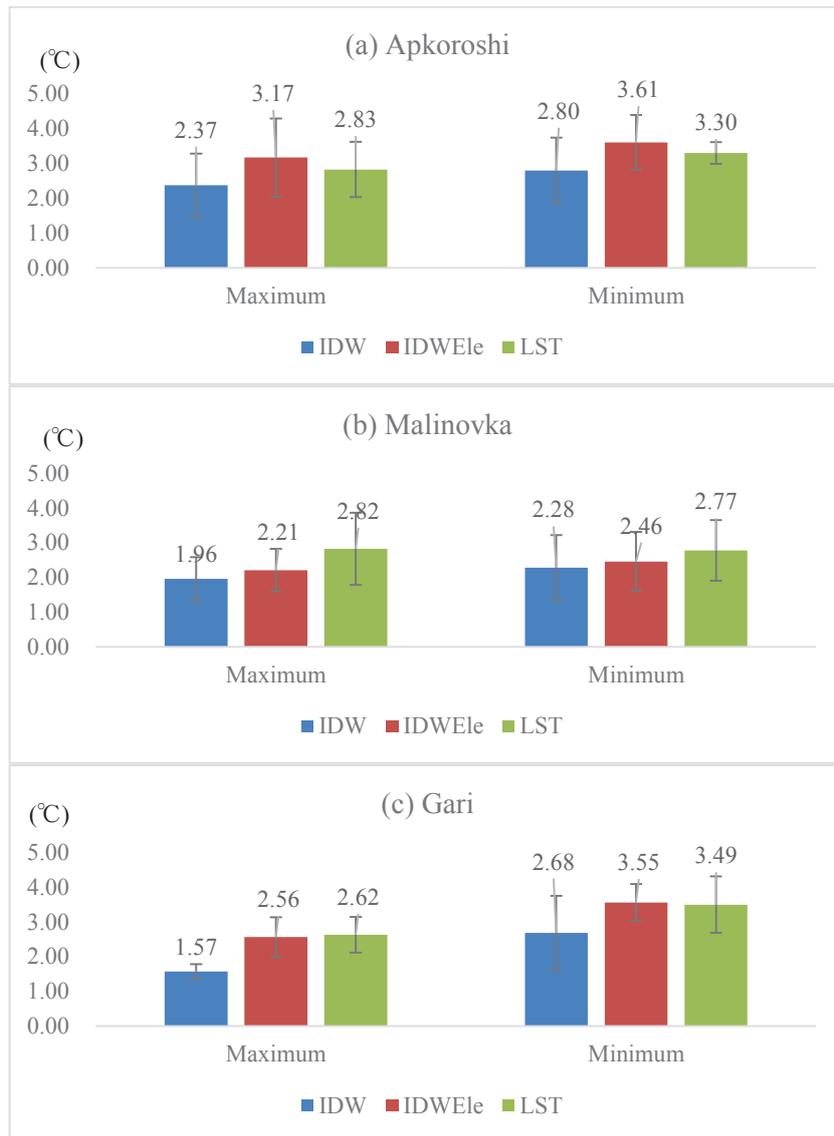
Figure 4.13 The results of *RMSE* at validation stations in different test basins during snowmelt

period (1983-1987), (a) Apkoroshi basin, (b) Malinovka basin, (c) Gari basin.

Table 4.7 The $R^2$ and $RMSE$ of daily maximum air temperature in all validation stations.

| Apkoroshi | IDW | | IDWEle | | LST-LST | |
|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 313480 | 0.88 | 2.36 | 0.86 | 2.45 | 0.93 | 2.04 |
| 314160 | 0.94 | 1.67 | 0.91 | 2.04 | 0.85 | 2.76 |
| 314180 | 0.95 | 1.40 | 0.89 | 2.33 | 0.85 | 3.34 |
| 314210 | 0.87 | 2.55 | 0.79 | 3.39 | 0.85 | 4.19 |
| 314740 | 0.91 | 2.61 | 0.87 | 3.04 | 0.85 | 2.48 |
| 314780 | 0.86 | 4.57 | 0.56 | 5.37 | 0.83 | 2.98 |
| 314840 | 0.88 | 2.59 | 0.80 | 3.74 | 0.83 | 2.98 |
| 314890 | 0.92 | 1.97 | 0.81 | 3.22 | 0.93 | 2.04 |
| Malinovka | IDW | | IDWEle | | LST-LST | |
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 318320 | 0.96 | 1.23 | 0.92 | 1.89 | 0.87 | 2.84 |
| 318450 | 0.89 | 2.22 | 0.87 | 2.69 | 0.87 | 2.60 |
| 318730 | 0.94 | 1.55 | 0.93 | 1.65 | 0.92 | 4.14 |
| 318780 | 0.96 | 1.40 | 0.94 | 1.89 | 0.93 | 1.81 |
| 319130 | 0.73 | 3.17 | 0.73 | 3.29 | 0.71 | 3.14 |
| 319210 | 0.83 | 2.56 | 0.85 | 2.44 | 0.71 | 3.36 |
| 319390 | 0.93 | 2.12 | 0.90 | 3.00 | 0.93 | 1.80 |
| 319810 | 0.93 | 2.31 | 0.93 | 1.80 | 0.93 | 1.87 |
| 508880 | 0.95 | 1.69 | 0.93 | 1.95 | 0.91 | 1.88 |
| 509830 | 0.96 | 1.32 | 0.94 | 1.50 | 0.92 | 4.73 |
| Gari | IDW | | IDWEle | | LST-LST | |
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 312950 | 0.94 | 1.74 | 0.90 | 2.20 | 0.92 | 3.21 |
| 313000 | 0.94 | 1.83 | 0.85 | 3.07 | 0.88 | 3.14 |
| 313710 | 0.96 | 1.31 | 0.92 | 2.04 | 0.92 | 3.15 |
| 313880 | 0.91 | 1.90 | 0.81 | 2.82 | 0.89 | 2.07 |
| 314420 | 0.92 | 1.81 | 0.79 | 3.54 | 0.90 | 2.05 |
| 314450 | 0.95 | 1.44 | 0.93 | 1.74 | 0.90 | 2.12 |
| 314590 | 0.92 | 1.40 | 0.88 | 2.33 | 0.91 | 2.75 |
| 503530 | 0.93 | 1.82 | 0.91 | 2.19 | 0.89 | 2.61 |

Table 4.8 The $R^2$ and $RMSE$ of daily minimum air temperature in all validation stations.

| Apkoroshi | IDW | | IDWEle | | LST-LST | |
|---|---|---|---|---|---|---|
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 313480 | 0.67 | 4.16 | 0.54 | 4.58 | 0.57 | 3.56 |
| 314160 | 0.76 | 1.79 | 0.65 | 2.76 | 0.65 | 2.88 |
| 314180 | 0.66 | 3.07 | 0.57 | 4.65 | 0.65 | 2.73 |
| 314210 | 0.75 | 2.32 | 0.60 | 3.00 | 0.67 | 2.43 |
| 314740 | 0.85 | 1.83 | 0.74 | 2.43 | 0.76 | 3.42 |
| 314780 | 0.77 | 4.61 | 0.73 | 3.85 | 0.76 | 3.62 |
| 314840 | 0.70 | 2.30 | 0.63 | 3.20 | 0.51 | 3.14 |
| 314890 | 0.73 | 1.96 | 0.63 | 2.95 | 0.53 | 3.79 |
| Malinovka | IDW | | IDWEle | | LST-LST | |
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 318320 | 0.80 | 2.86 | 0.72 | 3.05 | 0.60 | 4.40 |
| 318450 | 0.71 | 4.23 | 0.67 | 4.24 | 0.60 | 4.35 |
| 318730 | 0.87 | 1.81 | 0.84 | 2.26 | 0.82 | 2.47 |
| 318780 | 0.90 | 1.75 | 0.87 | 1.96 | 0.81 | 2.74 |
| 319130 | 0.85 | 1.72 | 0.85 | 1.78 | 0.80 | 2.03 |
| 319210 | 0.88 | 1.57 | 0.85 | 1.85 | 0.80 | 2.10 |
| 319390 | 0.84 | 3.45 | 0.80 | 3.46 | 0.83 | 2.65 |
| 319810 | 0.85 | 2.04 | 0.87 | 1.92 | 0.83 | 2.41 |
| 508880 | 0.86 | 2.06 | 0.84 | 2.31 | 0.83 | 2.29 |
| 509830 | 0.93 | 1.32 | 0.88 | 1.74 | 0.82 | 2.30 |
| Gari | IDW | | IDWEle | | LST-LST | |
| WMOID | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) | $R^2$ | $RMSE$ (℃) |
| 312950 | 0.70 | 2.56 | 0.54 | 4.19 | 0.64 | 4.03 |
| 313000 | 0.81 | 2.06 | 0.61 | 3.82 | 0.65 | 2.90 |
| 313710 | 0.78 | 2.29 | 0.53 | 3.50 | 0.64 | 4.18 |
| 313880 | 0.68 | 2.49 | 0.52 | 4.13 | 0.51 | 3.46 |
| 314420 | 0.82 | 2.16 | 0.53 | 3.53 | 0.70 | 2.64 |
| 314450 | 0.80 | 2.46 | 0.68 | 2.99 | 0.70 | 2.71 |
| 314590 | 0.20 | 4.65 | 0.51 | 3.56 | 0.36 | 3.82 |
| 503530 | 0.86 | 1.56 | 0.57 | 3.25 | 0.79 | 2.03 |

### 4.2.3. Influence factors for air temperature estimation at the observation station

#### 4.2.3.1. Influence factors for IDW and IDWEle method

For IDW and IDWEle methods, the elevation information of the validation stations and their interpolation stations are shown in Figure 4.14.

The average elevations of interpolation stations and the validation stations are lower than 300m for all test basins (Figure 4.14 (a)). It is clear that both the interpolation stations and estimated stations have low elevation in all test basins. Moreover, the differences between the validation stations and their interpolation stations relatively small, the values are 19 m, 15 m and 9 m for Apkoroshi basin, Malinovka basin and Gari basin, respectively.

And the average number of interpolation stations was also calculated for each test basin, there are six interpolation stations in the Apkoroshi basin and the Gari basin for each validation station, and for the Malinovka there are nine interpolation stations for each validation station. According to the Eq.(3), the regression analysis between the elevation and the air temperature data requires a certain amount of interpolation stations, thus, this is another reason that the IDWEle method can obtain relatively good results in Malinovka basin compared with other basins.
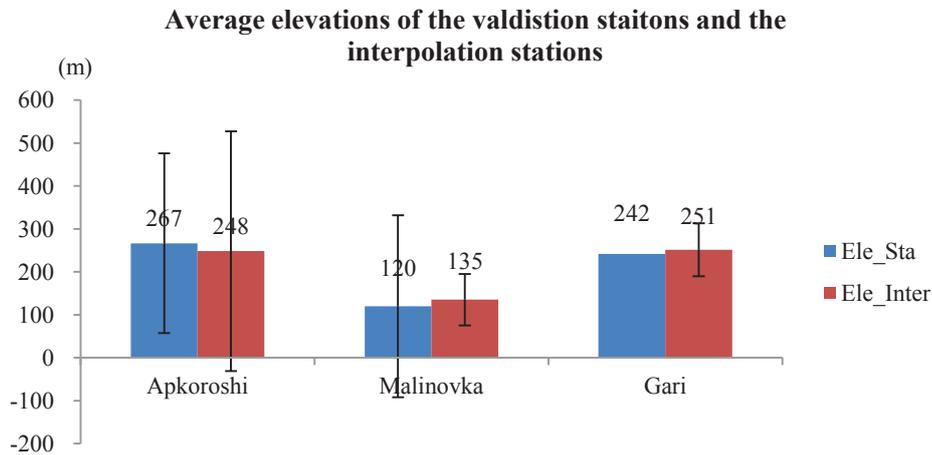


Figure 4.14 The elevation information of validation stations and the interpolation stations.
Ele_Sta stand for the validation station and Ele_Inster stand for the interpolation stations.

For the IDW method which has no corrections on the topographic effects, their relative good performances are relying on the similarity of elevations between the validation stations and the interpolation stations (Stahl *et al.,* 2006).

For the IDWEle method, the elevation variation is important for constructing the linear regression of Eq.(3). Previous studies (Stahl *et al*., 2006; Thornton *et al.,* 1997) also proved that for the IDWEle method, the accurate results are obtained when there are stations located at high elevations. Stahl *et al.,* 2006 demonstrated that without the stations located in the high elevation, the performances of IDWEle method were even worse than directly assigned the data of nearest stations. Thus, considering the results of the IDW method are always better than IDWEle method in different basins, it is clear the interpolation station with low station density and low variations can finally lead to larger errors for the IDWEle method.

In our study, we calculated the In the IDWEle method, the temperature lapse rate (TLAPS, $b_{1,d}$ of Eq. (4)) is used to estimate the effects of topography on air temperature, this study calculated the average value of TLAPS and the results are shown in Table 4.9 for both daily maximum and minimum data in all test basins.

Most of the TLAPS values are positive for daily maximum data in Apkorohsi basin and Malinovka basin and negative in the Gari basin. For daily minimum data, the negative values can be detected for all three test basins. It is common that the air temperature are decreased with the elevation, however, the positive TLAPS values of IDWEle method at the validation stations indicate that the air temperature are increasing with the elevation in several validation stations. In addition, judged by the values of the TLAPS, it is clear that these results are different from the normal value that usually employed for air temperature interpolations (-6.0 °C/km or -6.5 °C/km, Dodson and Marks, 1997).

By using the interpolation stations (Table 4.9), this study conducted the liner regression analysis between elevation and the average air temperature data. The results are shown in Figure 4.15. Though the *p* values are larger than 0.05, positive TLAPS can be detected for daily maximum data and negative TLAPS can detected for daily minimum air temperature data in Malinovka basin (Figure 4.15 (b) and Figure 4.15 (d)). It is clearly that the results shown in Table 4.9 are consistent the results of Figure 4.15 for Malinovka basin. However, it is interesting that the though the interpolation stations used for Apkoroshi basin presented a negative TLAPS for daily maximum data at interpolation stations, the positive TLAPS values can still be detected for IDWEle method in some validation stations for daily maximum data (Stations: 313480, 314160,

314180, 314890). Only the results of the Gari basin shows the consistent trend for TLAPS of interpolation stations and the IDWEle methods.

For the IDWEle method, the accuracy of interpolation depends on the difference in elevation between the observation station and the location where one applies the IDWEle method (Thornton *et al.,* 1997; Stahl *et al.,* 2006). For the IDWEle method, elevation differences between the interpolation stations and validation stations are small (Figure 4.14) and this may lead to worse results for the IDWEle method compared with the IDW method.

