

High-resolution Climate Data From an Improved GIS-based Regression Technique for South Korea

Hyung-Il Eum*, Jong Pil Kim**, and Jaepil Cho***

Received October 12, 2017/Revised April 25, 2018/Accepted May 1, 2018/Published Online

Abstract

This study presents an improved GIS-based regression model that incorporates an adaptive effective radius algorithm into the structure of previous regression model (Kongju National University Regression Model, KNU/RM), named of IGISRM. The performances of IGISRM were evaluated at various spatial resolutions to select the most suitable grid-spacing considering accuracy and computational efficiency. In addition, the performances between IGISRM and KNU/RM were inter-compared based on various performance measures related to amount and occurrence of precipitation and extreme indices. The results showed that both regression models produced the compatible performance in amount-related measures. However, IGISRM outperformed KNU/RM in occurrence-related measures mainly by the adaptive effective radius. Especially, IGISRM improved the ability in reproducing spatial distribution of precipitation events when the spatial variability in precipitation occurrence highly exists. In contrast, KNU/RM predicted overly wet days due to drizzle effect. Such results indicate that IGISRM has the better skill to capture the heterogeneity of spatial distribution of precipitation. In addition, IGISRM showed higher skill in reproducing the extreme indices, in particular those related to wet and dry spell of precipitation, number of wet days and annual precipitation intensity mainly due to over-predicted wet days of KNU/RM induced by the longer radius of influence circle.

Keywords: *gridded climate data, geo-statistical interpolation, influence radius, South Korea*

1. Introduction

As advances in Atmosphere-Ocean General Circulation Models (AOGCMs) and Regional Circulation Models (RCMs) incorporating the state-of-the-art techniques, high-resolution climate data in the form of regularly spaced grid cells have been generated for projection of spatial climate pattern at global and regional scales (IPCC, 2007). Robust assessment of reliability on the gridded climate data is a prerequisite step to ensure their applicability to various analyses at post stages. Therefore, reliable surface meteorological records are a crucial factor not only to assess the performance of climate models but also to generate feasible projections from ecological, environmental, and hydrological models (Maurer *et al.*, 2002; Christensen and Lettenmaier, 2007; Crozier *et al.*, 2007; Monk *et al.*, 2011; Bennett *et al.*, 2012). In particular, process-based land-surface schemes in a distributed hydrologic model require gridded input data that are typically interpolated to each grid cell using observational station data (Eum *et al.*, 2014). However, a region where sparse surface station density exists, e.g., mountainous regions (Rinke *et al.*, 2004), has a difficulty to appropriately estimate values by a general interpolation scheme due to poor data, complex topography, and

corresponded higher uncertainty in input data (Murphy *et al.*, 2004; Wilby and Harris, 2006; Choi *et al.*, 2009; Kay *et al.*, 2009; Chen *et al.*, 2011; Pouline *et al.*, 2011; Eum *et al.*, 2012; Eum *et al.*, 2014). In addition, many studies have recently applied statistical downscaling methods that require a reliable long-term historical dataset to downscale a coarse scale output of AOGCM for regional or local scale applications such as ecological and hydrologic impacts studies (Brekke *et al.*, 2009; Barsugli *et al.*, 2013; Hanson *et al.*, 2012; Miller *et al.*, 2013; Werner *et al.*, 2013; Gutmann *et al.*, 2014; Hay *et al.*, 2014). Because the spatial resolutions of the statistical downscaling methods corresponds to those of observational data, spatial climate data created by a process of statistically interpolating observational station data have been employed into the statistical downscaling methods to provide a downscaled gridded climate dataset, a finer spatial resolution from AOGCM's outputs (Salathe *et al.*, 2009; Maurer *et al.*, 2010; Abatzoglou and Brown, 2012; Murdock *et al.*, 2013).