Table 4.9 The TLAPS of IDWEle method for daily temperature creation in different test basins.

| WMOID | Apkoroshi | | WMOID | Malinovka | | WMOID | Gari | |
| | Maximum (℃/km) | Minimum (℃/km) | | Maximum (℃/km) | Minimum (℃/km) | | Maximum (℃/km) | Minimum (℃/km) |
|---|---|---|---|---|---|---|---|---|
| | | | 318320 | -11.7 | -4.96 | | | |
| 313480 | 2.8 | -5.1 | 318450 | -3.1 | -3.82 | 312950 | -7.7 | -17.0 |
| 314160 | 1.9 | -4.9 | 318730 | 1.1 | -14.7 | 313000 | -11.8 | -1.5 |
| 314180 | 1.8 | -4.5 | 318780 | 7.4 | -8.9 | 313710 | -9.9 | -11.0 |
| 314210 | -7.1 | -8.1 | 319130 | -1.0 | -4.2 | 313880 | 0.0 | -2.61 |
| 314740 | -8.3 | -8.3 | 319210 | -0.2 | -3.4 | 314420 | -2.86 | -12.6 |
| 314780 | -2.0 | -2.0 | 319390 | 12.2 | 2.5 | 314450 | -11.3 | -11.5 |
| 314840 | -4.8 | -5.4 | 319810 | 1.0 | -5.0 | 314590 | -2.5 | -7.2 |
| 314890 | 2.34 | -4.5 | 508880 | 5.4 | 5.1 | 503530 | -14.7 | -16.4 |
| | | | 509830 | 7.6 | -5.5 | | | |
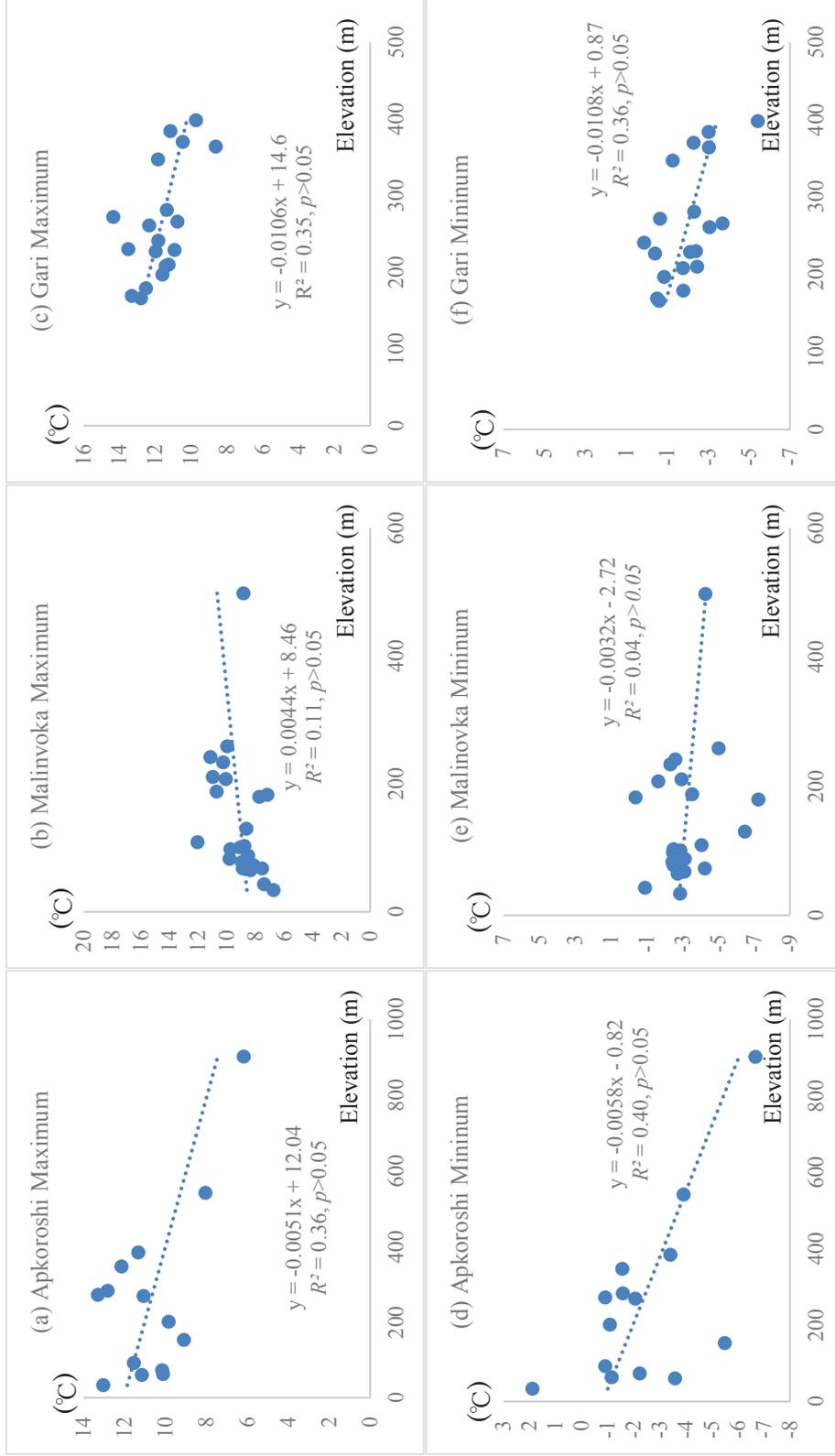
Figure 4.15 The linear regression of elevations and the average temperature at the interpolation stations of all test basins.

### 4.2.3.2.  Influence factors for LST-LST method

As the values of $a_1$ and $a_2$ were close to 1 (Table III), the differences in temperature at the two stations can be approximated using the values of $b_1$ and $b_2$. They are the TLAPS of the $T_a$-$T_a$ regression equation and LST-LST method. The linear relationships between elevation differences and $b_1$, $b_2$ are shown in Figures 4.16, 4.17 and 4.18.

The results of Apkoroshi basin indicate that $b_1$ and $b_2$ had a strong linear relationship with the elevation difference for daily minimum air temperature and LST data, the linear relationship is insignificant for daily maximum data. The slopes of $b_2$ (Figure 1 (c) and Figure1 (d)) were –4.3°C/km and –8.4°C/km for daily maximum and minimum LST data, respectively and these values are –4.8°C/km and –5.7°C/km. According to the result shown in the Figure 4.15, these values are –5.1°C/km and –3.2°C/km for available interpolation stations respectively. The daily maximum values of $b_1$ and $b_2$ are very close to the TLAPS of the interpolation stations, while the differences are larger for the daily minimum data, and this might be the results that lead to the better results of *RMSE*s for the LST-LST method compared with the IDWEle method (Figure 4.13(b)).



(a) Maximum, $b_1$ (℃), y = -0.0044x + 0.094, $R^2$ = 0.48, p>0.05, Elevation (m); (b) Minimum, $b_1$ (℃), y = -0.0057x - 0.93, $R^2$ = 0.51, p <0.05, Elevation (m)

(c) Maximum          (d) Minimum

$b_2$ (℃)             $b_2$ (℃)

6                     6

4    ●   y = -0.0043x + 0.22        4    ●   y = -0.0084x - 0.72
         $R^2 = 0.43, p>0.05$                 $R^2 = 0.82, p<0.05$
2                     2

0                     0

-2                    -2

-4         Elevation (m)    -4         Elevation (m)

-6                    -6

-1000  -500   0   500  1000      -1000  -500   0   500  1000
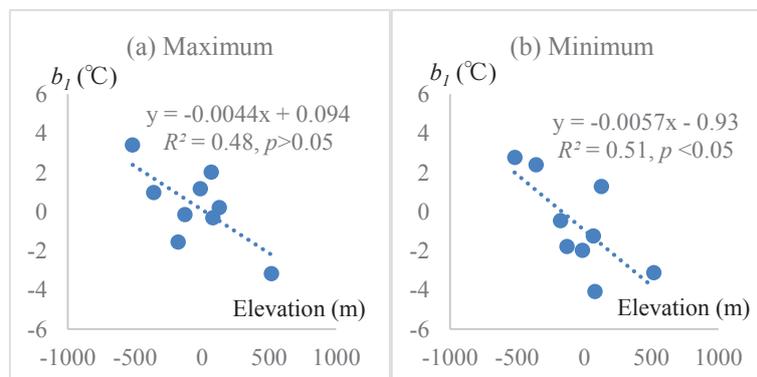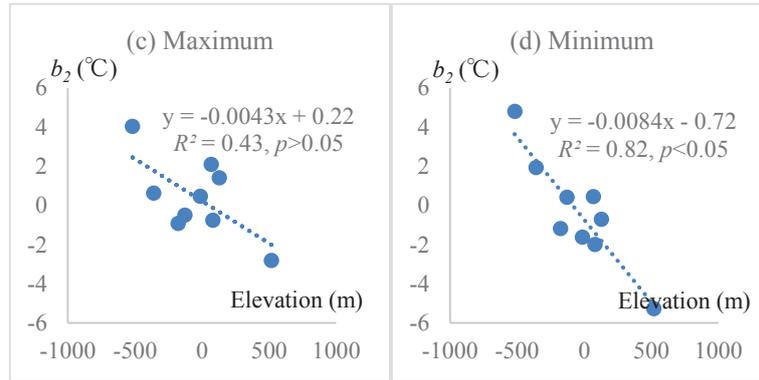
Figure 4.16 Regression analysis between the temperature differences ($b1$ and $b2$) and the elevation difference in the LST- LST method of Apkoroshi basin, (a) daily maximum air temperature, (b) daily minimum air temperature, (c) daily maximum LST, and (d) daily minimum LST.

The results of Malinovka basin indicate that $b_2$ had a strong linear relationship with the elevation difference for both daily maximum and minimum LST data, while for the $b_1$, the linear relationship was insignificant for daily maximum LST data. The slopes of $b_2$ (Figure 1c and 1d) were −18.6°C/km and −24.8°C/km for daily maximum and minimum LST data, respectively. Different slopes were observed for $b_1$ (Figure 4.17(a) and 4.17(b)), −7.1°C/km and −30.8°C/km. For the interpolation station, these values are 4.4°C/km and −5.8°C/km (Figure 4.15). It is clear that the differences between the slope of $b_2$ and the interpolation station are larger compared the slope of IDWEle method and TLAPS of interpolation stations, and according to the results shown in the Figure 4.13(b), it seems the differences can explain the larger error of LST data compared with the IDWELE method. However, it is common that the temperature is decreasing with the elevation. Though in the Fareast Area, the inversion layer can be detected (Fan *et al*., 2002) during the winter and spring, according to basic information of the Malinovka basin, the elevation variations are large and its slope are relatively large (Table 2.1). Thus, the areal accuracy that effected by different creation methods are still need to be considered in the Malinovka basin.
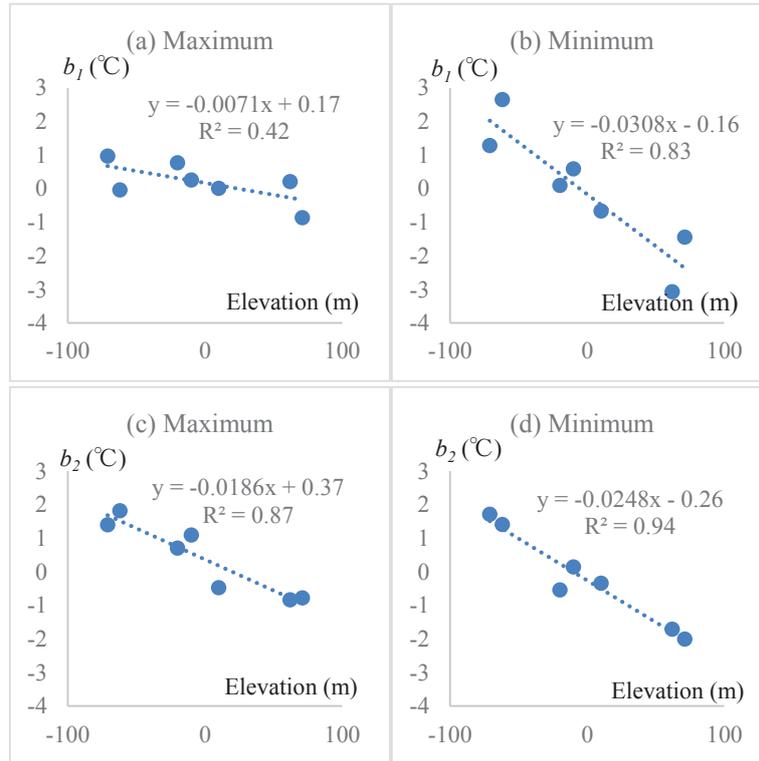
Figure 4.17 Regression analysis between the temperature differences ($b1$ and $b2$) and the elevation difference in the LST- LST method of Malinovka basin, (a) daily maximum air temperature, (b) daily minimum air temperature, (c) daily maximum LST, and (d) daily minimum LST.

The results of Gari basin indicate that $b_1$ had no strong linear relationship with the elevation difference for both daily maximum and minimum data. And $b_2$ only had linear relationship in the daily minimum data, and increasing trend is detected LST data. The slopes of $b_1$ (Figure 4.18(a) and Figure 4.18(b)) were -1.1°C/km and –5.9°C/km for daily maximum and minimum LST data, respectively. The slopes of $b_2$ (Figure 4.18(c) and Figure 4.18(d)) were 6.0°C/km and –17.6°C/km for daily maximum and minimum LST data, respectively. And these values are -10.6°C/km and –10.8°C/km for interpolation stations (Figure 4.15).

According to the results of *RMSE* shown in the Figure 4.13 (c), the results of the IDWEle method is better than the LST-LST method for daily maximum data. Considering the decreasing trend of the interpolation stations and the IDWEle method,

73

the increasing trend of $b_2$ for LST-LST method may lead to more errors in this basins.
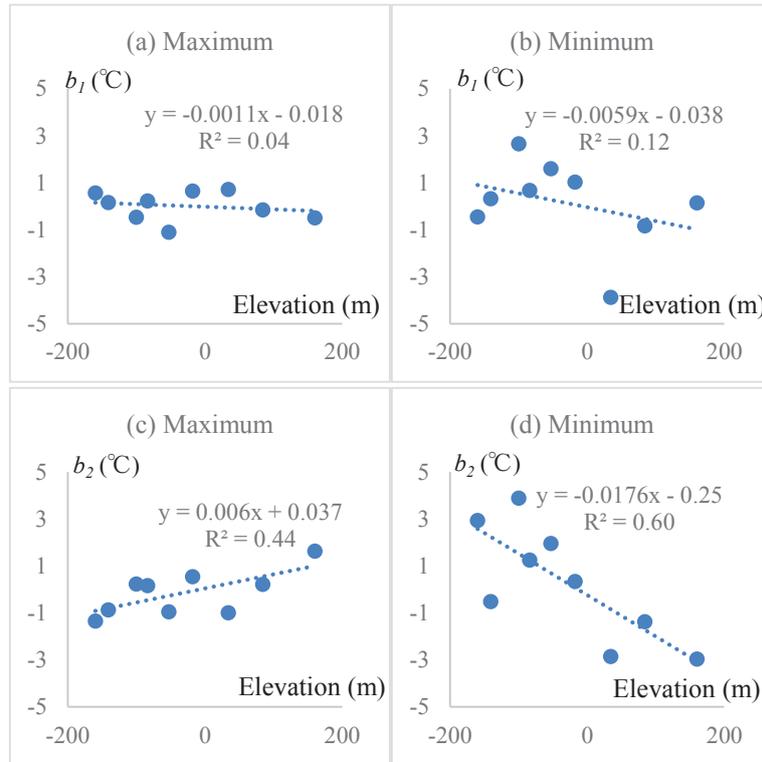


Figure 4.18 Regression analysis between the temperature differences ($b1$ and $b2$) and the elevation difference in the LST-LST method of Gari basin, (a) daily maximum air temperature, (b) daily minimum air temperature, (c) daily maximum LST, and (d) daily minimum LST.

Generally, the results indicate that the estimation of the elevation effects may influence the accuracy of LST-LST method, especially for the test basins of which the validation stations and the interpolation stations has low elevations.

Further, Westermann *et al*. (2011) once presented that MODIS LST data exits the high difference between dry area and wet area during snowmelt season in the tundra of Svalbard, Norway. Hachem *et al*. (2012) also concluded that the stagnant surface water which can modify the heat exchanges between the ground surface and the atmosphere leads to the difference between the air temperature and LST in Alaska. Moreover, due to the variations of land covers, mixed pixels containing different objects of satellite data can finally lead to relatively rough accuracy for the LST calculation compared with the

74

observed station that only located in one point (Liu *et al.,* 2006; Yang *et al.,* 2011). The "pixel" of the MODIS LST data is at least 1km$^2$, which is a comprehensive value compared with the observed data in the point. Thus, this is another probable reason for the errors of the estimated results of MODIS data. However, in this study, limited to the observed stations in these study regions, the influence of water content and mixed pixel effects cannot be detected at now.

In addition, the statistical analysis method always need abundant training data to generate stable results (Zakšek and Schroedter-Homscheidt 2009), the amount of data are less than 2,000 for 2002-2010 of all stations (Table 4.6). The short observation period of the MODIS LST data is an important factor that induces the larger estimated errors of this method. It is also noteworthy that the cloud and hazes are treated by not all exclude during the generation of LST data (Wan 2008; Vancutsem *et al*., 2010), the quality of the LST data also has influence on the final estimated results.

Though the errors are extended by the LST-LST method, the *RMSEs* obtained herein were still within a reasonable range compared with earlier research in different regions. For example, based on the multi-steps linear regression method, Colombi *et al.* (2007) predicted air temperature in alpine areas of Italy. Their *RMSE* = 1.89°C for daily average temperature and maximum and minimum air temperature were 2.47°C and 3.36°C, respectively, similar to our results. Shen *et al.* (2011) used a single linear regression method to predict daily maximum and minimum air temperature based on LST data, with error 2–3℃.

Generally, for the IDW method it shows the best results of all the different methods, the IDWEle method obtained the relatively worse results. However, according to the different topographic characters of the test basins; this study still employed the IDWEle

method to test the estimated data for driving the snowmelt model. Compared with the IDW method, though the errors are enlarged in LST-LST method, the two indices ($R^2$ and $RMSE$) still show that errors of the LST−LST method were within a reasonable range, and overall performance of the method demonstrated its capability for predicting daily maximum and minimum temperature. And the analysis results of the temperature lapse rate for both IDWEle method and LST-LST method indicate that the worse results of the IDWEle method and LST-LST method is highly probable induced from the estimation of topographic effects on air temperature at the low elevation stations.

### 4.2.4. The created air temperature in the test basins

According to the percentage of area of each subbasin and its created air temperature, this study calculated the value of the daily maximum and daily minimum temperature for each scenarios. The results are shown in Figure 4.19.

In the Apkoroshi basin, for the daily maximum data, the T-IDWEle scenarios obtained the lowest temperature (5.78°C), followed by T-LST (7.46°C), T-Observed (8.64°C) and T-IDW (9.46°C). The order is T-IDWEle (-6.33°C), T-Observed (-4.87°C), T-LST (-4.02°C) and T-IDW (-3.74°C) for the daily minimum data.

In the Malinovka basin, for the daily maximum data, the T-LST scenarios obtained the lowest temperature (7.33°C), followed by T-Observed (7.87°C), T-IDW (8.79°C) and T-IDWEle (12.56°C). The order is T-IDWEle (-7.01°C), T-Observed (-4.05°C), T-IDW (-3.47°C) and T-LST (-1.68°C) for the daily minimum data.

In the Gari basin, for the daily maximum data, the T-IDWEle scenarios obtained the lowest temperature (7.33°C), followed by T-LST (7.87°C), T-IDW (11.26°C) and T-Observed (11.26°C). The order is T-KST (-3.83°C), T-IDWEle (-2.61°C), T-IDW (-2.06°C) and T-Observed (-1.99°C) for the daily minimum data.

76
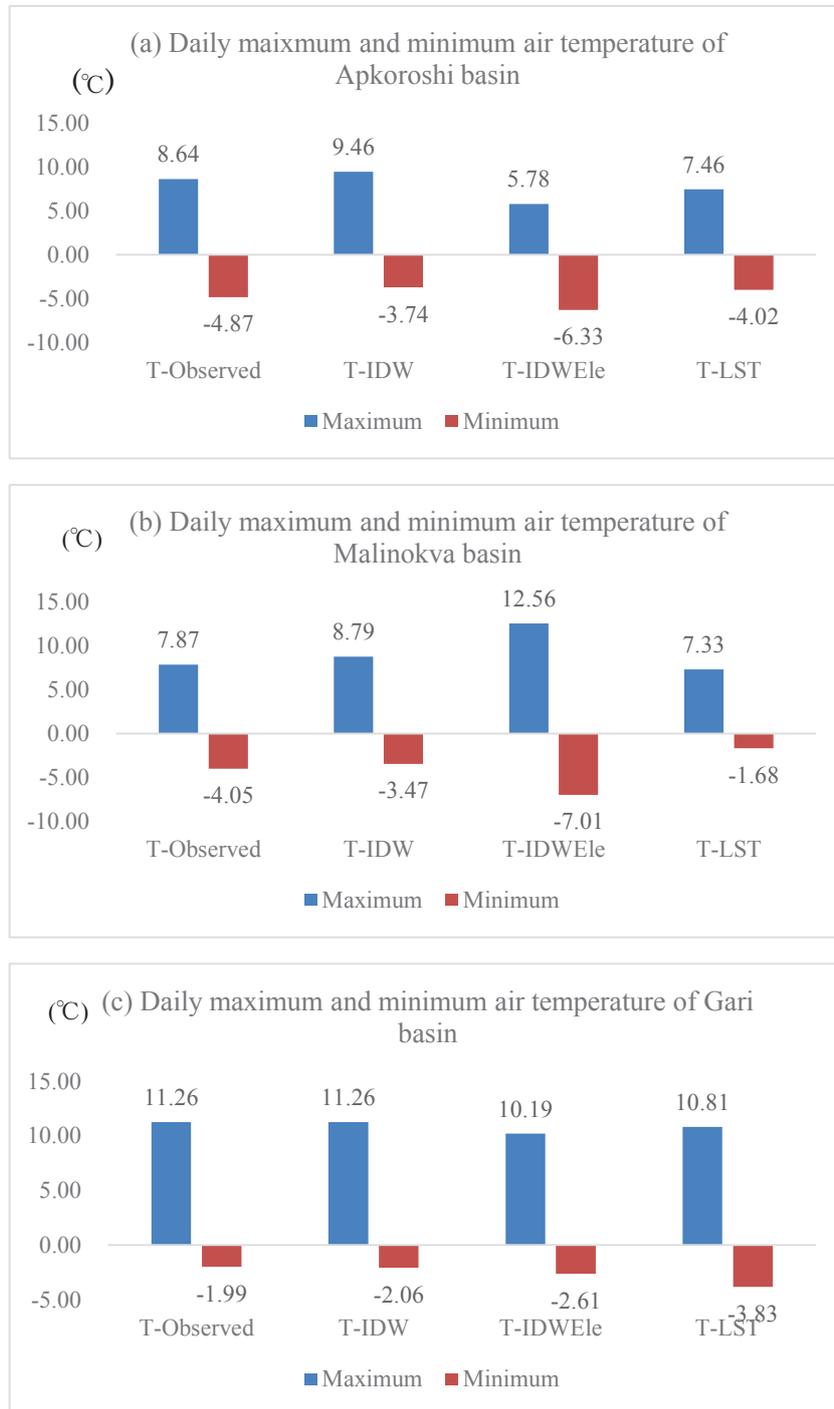
Figure 4.19 Daily maximum and minim air temperature of created data in different test basins, (a) Apkoroshi basin, (b) Malinovka basin, and (c) Gari basin.

In addition, according to the geo-information listed in Table 2.1, it is clearly that the elevation are different in each test basins, and considering the elevation influence on

the air temperature, this study also analyzed the linear relationship between air temperature and the elevation in each test basin

The relationship between the created air temperature and elevation of Apkoroshi basin are shown in Figure 4.20 and Figure 4.21. It is clearly that for the daily maximum data of the T-IDWEle scenario and T-LST scenario shows a significant linear relationship with the elevation, and the TLAPS of these two scenarios are very similar, the values are -4.8°C/km and -4.6°C/km, respectively. For the T-Observed scenario and T-IDW scenario, the results indicate they has no significant relationship with elevations.

For the daily minimum data only the T-IDWEle method gave a significant relationship with the elevation, the TLAPS is -2.7°C/km, and for the T-LST data a positive TLAPS is detected though the relationship is insignificant. The T-Observed and T-IDW method presented no significant relationship with the elevation.
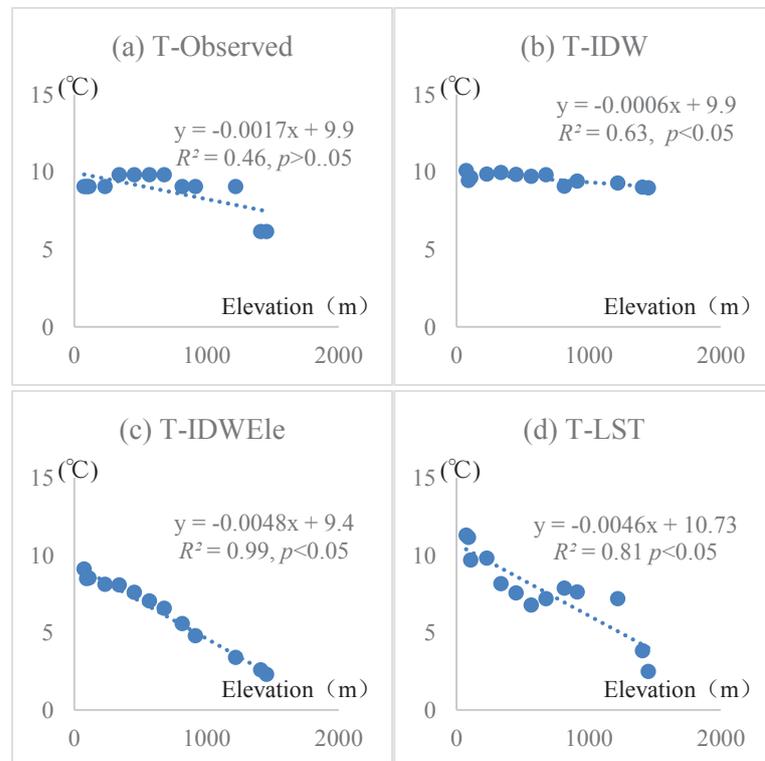


Figure 4.20 The regression analysis between created daily maximum air temperature and elevation

**(a) T-Observed**

(℃)

y = -0.0015x - 3.39
R² = 0.10, p>0.05

Elevation（m）

**(b) T-IDW**

(℃)

y = 7E-06x - 3.63
R² = 3E-05, p>0.05

Elevation（m）

**(c) T-IDWEle**

(℃)

y = -0.0027x - 4.28
R² = 0.80, p>0.05

Elevation（m）

**(d) T-LST**

(℃)

y = 0.0008x - 4.26
R² = 0.3, p>0.0.5

Elevation（m）

Figure 4.21 The regression analysis between created daily minimum air temperature and elevation in Apkoroshi basin, (a) T-Observed, (b) T-IDW, (c) T-IDWEle, and (d) T-LST.

The relationship between the created air temperature and elevation of Malinovka basin are shown in Figure 4.22 and Figure 4.23. It is clearly that for the daily maximum data the T-IDWEle data and T-LST data presented a significant linear relationship with the elevation, for the T-Observed and T-IDW scenarios the results indicate they has no significant relationship with elevations. And it is also interesting that the T-IDWEle data presented an increasing trend with high value of TLAPS (12.3°C/km), and for the T-LST data the negative TLAPS is detected (-7.0°C/km).

For the daily minimum air temperature only the T-IDWEle method gave a significant relationship with the elevation, and the TALPS is -13.2°C/km. For the T-LST data a positive TLAPS is detected though the relationship is insignificant as same as the Apkoroshi basin, and the TLAPS is 5.8°C/km. They the T-IDW presented a significant

relationship with the elevation, the TALPS is only -0.8°C/km which indicate that with the elevation increasing the decrease of temperature is small.
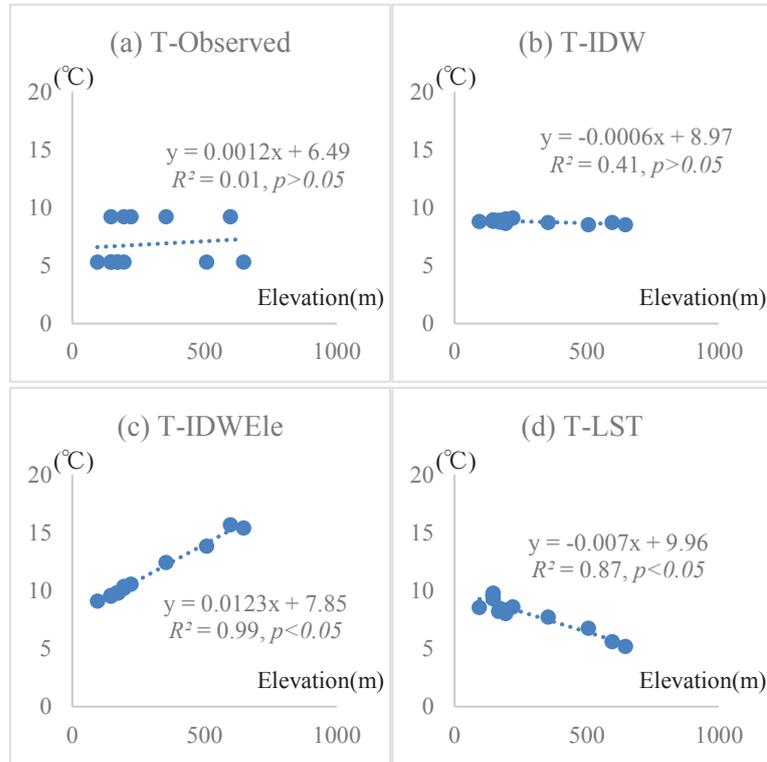


Figure 4.22 The regression analysis between created daily maximum air temperature and elevation in Malinovka basin, (a) T-Observed, (b) T-IDW, (c) T-IDWEle, and (d) T-LST.
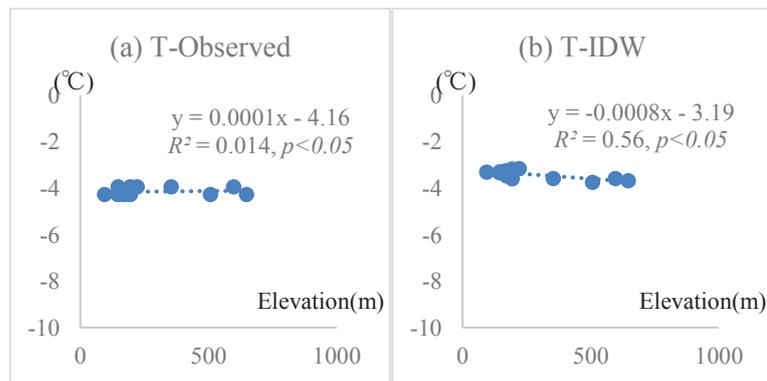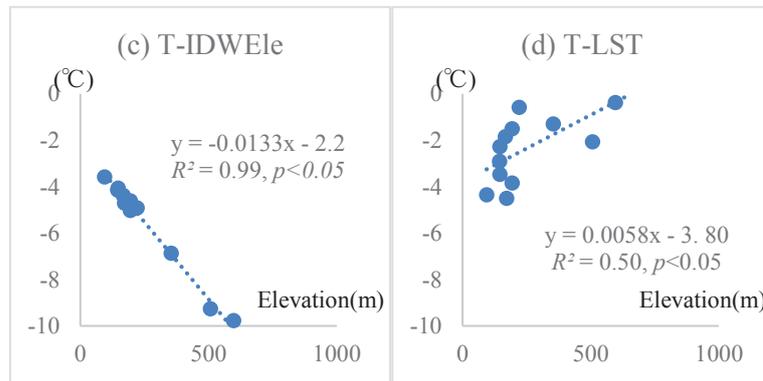
Figure 4.23 The regression analysis between created daily minimum air temperature and elevation in Malinovka basin, (a) T-Observed, (b) T-IDW, (c) T-IDWEle, and (d) T-LST.