Most spatial climate datasets employed in hydro-climatic studies have been generated mainly by simple interpolation such as Inverse Distance Weighting (IDW) and kriging (Eum and Simonovic, 2012; Boer *et al.*, 2001; Wu and Li, 2013), regional

*Surface Water Modeller, Environmental Monitoring and Science Division, Alberta Environment and Parks, 2938 11th Street NE, Calgary, Alberta, Canada, T2E 7L7 (Corresponding Author, E-mail: hyung.eum@gov.ab.ca)

**Research Fellow, National Disaster Management Research Institute, Ulsan 44538, Korea (E-mail: jpkim97@korea.kr)

***Member, Research Fellow, APEC Climate Center (APCC), Busan 612-020, Korea (E-mail: jpcho89@apcc21.org)

regression (Vogel *et al.*, 1999; Johansson and Chen, 2005; Roman *et al.*, 2012), a trivariate thin-plate smoothing spline technique termed ANUSPLIN developed by the Australia National University (Hutchinson, 2004; Hutchinson *et al.*, 2009; Hopkinson *et al.*, 2011), the Daymet model (Thornton *et al.*, 1997), and the Parameter-elevation Regression on Independent Slopes Model (PRISM) (Daly *et al.*, 1994). Among them, the previous studies demonstrated that PRISM provides improved climate grids especially in mountainous and coastal areas by reflecting physiographic similarity to a target grid cell such as distance, elevation, terrain orientation, coastal proximity, and topographic position (Daly *et al.*, 2002; 2003; 2008).

A gridded precipitation and temperature datasets at 5 km spatial resolution has been generated with the PRISM structure based on 526 stations in South Korea (Hong *et al.*, 2007; Shin *et al.*, 2008). The Kongju National University (KNU) has improved the prior datasets in terms of spatial resolution and extension of time window (Lee *et al.*, 2014), called KNU/Regression Model (KNU/RM) in this study, employing 526 station data. The Climate Change Information Center (CCIC) of the Korea Meteorological Administration (KMA) has recently employed KNU/RM to generate high-resolution observational gridded climate data from 2001 to 2014 at 1 km resolution over South Korea (refer to www.climate.go.kr). Because 450 stations (86%) among 526 stations employed in KNU/RM have started to measure climate variables as of the year of 2000, the length of observation period is not enough to be used for various regional applications such as hydrologic modelling and statistical downscaling that require a reliable long-term dataset for calibration and validation. Moreover, KNU/RM used a fixed radius of influence over a domain, which may cause an unexpected bias in a case of an excessively longer radius of influence for a spatially strong heterogeneous region in climatology. Therefore, we generate a long term gridded climate dataset based on 60 stations with more than 30-year observations using an Improved GIS-based Regression Model (IGISRM) that combines the KNU/RM structure with an iterative station density algorithm from the Daymet model. In other words, local regression is implemented based on the KNU/RM structure with the radius of influence varying with the density of stations at each grid cell. To validate the applicability of IGISRM, we compared IGISRM's output with other station data which were not used in modelling during recent years from 2000 to 2010.

A decision spatial resolution for gridded data to be interpolated is not straight-forward because it needs to take into account the reliability of model outputs, the density of observational stations employed in a regression model and a required resolution for applications at post stages. Therefore, we suggest a systematic procedure to select a suitable spatial resolution that generates an acceptable level of accuracy evaluated by Correlation Coefficient (CC), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and bias. This study also validated the performance of IGISRM at the selected spatial resolution by comparing that of

KNU/RM based on multiple performance measures (described in section 3.4). In addition, this study inter-compared the ability of IGISRM and KNU/RM to reproduce climate extreme events that are associated with impacts on socioeconomic and natural hazards by evaluating the CLIMDEX indices (Karl *et al.*, 1999; Shillmann *et al.*, 2013; Zhang *et al.*, 2011).

Therefore, the main objectives of this study are to: 1) develop the IGISRM method and generate a long-term gridded daily climate dataset over South Korea, 2) evaluate the accuracy of generated climate data corresponding to various spatial resolutions and select the most suitable spatial resolution with regard to the accuracy and computational efficiency, and 3) assess the performance of IGISRM by comparing that of the KNU/RM based on multiple performance measures and the CLIMDEX indices.