The relationship between elevation and the created data of Gari basin are shown in Figure 4.24 and Figure 4.25. Only the T-IDWEle data presented a significant linear relationship the elevation, and the TLAPS is -14.6°C/km, and for the T-LST data though the relationship is insignificant the TLAPS is very large compared other scenarios and the T-LST scenarios of Apkoroshi basin and Malinovka basin (-26.5°C/km).

For the daily minimum air temperature only the T-IDWEle method shows a significant relationship with the elevation, for the T-LST data a positive TLAPS is detected though the relationship is insignificant the TLAPS is 4.0°C/km.
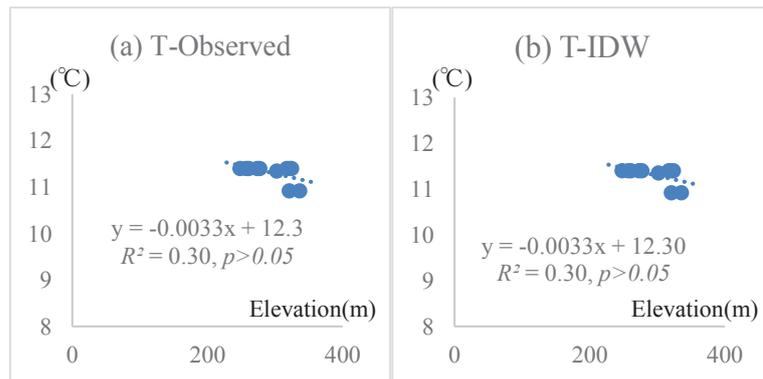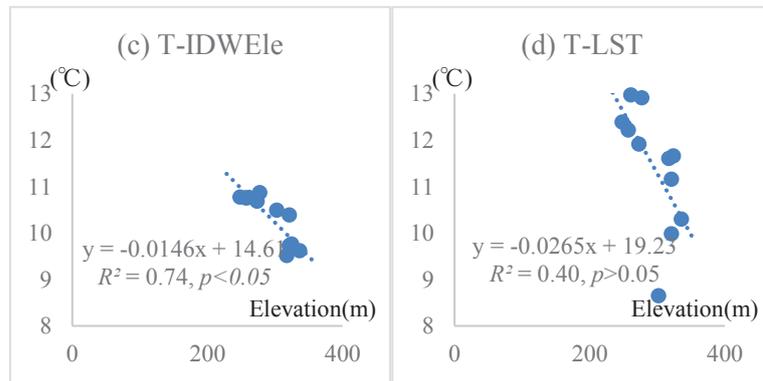
Figure 4.24 The regression analysis between created daily maximum air temperature and elevation in subbasins of Apkoroshi basin, (a) T-Observed, (b) T-IDW scenario, (c) T-IDWEle, and (d) T-LST.
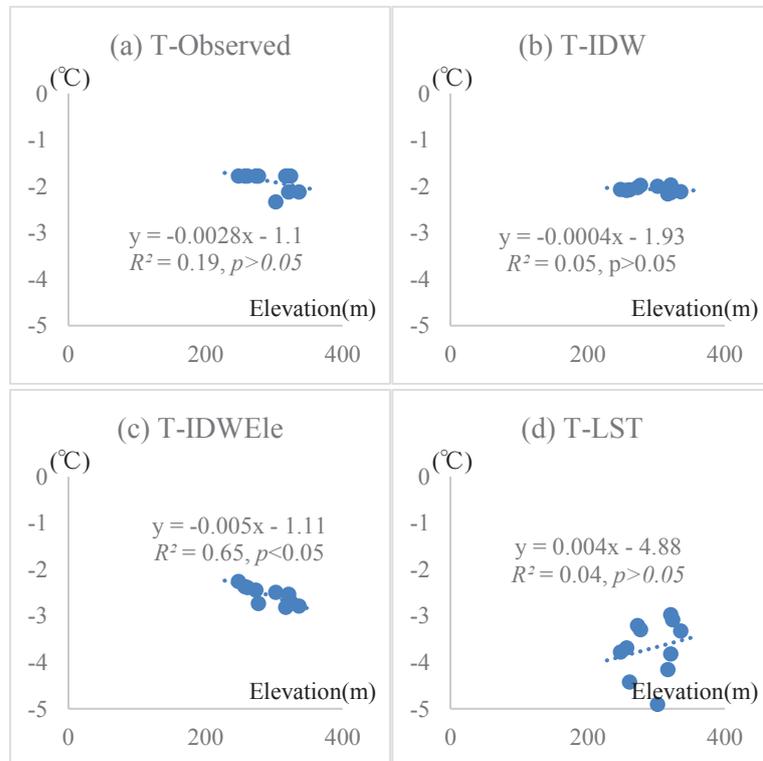


Figure 4.25 The regression analysis between created daily minimum air temperature and elevation in subbasins of Apkoroshi basin, (a) T-Observed, (b) T-IDW, (c) T-IDWEle, and (d) T-LST.

For the T-IDW data and T-IDWEle data, this study also presented the linear regression analysis between the air temperature of interpolation stations and elevations in Figure 4.26. And the elevation information of the interpolation stations and the geo center point of subbasins are shown in Figure 4.27.

Considering the results of T-IDWEle scenario which are shown in the Figure

4.21~4.26, it is clear that the relationships between the elevation and air temperature of the subbasins are influenced by those relationships of interpolation stations. However, for the IDW method, because there is no topographic corrections, thus, its created values are only rely on the low elevation stations, and this is reason for its higher values of created data compared with the T-IDWEle method. Moreover, because the T-Observed data also used the low elevation stations for data creation, thus its value is also higher than T-IDWEle data, except for the daily maximum data in the Malinovka basin. According to the results shown in Figure 4.26 (c), it is clear that the positive TLAPS of the T-IDWEle data (Figure 4.22 (c)) is inducted by the interpolation stations, however it is well known that the TLAPS should be negative value, thus for the Malinovka basin, though the linear relationship between elevation and created maximum air temperature is high, this may include errors to the data creation.
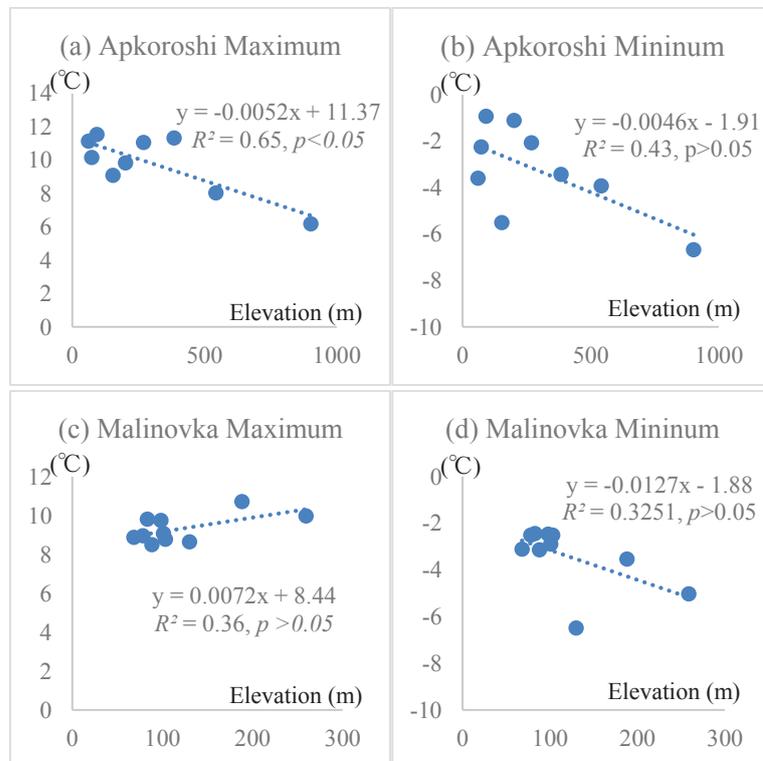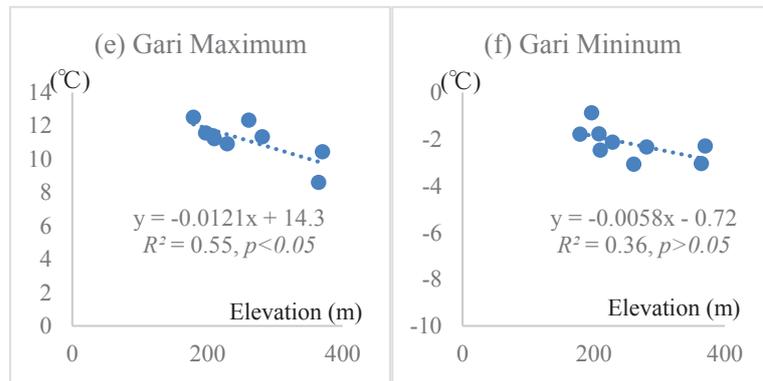
Figure 4.26 The linear regression analysis between air temperature and elevations of the interpolation stations in all test basins.
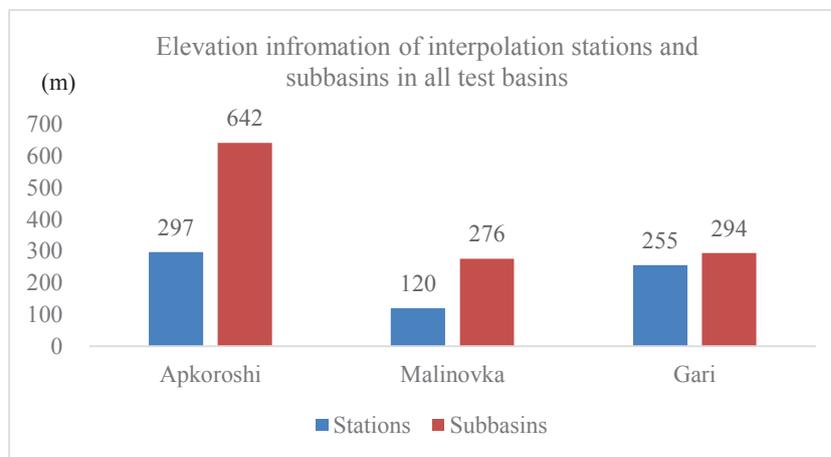


Figure 4.27 The elevation information of the interpolation stations and the subbasin in all test basins.

We also conducted the linear regression analysis for the $b_2$ of LST-LST methods and its relationship with the elevation, the results are shown in Figure 4.28.

It is clear that the linear regression analysis results of $b_2$ and the elevation are consistent with the results of created data in Apkoroshi basin and Malinovka basin, However, the trend is different for daily maximum data in the Gari basin, the TLAPS of Figure 4.28 (e) is positive while in the Figure 4.24 (d), the TLAPS is negative. For all three basins, it is clear that the decreasing TLAPS can be detected for the daily maximum air temperature. Considering the elevation variation in these basins, the relationship between daily maximum air temperature and elevation are

reasonable for T-LST scenario in these two basins.
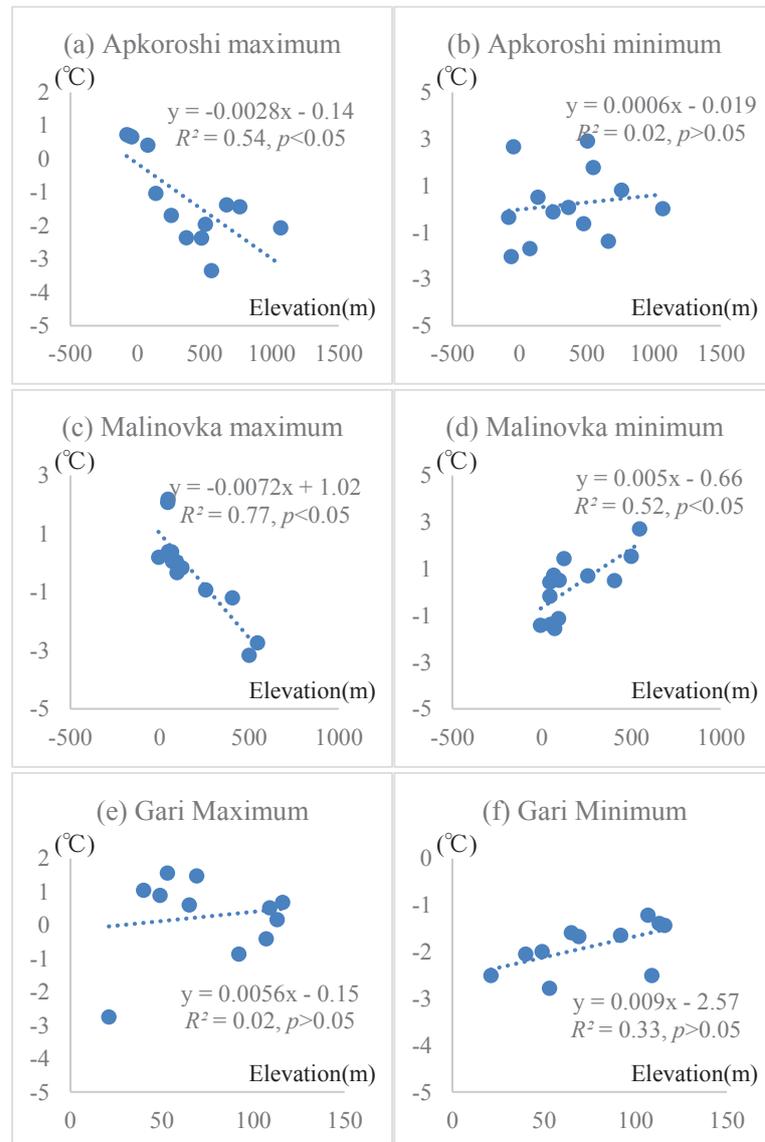


Figure 4.28 The linear regression analysis between air temperature and elevations of the interpolation stations in all test basins.

Generally, for the daily maximum air temperature, the T-LST scenario seems can present a relative stable relationship with elevation compared with the other data, while for the daily minimum data it is difficult to conclude which data is more reasonable. Because there is no air temperature data in the test basins, thus the accuracy of the different created data is need further discussion by analysis the snowmelt simulation results in the next section.

## 4.3. Evaluations of different estimated air temperature data by SWAT model

### 4.3.1. Performances of snowmelt simulation

Based on the settings in section 3.4, the SWAT model for different air temperature input data were operated in the test basins.