2. Physiographic and Climatic Data

Physiographical information is a crucial input in the IGISRM structure that estimates parameters in a regression based on physiographic similarity to a target grid cell. High resolution topographic data (~90 m pixels) generated by the NASA's Shuttle Radar Topography Mission (SRTM) for the South Korea region are employed in this study. Topographic facets [Fig. 1(c)] indicating contiguous terrain slope are extracted from the elevation data [Fig. 1(b)] by identifying a direction with maximum slope among eight slopes calculated by a distance and a slope between a target grid and adjacent grid cells at the eight directions. Coastal proximity is designed based on the elevation at a target grid and distance between a target grid and the ocean as shown in Table 1 and corresponded values of coastal proximity from 1 to 6 at each grid are assigned depending on the elevation and distance [Fig. 1(d)] as suggested in Daly *et al.* (2002).

In South Korea, there are mainly two observational systems; one is the Automatic Weather Station (AWS) and the other is the Automated Surface Observing System (ASOS). The average station-spacing including AWS and ASOS over South Korea is approximately 12 km [refer to Fig. 2(b)]. While ASOSs provide a long-term climate data information (more than 40 years), AWSs installed as of 2000 have a relatively short record length, 10 ~ 15 years. A gridded dataset generated by this study may be used for long-term-data-needed studies such as trend analysis and statistical downscaling to extract essential statistics from the dataset. Among all ASOS stations, therefore, this study selected 60 stations [Fig. 2(a)] that provide long-term precipitation, minimum and maximum air temperature data with few missing values and a common period of observational data from 1973 to 2010. While the selected ASOS stations are used to generate a gridded dataset, AWS stations are employed to validate IGISRM suggested in this study. In other words, ASOS stations are used to calibrate and generate a gridded dataset from 1973 to 2010, and validation is implemented by comparing the generated data with observational data at 236 AWS stations from 2000 to 2010.