The results of $R^2$ and $NSE$ for Apkoroshi basin is shown in Figure 4.29 and the hydrograph of the best simulation is shown in the Figure 4.30. The T-IDW ($R^2$=0.66, $NSE$=0.66) and T-IDWEle ($R^2$=0.64, $NSE$=0.59) scenarios obtained the relative worse results compared with the T-Observed scenario ($R^2$=0.72, $NSE$=0.66) and the T-LST scenario ($R^2$=0.75, $NSE$=0.73). The best results are obtained by the T-LST scenario.
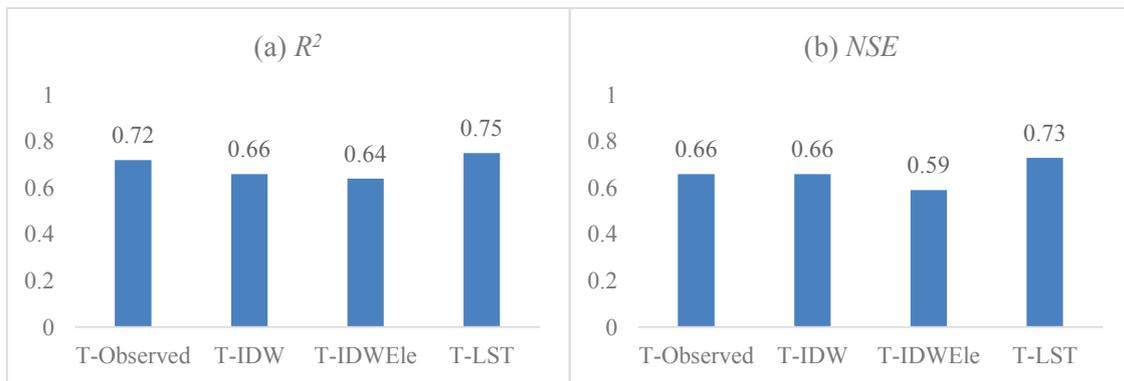


Figure 4.29 The results of $R^2$ and $NSE$ in Apkoroshi basin, (a) $R^2$, and (b) $NSE$.
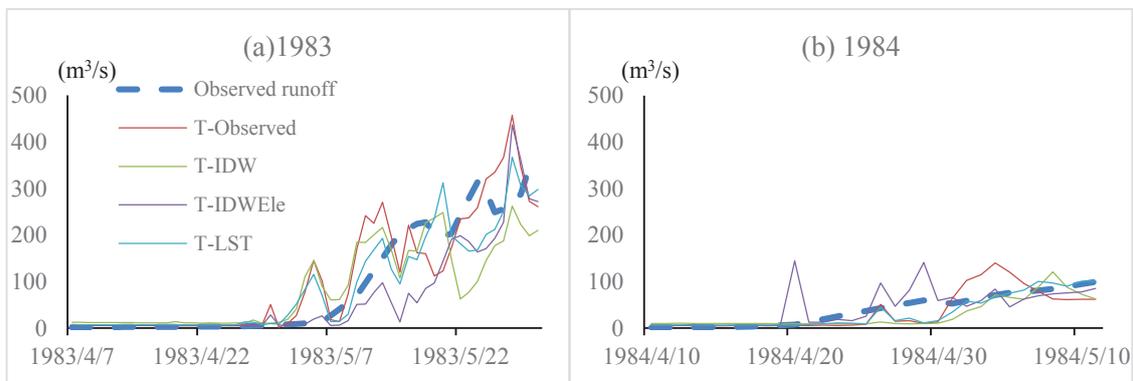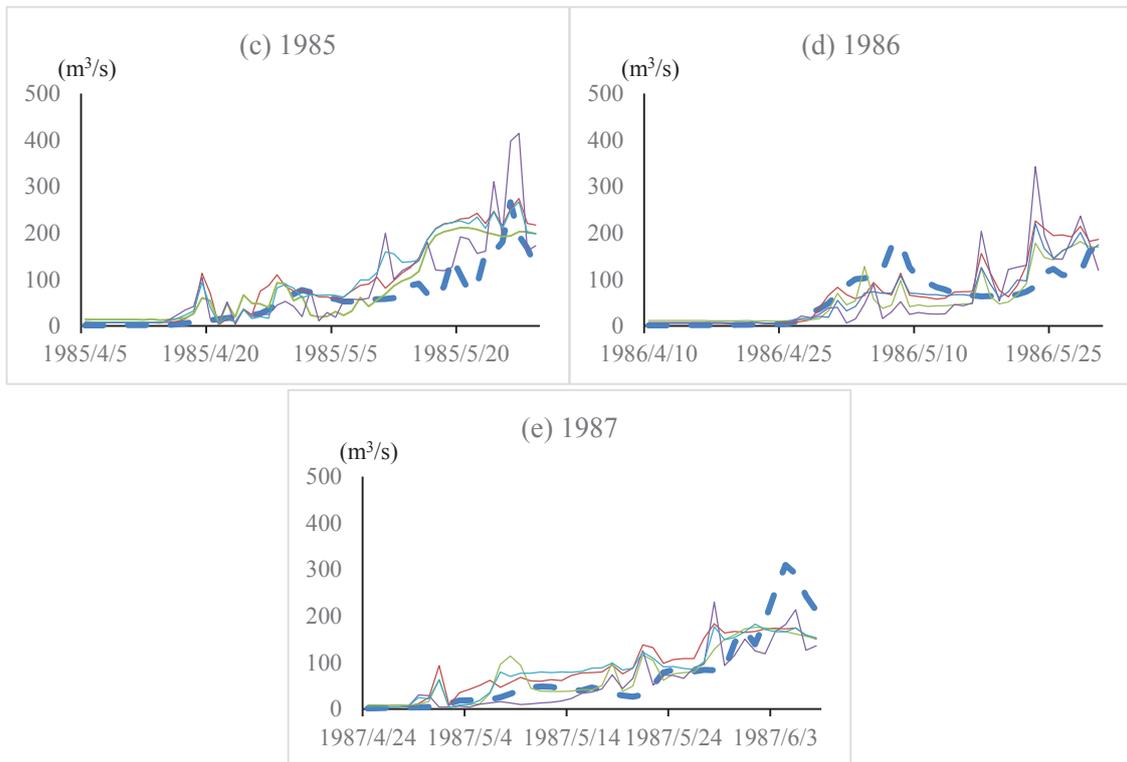
Figure 4.30 The best simulation hygdrographs for different air temperature in Apkoroshi basins from 1983 to 1987, (a) 1983, (b) 1984, (c) 1985, (d) 1986 and (e) 1987.

In this study, the daily absolute error and *MAE* were calculated and *ANOVA* analysis conducted for each period, the results are shown in Figure 4.31. For all period, the ANOVA analysis indicate that there is no significant difference between the different created air temperature data. According to the results of *MAE*s, the T-LST data presented smaller result for entire period and mainly melting period than the early period compared with other data, the T-IDWEle data obtained the largest error for main period and entire period, and however, the errors are smallest during the early melting period. The T-Observed data and T-IDW data obtained similar results and the results are larger than T-LST data for all periods.
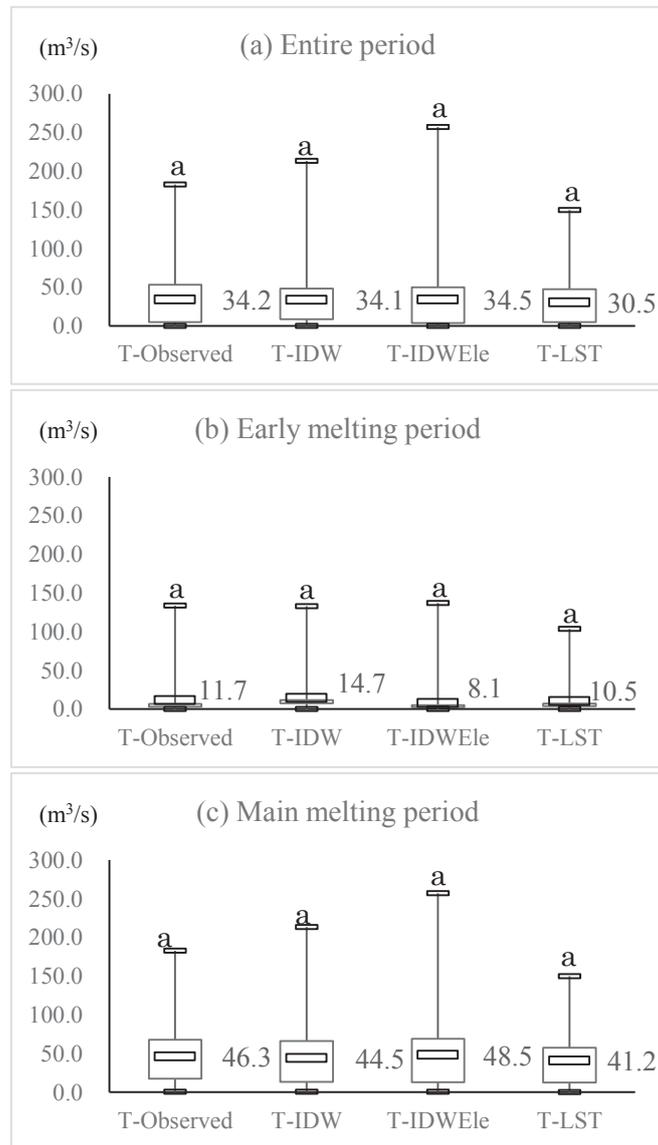
Figure 4.31 Absolute error of simulated discharges for different air temperature data in Apkoroshi basin, (a) entire period, (b) early melting period, (c) main melting period.

The results of $R^2$ and $NSE$ for Malinovka basin is shown in Figure 4.32 and the hydrograph of the best simulation is shown in the Figure 4.33. The T-IDW ($R^2$=0.66, $NSE$=0.63) and T-IDWEle ($R^2$=0.63, $NSE$=0.53) scenarios obtained the relative worse results compared with the T-Observed scenario ($R^2$=0.68, $NSE$=0.67) and the T-LST scenario ($R^2$=0.80, $NSE$=0.79). The best results are obtained by the T-LST method, while the worst results is obtained in the T-IDWEle scenario.
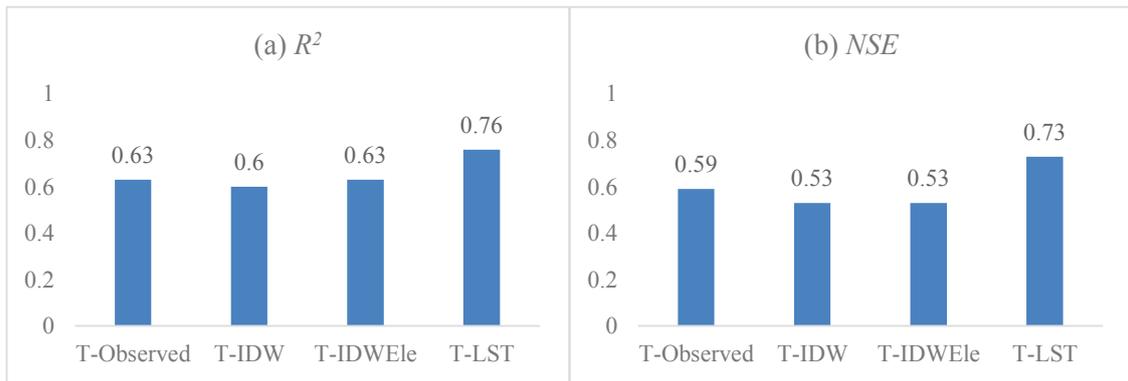
Figure 4.32 The results of $R^2$ and $NSE$ o in Malinovka basin, (a) $R^2$ and (b) $NSE$.
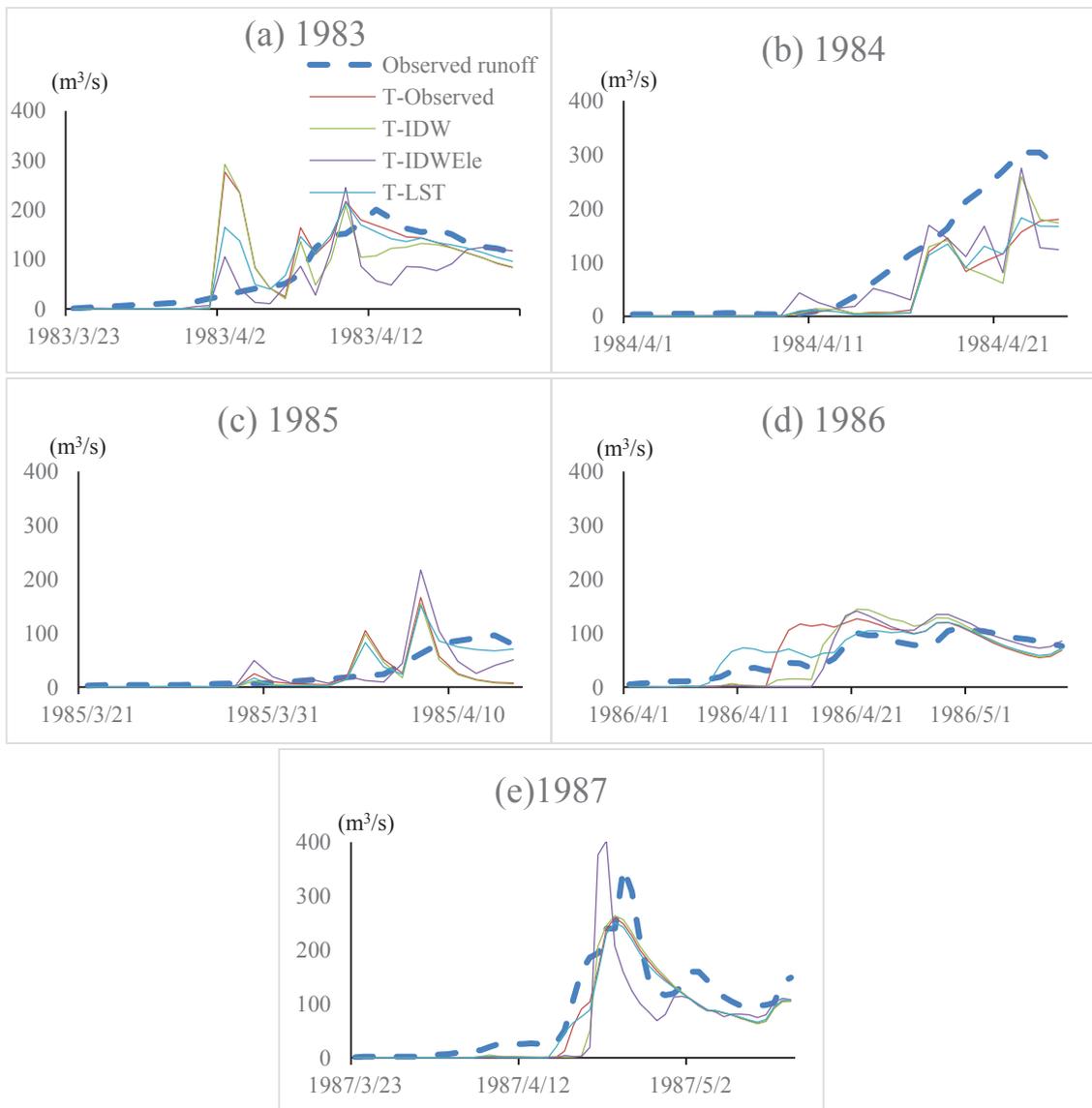


Figure 4.33 The best simulation hygdrograph for different air temperature in Malinvoka basin in different years from 1983 to 1987, (a) 1983, (b) 1984, (c) 1985, (d) 1986 and (e) 1987.