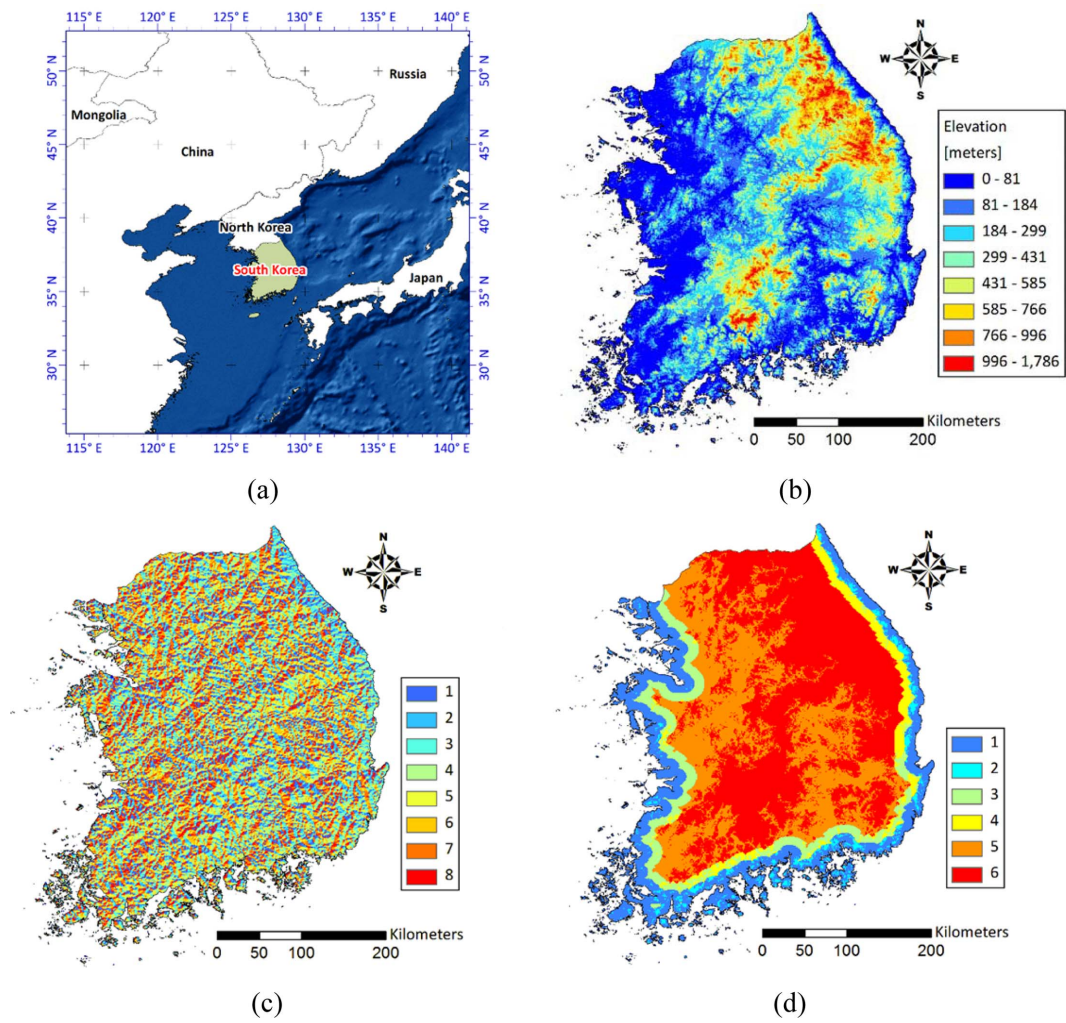


Fig. 1. Schematic Maps of: (a) Study Area, (b) Elevation, (c) Topographic Facet, (d) Coastal Proximity Over South Korea

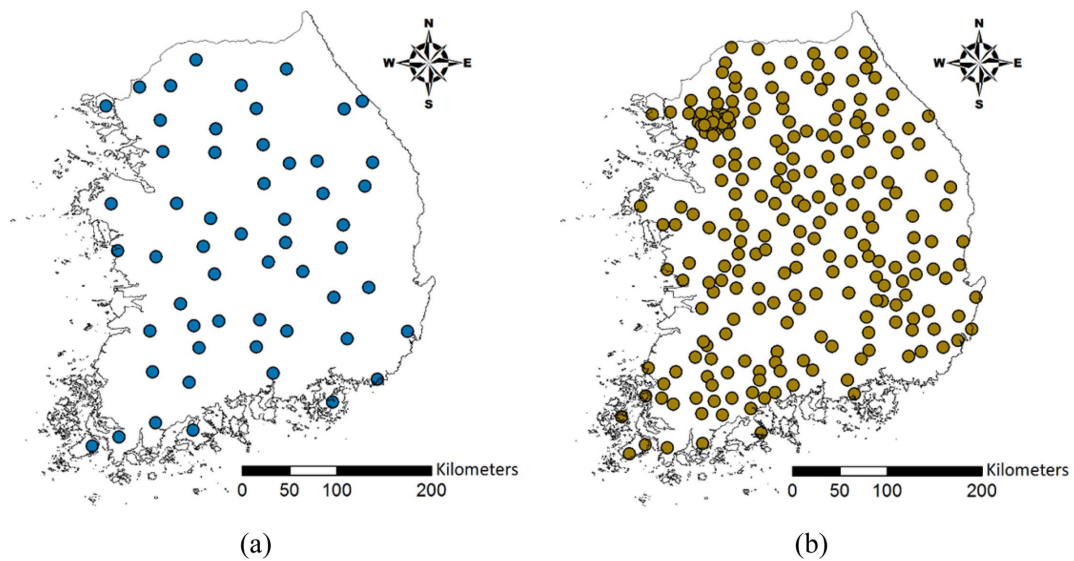


Fig. 2. Surface Networks for: (a) ASOS, (b) AWS Used to Calibrate and Validate Both the IGISM and KNU/RM Models

3. Methodology

3.1 IGISRM

IGISRM newly introduces an iterative station density algorithm into the KNU/RM structure to decide the radius of influence varying with the density of stations at each target grid cell. It is assumed that the spatial distribution of climatology can be captured mainly by elevation, i.e., elevation is the most important independent variable in a regression function since precipitation tends to generally increase as altitude increases and temperature tends to decrease (Barry and Chorley, 1987). Moreover, IGISRM incorporates the effects of terrain-induced climate transitions (facets) and coastal moisture sources into a regression function to determine station weighting in the algorithm.