In the Malinovka basin, the daily absolute error and *MAE* were calculated and

*ANOVA* analysis conducted for each period. Except the main melting period ($p$=0.002), the results were insignificant ($p$>0.05). The Steel-Dwass test was calculated for the middle period (Figure 4.34). The $p$ values between the T-LST case and the T-IDW, T-IDWEle cases were less than 0.05, respectively, which indicate that the T-LST case can produce less simulation errors than T-IDW and T-IDWEle. Though the differences between the T-LST case and T-Observed case were insignificant, the results still indicate that the *MAE*s of T-LST case were smaller than T-Observed in every period.
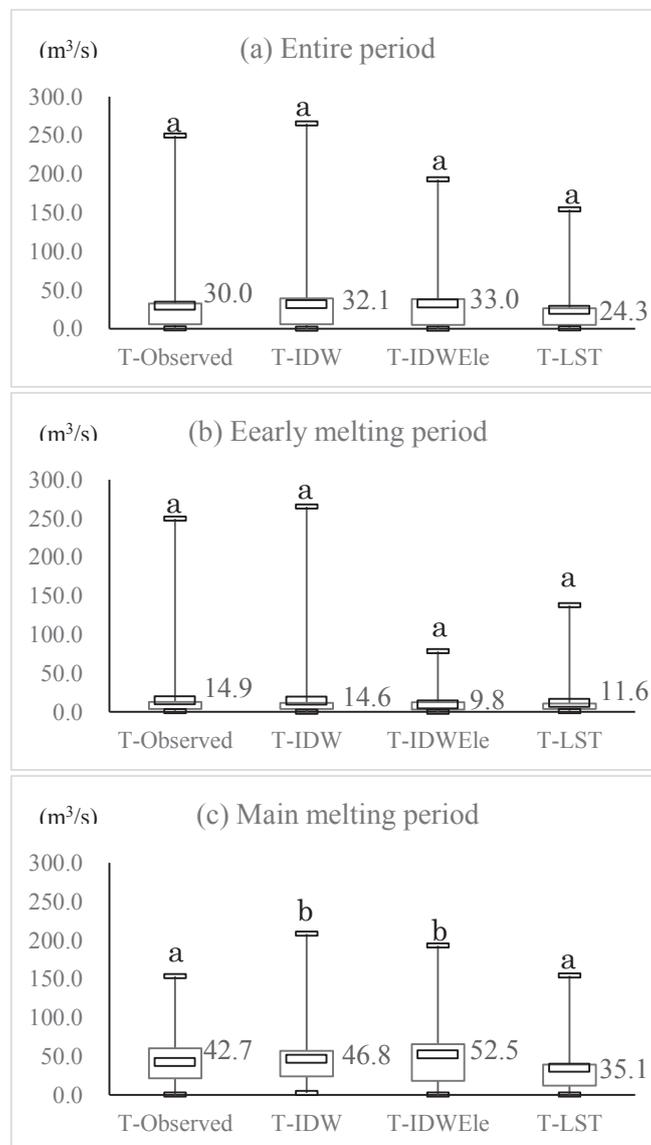


Figure 4.34 Absolute error of simulated discharges for different air temperature data in Malinovka basin, (a) entire period, (b) early melting period, (c) main melting period.

The results of $R^2$ and $NSE$ of Gari basin is shown in Figure 4.35 and the hydrograph of the best simulation is shown in the Figure 4.36. The performances of the $NSE$ are similar in Gari basin compared with other two basins. The T-Observed ($R^2$=0.66, $NSE$=0.66) and T-IDW ($R^2$=0.68, $NSE$=0.66) scenarios obtained the relative worse results compared with the T-IDWEle scenario ($R^2$=0.72, $NSE$=0.70) and the T-LST scenario ($R^2$=0.70, $NSE$=0.66). The best results are obtained by the T-IDWEle method, while other method performed the similar results.
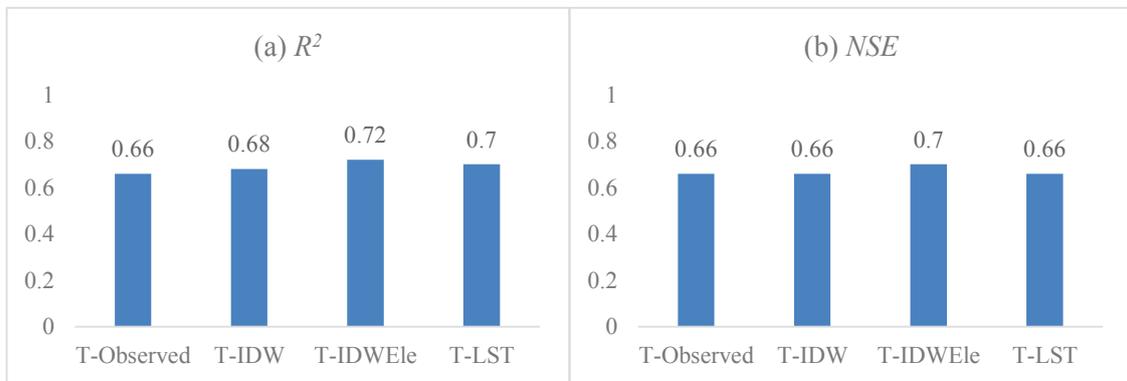


Figure 4.35 The results of $R^2$ and $NSE$ in Gari basin, (a) $R^2$ and (b) $NSE$.
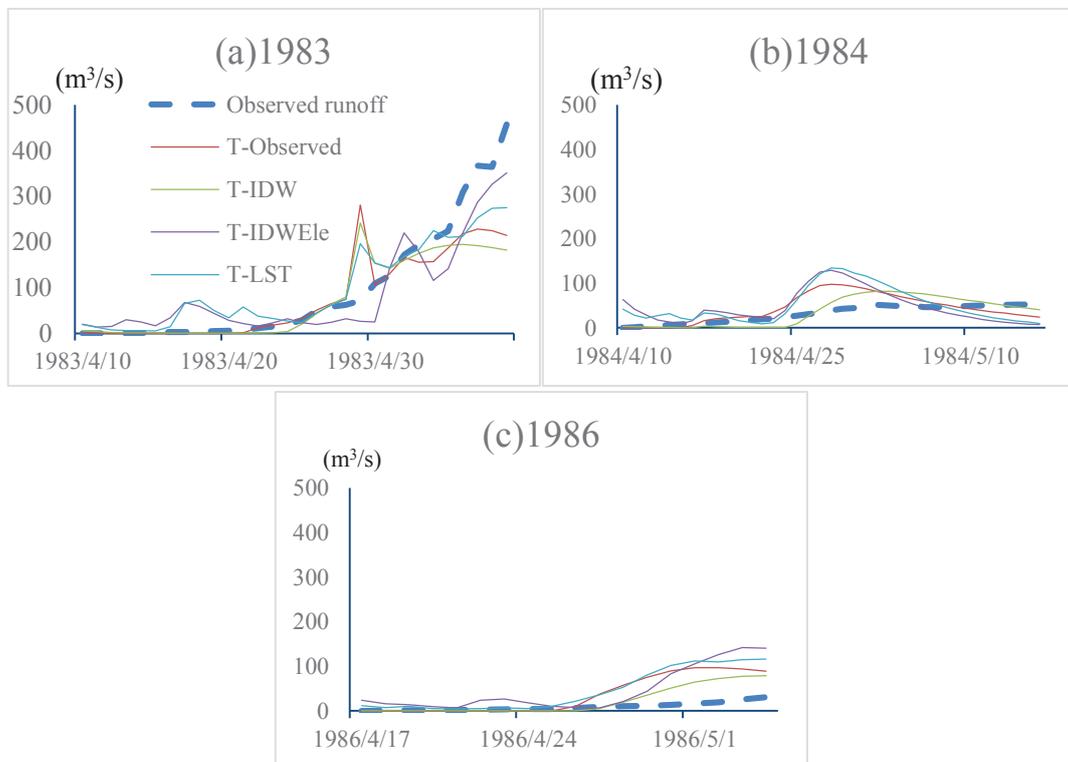


Figure 4.36 The best simulation hygdrograph for different air temperature in Gari basin in different years

In the Gari basin, the daily absolute error and *MAE* were calculated and *ANOVA* analysis conducted for each period. Except the early melting period ($p<0.001$), the results were insignificant ($p>0.05$). The Steel-Dwass test was calculated for the early melting period (Figure 4.37). The *p* values between the T-LST case and the T-IDW, T-IDWEle cases were less than 0.05, respectively, which indicate that the T-LST and T-IDWEle cases produced more simulation errors than T-IDW and T-IDWEle. And for the main melting period, the same trend is also detected.
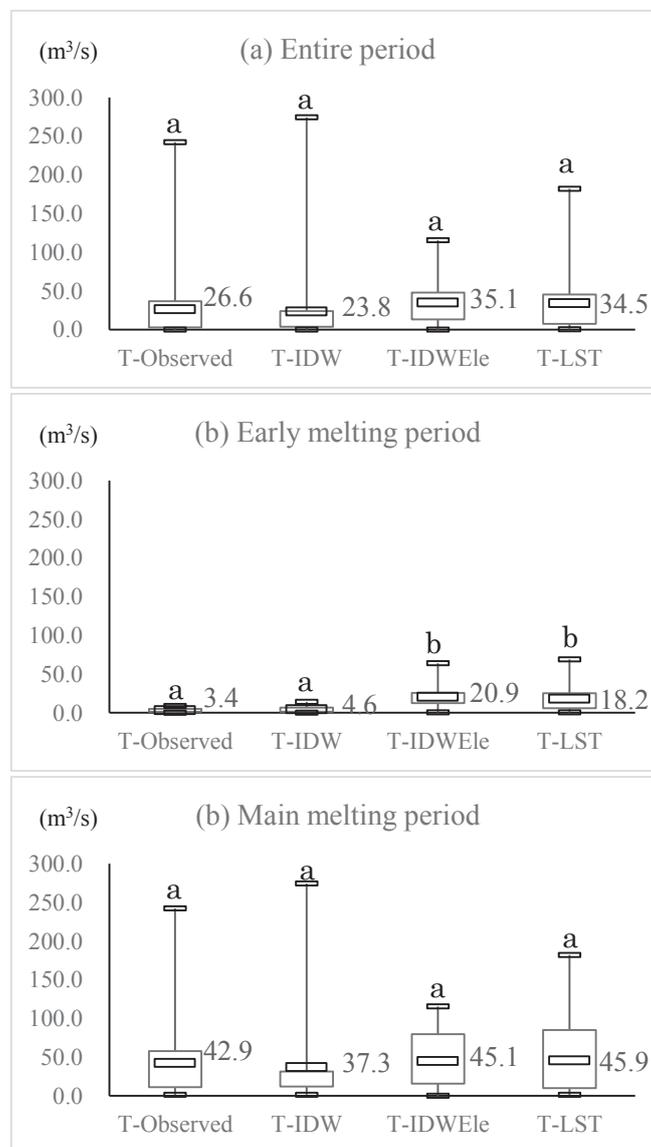


Figure 4.37 Absolute error of simulated discharges for different air temperature data in Gari basin, (a)

entire period, (b) early melting period, (c) main melting period.

According to the simulation hydrographs (Figures 4.30, 4.33, and 4.36), it is clear that the starting melt point cannot be accurately reflected by all the data. The melting processes occurred earlier for the temperature index data. For example, in Apkoroshi basin, in the year 1983, the simulated snowmelt hydrograph exist earlier peaks than the monitoring data. In the Gari basin, during the period of 1984 and 1986, the first peaks are occurred earlier in simulated hydrograph than the monitoring data. Moreover, the hydrographs of the estimated data in the Apkoroshi basin and Gari basin are more fluctuant compared with the observed runoff data.

It is clear the snowmelt is also influenced by other hydrological processes such as percolation and refreezing, especially at the beginning melt period. (Senese $et\ al.$, 2014). The results of Ahl $et\ al.$ (2006) indicated that the melt water in the river channel is mainly contributed by the lateral flow rather than overland flow in the snow-dominated area of Rocky Mountain. Quinton $et\ al.$ (2004) also reported the variation of the soil conductivity during the snowmelt period will influence the subsurface and surface flow.

Aside from meteorological influences, land cover is another influence on snowmelt. Wetlands are believed to be a major factor in snowmelt simulation, because of its capacity for water storage (Hayashi $et\ al.,$ 2004). Hydrological properties of these wetlands are very sensitive to variations of $T_a$, seasonal precipitation, and other climatic factors (Fang and Pomeroy, 2008). To simulate snowmelt processes in wetlands, Fang $et\ al.$ (2010) applied a physically based approach to a wetland-dominated prairie basin in Canada. They found that the ability of wetlands to trap blowing snow in winter and store runoff water is a crucial feature of the hydrology, and this poses a substantial challenge to hydrological modeling. Wang $et\ al.$ (2008) and Yang $et\ al.$ (2010)

93

demonstrated that the SWAT model's hydrological simulation component should be improved for wetland-dominated areas based on detailed wetland measurement data, such as annual water table depth and normal and maximum water storage capacities. Wetland covered over 25% of the entire basin Gari. However, in a data-sparse area, it is difficult to acquire enough data to conduct physically based snowmelt simulation of wetlands, and this is the likely reason that the NSE is weak in that basin. And also for other snowmelt parameters, according to the different basins, the variations of these parameters can be greatly influence by the basin characters such as land cover types, aspects (Horton 2003; Tekeli $et$ $al.$, 2005).

In addition, it is well-known that the simulation results of hydrological model are generated by both input data and model parameters. Except the temperature data, the precipitation data also directly decides the water amount and the significantly influence the simulation results. In our study, we employed the APHRODITE data as the input data. Though the applicability of this data in this regions was proved by other researchers (Gillies $et$ $al.,$ 2012; Onishi $et$ $al.$, 2010), considering the areas of the test basins (3,000-5,000 km$^2$) and the spatial resolution (0.25 degree) of the input rainfall data, the precipitation data may also lead to the errors of simulation results. In this study, the simulation result of Gari basin, during the year 1986, the simulated results are significantly higher than the observed runoff.

## 4.3.2. Uncertainty analysis
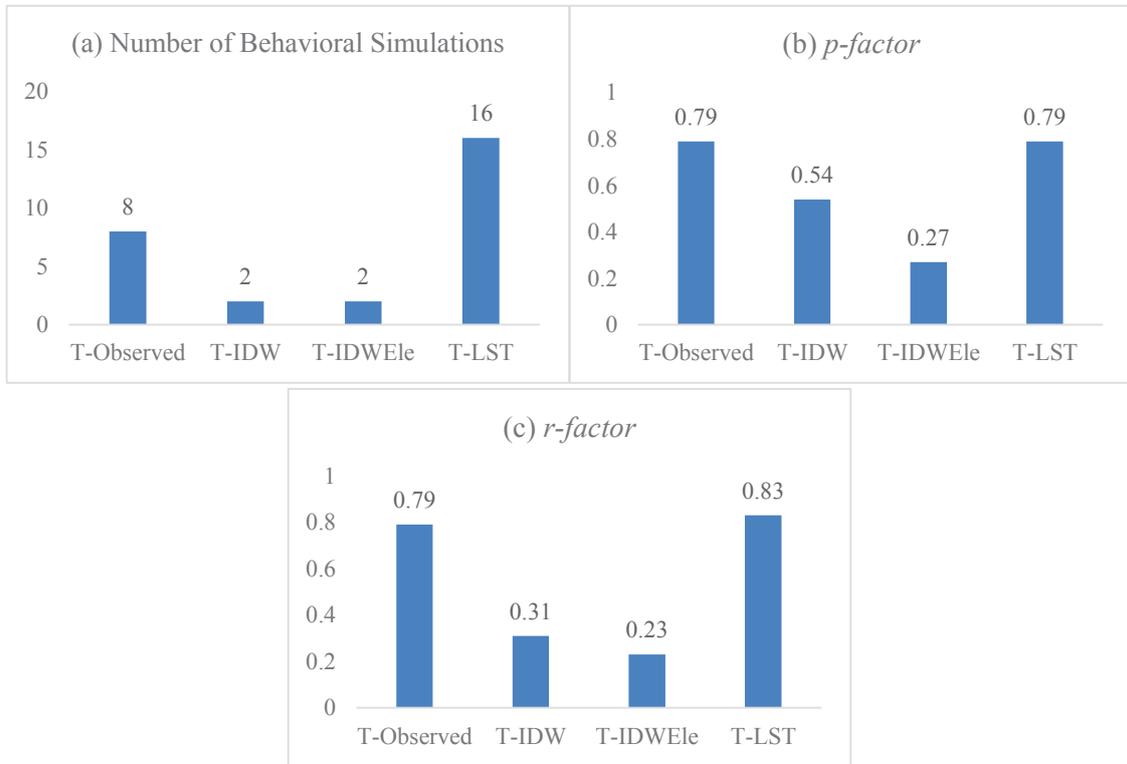


Figure 4.38 The uncertainty analysis results in Apkoroshi basin, (a) Number of behavioral simulations, (b) *p-factor* and (c) *r-factor*.

The uncertainty analysis results of Apkoroshi basin is shown in Figure 4.38. For the number of behavioral simulations, the T-LST data (16 simulations) obtained with the best results followed with T-Observed data (8 simulations), T-IDW data (2 simulations) and T-IDWEle data (2 simulations). The same orders can be detected for the *p-factor* and *r-factor*. For the *p-factor* the values are 0.79, 0.79, 0.54, and 0.27 for T-LST data, T-Observed data, T-IDW data, T-IDWEle data, respectively; and the values of *r-factor* are 0.83, 0.79, 0.31, and 0.23, respectively. The results clearly indicate that the different data creation methods has different performances in this test basin for model uncertainties, the T-LST obtained the relatively good results compared with the T-IDW and T-IDWEle scenarios. However, it is also noticed that the T-Observed data which only depend on the sparse observed data obtained the similar results for both model

simulation and uncertainties analysis in this test basin.
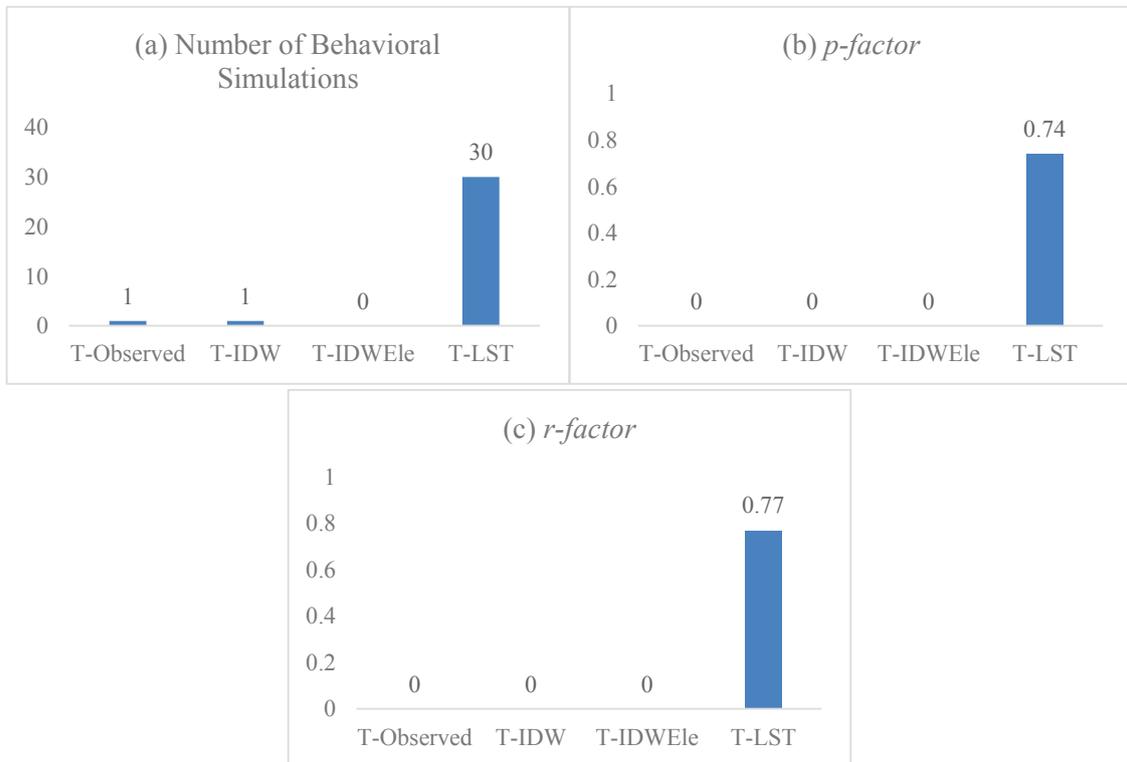


Figure 4.39 The uncertainty analysis results in Malinovka basin, (a) Number of behavioral simulations, (b) *p-factor* and (c) *r-factor*.

The uncertainty analysis results of Malinovka basin is shown in Figure 4.39. For the number of behavioral simulations, the T-LST data (30 simulations) obtained with the best results followed with T-Observed data (1 simulations), T-IDW data (1 simulations) and T-IDWEle data (0 simulations). The same orders can be detected for the *p-factor* and *r-factor*. For the *p*-factor, except T-LST scenario other scenarios are 0.00 and the values of *r-factor* is 0.77 for T-LST scenario. It is clear that judged by the uncertainly analysis results, the T-LST shows better performances compared the original observed and other interpolations methods.
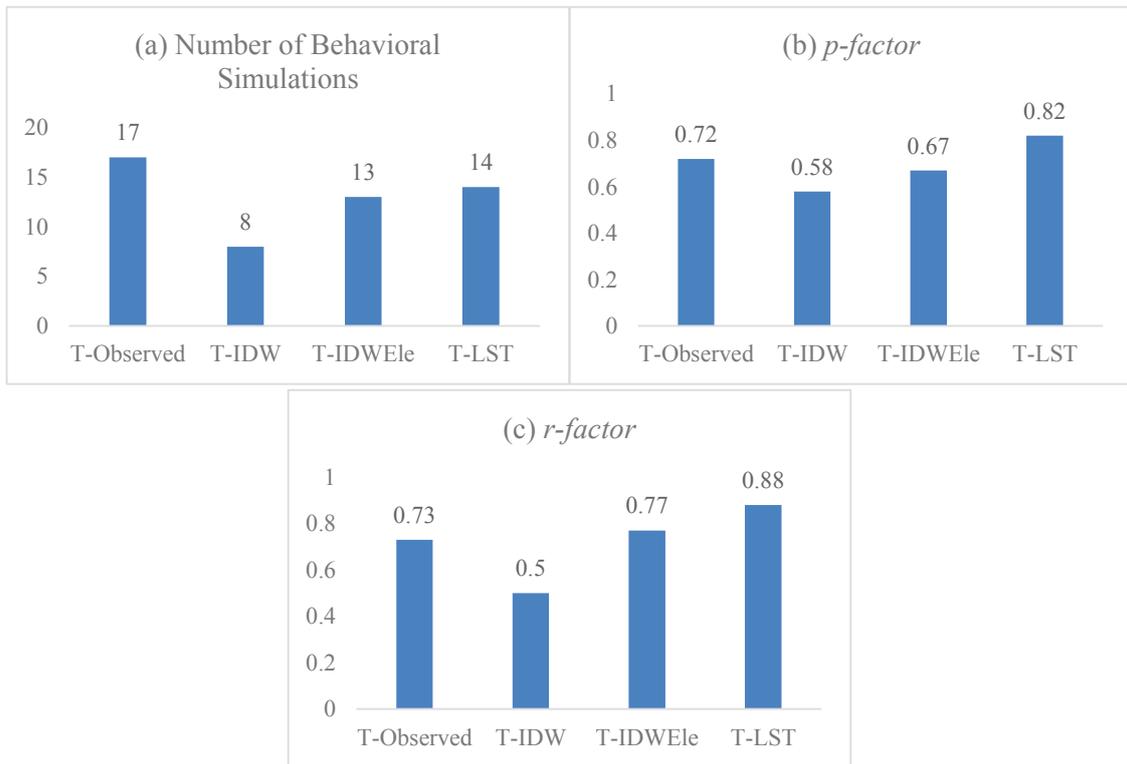
Figure 4.40 The uncertainty analysis results in Gari basin, (a) Number of behavioral simulations, (b) *p-factor* and (c) *r-factor*.

The uncertainty analysis results of Gari basin is shown in Figure 4.40 For the number of behavioral simulations, the T-Observed data (17 simulations) obtained with the best results followed with T-LST data (14 simulations), T-IDWEle data (13 simulations) and T-IDW data (8 simulations). The orders are changed for the *p-factor* and *r-factor*. For the *p-factor* the values are 0.82, 0.72, 0.67, 0.58 for T-LST data, T-IDWEle data, T-IDW data, T-IDWEle data, respectively; and the values of *r-factor* are 0.88, 0.73, 0.77, and 0.50, respectively. It is clear that judged by the uncertainly analysis results, the T-LST shows better performances compared the original observed and other interpolations methods.

Though the applicability of SWAT model for snowmelt simulation are proved by previous studies, the results of our study still suggest that there are many other influence factors that may lead to the uncertain results of this model.

However, according to the setting for snowmelt simulation this study keep the other factors as the constant value and made the input air temperature as the only variation for a fixed parameter set, thus the results still indicate that the input air temperature data has great influence on the hydrological simulation results. The best simulation results in indicate that the T-LST data can drive the model to obtain better results compared with other data. The data estimated only based on the air temperature data performed rather worse results for model uncertainty analysis.

Considering the simulation results of different basins, it is clear that the T–LST data showed more advantage for the snowmelt simulation in the study basin. However, for the T-IDW and T-IDWEle data, the results are worse compared with the T-LST data.

It is interesting that the IDWEle and IDW methods provided better results than the LST-Linear method at the verification stations, while the same methods cannot provide better snowmelt simulations. Elevation differences between the subbasins and their interpolation stations were larger than elevation differences between verification stations and interpolation stations.

In addition, the snowmelt simulations indicate that based only on sparsely observed air temperature data, the spatial interpolation method (IDW and IDWEle) cannot always generate accurate temperature data. In fact, the station density of the region was less than $1/20,000$ km$^2$ (87 stations in 2,040,700 km$^2$), significantly less than previous studies based on the IDW method (Courault and Monestiez, 1999, with station density $1/1,250$ km$^2$; Dodson and Marks, 1997, with station density $1/1,000$ km$^2$).

The results indicate that the data density and elevations directly limits the applicability of the IDW and IDWEle methods for driving SWAT snowmelt simulation in a data-sparse basin. Though the performance of the LST-Linear method is relatively

poor at observation stations, the simulation results indicate that the data created by the LST-Linear method are stable and accurate approximations of the actual air temperature in the study basin.

# 5 Conclusions

Snowmelt is one the most important hydrologic processes in mid- and high-latitude regions, especially for the Amur River basin. Snowmelt is recognized has a significant influence on nutrient transport in river channels during spring snowmelt period. The temperature index model is a common tool for simulating basin snowmelt processes because of its generally good performance and fewer data requirements. However, in data-sparse basins, the density of air temperature monitoring data cannot meet the requirements of accurate applications Thus, to create high-density and accurate air temperature data, many data resources and methods have been developed. The main objective of this study was to evaluate a new air temperature data estimation method to generate air temperature data with high spatial density and accuracy, for improving the performance of snowmelt modeling using the temperature index method in Amur River basin.

The research used a simple linear regression equation and MODIS land surface temperature data between two locations to estimate the air temperature data for unknown point. This method was tested at 26 pairs of air monitoring stations and compared with the results of spatial interpolation methods: one is original inverse distance method, the other one is the improved inverse distance method that considering

the topographic influence.

The linear regression relationship between air temperature and land surface temperature of were firstly proved at the observed stations, the results indicated that the linear relationship of both land surface temperature and air temperature are strong. Then, both the linear regression method and the spatial interpolation methods were applied and evaluated by using the observed stations. The inverse distance method is more advanced according to the low elevation of the location of both validation stations and their interpolation stations. Contrarily, the inverse distance weighted method that considering the topographic influence performed relatively large errors according to the lack of stations locating in high elevation areas. Although the approximation of linear regression method might extend errors for air temperature estimation based on land surface temperature data, the results still demonstrate that this simple linear regression approach can estimated the air temperature data with limited errors range during the snowmelt periods, and spatial density of the created data is also very high.

Three test basins (Apkoroshi basin, Malinovka basin, Gari basin) were selected for model testing, which are located in the upper, middle and lower stream of the Amur River Basin. Snowmelt simulations with the newly estimated air temperature data from spatial interpolation method, the linear regression method combined the MODIS LST data and original data were compared in three test basins with varying slope and land cover types. The results of snowmelt simulation clearly demonstrated that only based on the sparse observed data, which cannot represent the spatial distribution of the air temperature in the test basin, the spatial interpolation methods included more errors to air temperature estimation, and finally lead to worse results of the snowmelt simulation. The data generated by the LST-Linear method can obtained the best simulation results

in the basins with relatively high variations of elevations, for the relative flat basin the LST data also need further improvement. Though the LST data obtained relatively weak results for air temperature estimated at a fixed point, the high spatial dense and its accuracy make it as an attractive candidate data source for estimated air temperature data in the basin with topographic influence, and this may expand the applicability of the SWAT model in snowmelt-dominated areas.

# Acknowledgements

The three and half year study in Gifu will be finished with the time flying away, during the study period, I received great help form the teachers and classmates of the United Graduate School of Agricultural Sciences, Gifu University. This dissertation received great help suggestions and encouragement from many kind people, who I cannot write my appreciations one by one in here. I want to take this opportunity to explain my thanks to:

(1) Associated Prof. Takeo Onishi: I would give my sincerest thanks to Associated Prof. Dr. Takeo Onishi, the supervisor and the guider during my doctoral study. He paid a large amount of time and great patient on my study. I owe my gratitude to Dr. Onishi for his kind help to introduce me to Gifu University and setting an exmaple for me to how can I become a qualified doctor. And I would extend my thanks to him for his superb teaching skills on the detail skills for hydrological experiments, constructing the structure of research and English writing. Without his patient and help this dissertation will not be possible.

(2) Prof. Ken Hiramatsu, I would take this opportunity to express my thanks to Prof.Ken Hiramstsu, thanks for his great help to assistant my application to Gifu

(10) I would thanks to the members of my lab and officers of Renno and the BWEL project, they gave me great help for my study and live during these years in Japan.

(11) Finally, I should give my sincerest gratitude to my parents, for their great patient, unrequited love and supports during my study in Japan. As the only child of them; I left them for struggling my own future with the white dyeing their hairs and wrinkle climbing on their faces.

# References

Adam JC, Hamlet AF, Lettenmaier DP. 2009. Implications of global climate change for snowmelt hydrology in the twenty-first century. *Hydrological Processes* 23: 962-972.

Ahl RS, Woods SW, Zuuring HR. 2008. Hydrologic calibration and validation of SWAT in a snow-dominated Rocky Mountain watershed, Montana, USA 1. *JAWRA Journal of the American Water Resources Association* 44: 1411-1430.

Anderson EA. 1976. A point energy and mass balance model of snow cover. *NOAA Technical Report NWS,* 19, U.S. Dept. of Commerce, National Weather Service.

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* 333: 413-430.

Arnold JG, Fohrer N. 2005. SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes* 19: 563-572.

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modeling and assessment part I: Model development1. *Journal of the American Water Resources Association* 34: 73–89.

Bai J, Chen X, Li, J, Yang L, Fang H. 2011. Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. *Environmental monitoring and assessment* 178:247-256.

Benali A, Carvalho AC, Nunes JP, Carvalhais N, Santos A. 2012. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sensing of Environment* 124: 108-121.

Bracmort KS, Arabi M, Frankenberger JR, Engel BA, Arnold JG. 2006. Modeling long-term water quality impact of structural BMPs.*Transactions of the ASAE* 49: 367-374.

Bøggild CE, Knudby CJ, Knudsen MB, Starzer W. 1999. Snowmelt and runoff modelling of an Arctic hydrological basin in west Greenland. *Hydrological Processes* 13: 1989-2002.

Bosilovich MG. 2013. Regional climate and variability of NASA MERRA and recent reanalysis: US summertime precipitation and temperature. *Journal of Applied Meteorology and Climatology* 52: 1939-1951.

Boone A, Etchevers P. 2001. An intercomparison of three snow schemes of varying complexity coupled to the same land surface model: local-scale evaluation at an Alpine site. *Journal of Hydrometeorology* 2: 374-394.

Brooks PD, Williams MW, Schmidt SK. 1998. Inorganic nitrogen and microbial biomass dynamics before and during spring snowmelt. *Biogeochemistry*, 43: 1-15.

Chen G, Iwasaki T, Qin H, Sha W. 2014. Evaluation of the Warm-Season Diurnal Variability over East Asia in Recent Reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. *Journal of Climate* 27:5517-5537.

Colombi A, De Michele C, Pepe M, Rampini A. 2007. Estimation of daily mean air temperature from MODIS LST in Alpine areas. *EARSeL eProceedings* 6: 38-46.

Corriveau J, Chambers PA, Yates AG, Culp JM. 2011. Snowmelt and its role in the hydrologic and nutrient budgets of prairie streams. *Water Science & Technology* 64: 1590-1596.

Courault D, Monestiez P. 1999. Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France. *International Journal of Climatology* 19: 365-378.

Cresswell MP, Morse AP, Thomson MC, Connor SJ. 1999. Estimating surface air temperatures, from Meteosat land surface temperatures, using an empirical solar zenith angle model. *International Journal of Remote Sensing* 20: 1125-1132.

Danilov-Danilyan VI, Gelfan AN, Motovilov YG, Kalugin AS. 2014. Disastrous flood of 2013 in the Amur basin: Genesis, recurrence assessment, simulation results. *Water Resources* 41: 115-125.

DeGaetano AT, Belcher BN. 2007. Spatial interpolation of daily maximum and minimum air temperature based on meteorological model analyses and

independent observations. *Journal of Applied Meteorology and Climatology* 46: 1981-1992.

Dodson R, Marks D. 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate Research* 81: 1-20. DOI: 10.3354/cr008001.

Douglas-Mankin KR, Srinivasan R, Arnold JG. 2010. Soil and Water Assessment Tool (SWAT) model: current developments and applications. *Transactions of the ASAB* 53: 1423-1431.

Duethmann D, Zimmer J, Gafurov A, Güntner A, Kriegel D, Merz B, Vorogushyn S. 2013. Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling.*Hydrology and Earth System Sciences* 17: 2415-2434.

Edwards AC, Scalenghe R, Freppaz M. 2007. Changes in the seasonal snow cover of alpine regions and its effect on soil processes: a review. *Quaternary International* 162: 172-181.

Elsner MM, Gangopadhyay S, Pruitt T, Brekke L, Mizukami N, Clark M. 2014. How does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations?. *Journal of Hydrometeorology* 15:1384-1403.

Essery R, Morin S, Lejeune Y, Bauduin-Ménard C. 2012. A comparison of 1701 snow models using observations from an alpine site. *Advances in Water Resources* 55: 131-148.

Eum HI, Dibike Y, Prowse T, Bonsai B.2014 Inter-comparison of high-resolution gridded climate data sets and their implication on hydrological model simulation over the Athabasca Watershed, Canada. *Hydrological Processes* 28:4250-4271.

Fan Y, Van den Dool H. 2008. A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research* *113*: D01103.

Fang X, Pomeroy JW. 2008. Drought impacts on Canadian prairie wetland snow hydrology. *Hydrological Processes* 22: 2858-2873. DOI: 10.1002/hyp.7074.

Fang X, Pomeroy JW, Westbrook CJ, Guo X, Minke AG, Brown T. 2010. Prediction of snowmelt derived streamflow in a wetland dominated prairie basin. *Hydrology and Earth System Sciences Discussions* 7: 1103-1141.

Fontaine TA, Cruickshank TS, Arnold JG, Hotchkiss RH. 2002. Development of a snowfall-snowmelt routine for mountainous terrain for the soil water assessment tool (SWAT). *Journal of Hydrology* 262: 209-223.

Fuka, DR, Walter MT, MacAlister C, Degaetano AT, Steenhuis TS, Easton ZM. 2013. Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes* 28: 5613-5623.

Gemmer M, Becker S, Jiang T. 2004. Observed monthly precipitation trends in China 19512002. *Theoretical and Applied Climatology* 77: 39-45.

Gillies RR, Wang SY, Huang WR. 2012. Observational and supportive modelling analyses of winter precipitation change in China over the last half century. *International Journal of Climatology* 32: 747-758. DOI: 10.1002/joc/2303.

Hachem S, Duguay CR., Allard M. 2012. Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain. *The Cryosphere* 6: 51-69.

Hayashi M, Quinton WL, Pietroniro A, Gibson JJ. 2004. Hydrologic functions of wetlands in a discontinuous permafrost basin indicated by isotopic and chemical signatures. *Journal of Hydrology* 296: 81-97.

Hong Y, Nix HA, Hutchinson MF, Booth TH. 2005. Spatial interpolation of monthly mean climate data for China. *International Journal of Climatology* 25: 1369-1379.

Hock R. 2003. Temperature index melt modelling in mountain areas. *Journal of Hydrology* 282: 104–115. DOI: 10.1016/S0022-1694(03)00257-9.

Hubbard KG, You J. 2005. Sensitivity analysis of quality assurance using the spatial regression approach: a case study of the maximum/minimum air temperature. *Journal of Atmospheric and Oceanic Technology* 22: 1520-1530.

Immerzeel WW, Van Beek LPH, Konz M, Shrestha AB, Bierkens, M. F. P. 2012. Hydrological response to climate change in a glacierized catchment in the Himalayas. *Climatic change* 110: 721-736.

Jarvis A, Reuter HI, Nelson A, Guevara E. 2008. Hole-filled SRTM for the globe Version 4. CGIAR-CSI SRTM 90m Database (http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1). Accessed September 11, 2014.

Jasper K, Kaufmann P. 2003. Coupled runoff simulations as validation tools for atmospheric models at the regional scale. *Quarterly Journal of the Royal Meteorological Society* 129: 673-692.

Jones P, Jedlovec G, Suggs R, Haines S. 2004. Using MODIS LST to estimate minimum air temperatures at night. In *13th Conference on Satellite Meteorology and Oceanography*, September, pp. 13-18.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83: 1631-1643.

Kloog I, Chudnovsky A, Koutrakis P, Schwartz J. 2012. Temporal and spatial assessments of minimum air temperature using satellite surface temperature measurements in Massachusetts, USA. *Science of the Total Environment* 432: 85-92.

Lauri H, Räsänen TA, Kummu M. 2014. Using reanalysis and remotely sensed temperature and precipitation data for hydrological modelling in monsoon climate: case Mekong River. *Journal of Hydrometeorology* 15:1532-1545.

Liang X, Lettenmaier DP, Wood EF, Burges SJ. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres (1984–2012)* 99(D7): 14415-14428.

Liston GE, Elder K. 2006. A distributed snow-evolution modeling system (SnowModel). *Journal of Hydrometeorology* 7: 1259–1276.

Liu Y, Hiyama T, Yamaguchi Y. 2006. Scaling of land surface temperature using satellite data: A case examination on ASTER and MODIS products over a heterogeneous terrain area. *Remote Sensing of Environment* 105: 115-128.

Makhinov AN. 2004. Natural and Anthropogenic Factors of the Amur River Runoff Formation.http://www.chikyu.ac.jp/AMORE/2004.3kyotoSympo/04.Makhinov.pdf. Accessed September 30, 2014.

Marks D, Domingo J, Susong D, Link T, Garen D. 1999. A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrological Processes* 13: 1935-1959.

Martinec J, Rango A. 1986. Parameter values for snowmelt runoff modelling. *Journal of Hydrology* 84: 197-219.

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG. 2012. An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology* 29: 897-910.

Monestiez P, Courault D, Allard D, Ruget F. 2001. Spatial interpolation of air temperature using environmental context: application to a crop model. *Environmental and Ecological Statistics* 8: 297-309.

Mostovoy GV, King RL, Reddy KR, Kakani VG, Filippova MG. 2006. Statistical estimation of daily maximum and minimum air temperatures from MODIS LST data over the state of Mississippi. *GISciences & Remote Sensing* 43: 78-110.

Muleta MK, Nicklow JW. 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology* 306: 127-145.

Nachtergaele F, Velthuizen HV, Verelst L, Wiberg D. 2012. Harmonized World Soil Database (HWSD). *Food and Agriculture Organization of the United Nations*. http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HWSD_ Documentation.pdf.   Accessed March 16, 2013.

Nagao S, Terashima M, Kodama H, Kim VI, Shestekin PV,  Makhinov AN. 2007. Migration behavior of Fe in the Amur River basin. *Report on Amur-Okhotsk Project* 4: 37-48.

Ninyerola M, Pons X, Roure JM. 2000. A methodological approach of climatological

modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology* 20: 1823-1841.

Nishioka J, Ono T, Saito H, Nakatsuka T, Takeda S, Yoshimura T, et al. 2007. Iron supply to the western subarctic Pacific: Importance of iron export from the Sea of Okhotsk. *Journal of Geophysical Research: Oceans* (1978–2012), 112:(C10).

Onishi T, Mitsudera H, Uchimoto K. 2012. Numerical Simulation of Dissolved Iron Production and Transport in the Amur River and the Sea of Okhotsk. *The Dilemma of Boundaries*. Taniguchi M and Shiraiwa T ed., Springer Japan, pp.87-105.

Perry M, Hollis D. 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology* 25(8):1041-1054.

Pierson DC, Samal NR, Owens EM, Schneiderman EM, Zion MS. 201). Changes in the timing of snowmelt and the seasonality of nutrient loading: can models simulate the impacts on freshwater trophic status?. *Hydrological Processes* 27: 3083-3093.

Piper SC, Stewart EF. 1996. A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling. *Global biogeochemical cycles* 10: 757-782.

Pradhanang SM, Anandhi A, Mukundan R, Zion MS, Pierson DC, Schneiderman EM, Frei A. 2011. Application of SWAT model to assess snowpack development and streamflow in the Cannonsville watershed, New York, USA. *Hydrological Processes* 25: 3268-3277.

Quinton WL, Hayashi M, Carey SK. 2008. Peat hydraulic conductivity in cold regions and its relation to pore size and geometry. *Hydrological Processes*, 22*:* 2829-2837.

Saha, S, Moorthi S, Pan H., Wu X, Wang J, Nadiga S, Reynolds RW. 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* 91: 1015-1057.

Saxton KE, Rawls WJ. 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Science Society of America Journal* 70: 1569-1578.

Senese A, Maugeri M, Vuillermoz E, Smiraglia C, Diolaiuti G. 2014. Using daily air temperature thresholds to evaluate snow melting occurrence and amount on Alpine glaciers by T-index models: the case study of the Forni Glacier (Italy). *The Cryosphere* 8: 1921-1933.

Seneviratne SI, Viterbo P, Lüthi D, Schär C. 2004. Inferring changes in terrestrial water storage using ERA-40 reanalysis data: The Mississippi River basin. *Journal of climate* 17: 2039-2057.

Shen S, Leptoukh GG. 2011. Estimation of surface air temperature over central and eastern Eurasia from MODIS land surface temperature. *Environmental Research Letters 6*: 045206.

Shen SS, Dzikowski P, Li G, Griffith D. 2001. Interpolation of 1961-97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. *Journal of applied meteorology* 40: 2162-2177.

Shibata H, Yoh M, Ohji B, Guo Y, Shi F, Cai T, et al. (2007). Biogeochemical processes of iron and related elements in terrestrial ecosystem of Amur River. *Report on Amur-Okhotsk Project* 4: 75-94.

Stahl K, Moore RD, Floyer JA, Asplin MG, McKendry IG. 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology* 139: 224-236.

Stewart IT, Cayan DR, Dettinger MD. 2004. Changes in snowmelt runoff timing in western North America under a business as usual climate change scenario. *Climatic Change* 62: 217-232.

Stieglitz M, Ducharne A, Koster R, Suarez M. 2001. The impact of detailed snow physics on the simulation of snow cover and subsurface thermodynamics at continental scales. *Journal of Hydrometeorology* 2: 228–242.

Sun YJ, Wang JF, Zhang RH, Gillies RR, Xue Y, Bo YC. 2005. Air temperature retrieval from remote sensing data based on thermodynamics. *Theoretical and Applied Climatology*, 80: 37-48.

Sun, H., Chen, Y., Gong, A., Zhao, X., Zhan, W., & Wang, M. (2014). Estimating mean air temperature using MODIS day and night land surface temperatures. *Theoretical and Applied Climatology* 118, 81-92.

Tekeli AE, Akyürek Z, Arda Şorman A, Şensoy A, Ünal Şorman A. 2005. Using MODIS snow cover maps in modeling snowmelt runoff process in the eastern part of Turkey. *Remote Sensing of Environment* 97: 216-230.

Thornton PE, Running SW, White MA. 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology* 190 : 214-251.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold V, Fiorino M, Woollen J. 2005. The ERA 40 Reanalysis. *Quarterly Journal of the Royal Meteorological Society* 131: 2961-3012.

Vancutsem C, Ceccato P, Dinku T, Connor SJ. 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment* 114: 449-465.

Vu MT, Raghavan SV, Liong SY. 2012. SWAT use of gridded observations for simulating runoff–a Vietnam river basin study. *Hydrology and Earth System Sciences* 16: 2801-2811.

Wan Z. 2008. New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sensing of Environment* 112: 59-74.

Wan Z, Dozier J. 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *Geoscience and Remote Sensing, IEEE Transactions on* 34: 892-905.

Wang X, Yang W, Melesse AM. 2008. Using hydrologic equivalent wetland concept within SWAT to estimate streamflow in watersheds with numerous wetlands. *Transactions of the ASABE* 51: 55-72.

Westermann S, Langer M, Boike J. 2011. Spatial and temporal variations of summer surface temperatures of high-arctic tundra on Svalbard—Implications for MODIS LST based permafrost monitoring. *Remote Sensing of Environment* 115: 908-922.

Xu J, Grumbine, RE, Shrestha A, Eriksson M, Yang X, Wang YUN, Wilkes A. 2009. The melting Himalayas: cascading effects of climate change on water, biodiversity,

and livelihoods. *Conservation Biology 23*: 520-530.

Yang G, Pu R, Zhao C, Huang W, Wang J. 2011. Estimation of subpixel land surface temperature using an endmember index based technique: A case examination on ASTER and MODIS temperature products over a heterogeneous area. *Remote Sensing of Environment* 115: 1202-1219.

Yang J, Reichert P, Abbaspour KC, Xia J, Yang H. 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology* 358: 1-23.

Yang W, Wang X, Liu Y, Gabor S, Boychuk L, Badiou P. 2010. Simulated environmental effects of wetland restoration scenarios in a typical Canadian prairie watershed. *Wetlands Ecology and Management* 18*:* 269-279.

Yang Y, Onishi T, Hiramatsu K. 2014. Improving the Performance of Temperature Index Snowmelt Model of SWAT by Using MODIS Land Surface Temperature Data. *The Scientific World Journal* 2014(ID: 823424): 1-20.

Yang Y, Onishi T, Hiramatsu K. 2015. Impacts of different steal temperature interpolation methods on snowmelt simulation. *Hydrological Research Letters* 9: 27-34.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society* 93: 1401-1415.

Yermoshin VV, Ganzei SS, Murzin AV, Mishina NV, Kudryavtzeva EP. 2007. Creation of GIS for Amur River basin: the basic geographical information. *Annual Report of Amur-Okhotsk Project* 4 : 151-159.

Yu LL, Xia ZQ, Li JK, Cai T. 2013. Climate change characteristics of Amur River. *Water Science and Engineering* **6**: 131-144.

Zakšek K, Schroedter-Homscheidt M. 2009. Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 414-421.

Zhang X, Srinivasan R, Bosch D. 2009. Calibration and uncertainty analysis of the SWAT model using genetic algorithms and Bayesian model averaging. *Journal of Hydrology* 374 : 307-317.

Zhu W, Lǚ A, Jia S. 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment* 130: 62-73.