PRISM takes a form of the simple linear regression as in Eq. (1) because vertical distributions of climate are often linear (Hanson *et al.*, 1982; Osborn, 1984) and simplicity and applicability to estimate climate variables compared with multiple regression schemes that have the complex relationship between multiple independent variables and climate.

$$\hat{y} = \beta_0 \hat{z} + \beta_1 \quad (1)$$

where $\hat{y}$ is the predicted climate variable; $\hat{z}$ is the elevation from GTOPO30, at a target grid cell; β_0 and β_1 are slope and intercept. The slope and intercept are determined by Eq. (2) and (3).

$$\beta_0 = \sum_{i=1}^n w_i (y_i - \bar{y}) (z_i - \bar{z}) / \sum_{i=1}^n w_i (z_i - \bar{z})^2; \beta_1 = \bar{y} - \beta_0 \bar{z} \quad (2)$$

$$\bar{y} = \sum_{i=1}^n w_i y_i / \sum_{i=1}^n w_i; \bar{z} = \sum_{i=1}^n w_i z_i / \sum_{i=1}^n w_i \quad (3)$$

In Eq. (2) and Eq. (3), n is the number of station within the radius of influence; w_i is weighting of i th station; y_i and z_i are climate variable and elevation at i th station, respectively. The predictand, a climate variable at unknown grid point, is determined by elevation and predictor at stations located within a radius of influence circle. In KNU/RM, an initial radius of influence is applied to all grid cells in a domain and is expanded until the number of stations within the radius of influence over all grid cells reaches a minimum number of stations users specified (Lee *et al.*, 2014). However, IGISRM calculates the radius of influence at each grid cell corresponding to the density of stations (the process will be discussed in the next section). At a certain grid cell, the local climate-elevation relationship is not properly captured by a sparse density of stations or a narrow range of elevation within the radius of influence. Such cases are possibly arising mainly from a short radius or the lack of the number of observational stations available in a domain. We employed only 60 stations in South Korea although the minimum number of stations in the radius is typically set to a range from 10 to 30 (Daly *et al.*, 2002). IGISRM increases the radius of influence until a minimum number of stations in the radius is set to 5 due to the limitation of station number in South Korea. Due to the

limited number of stations within an influence radius, an unreasonably steep slope such as a reverse relationship between climate variables and topographic properties might be obtained when sparse observational stations exist near a target grid cell. Based on the historical record over South Korea, Hong *et al.* (2007 and Shin *et al.* (2008) calculated the monthly maximum and minimum slopes, which are employed in IGISRM as a constraint in Eq. (1).

3.2 Combined Weighting of Station and Iterative Station Density Algorithm

The station weighting is obtained by combining weighting of distance, elevation, facet, and coastal proximity as shown in Eq. (4);

$$w_i = (F_r w_r^2 + F_z w_z^2)^{0.5} w_f w_c \quad (4)$$

where w_i is the combined weighting at i th station; w_r , w_z , w_f , and w_c are weights of distance, elevation, topographic facet, and coastal proximity, respectively; F_r and F_z are importance factor for distance and elevation, respectively, that represent a relative importance between horizontal and vertical distances. Because the horizontal distance among observational stations typically has a stronger influence in climate interpolation modelling rather than the vertical distance, F_r and F_z are set to 0.8 and 0.2 in this study as suggested in Daly *et al.* (2002).

The general form of the truncated Gauss filter to calculate distance weighting at the center of a grid cell (p) is shown in Eq. (5) (Thornton *et al.*, 1994).

$$w_r = \begin{cases} 0, & r > R_p \\ \exp\left[-\left(\frac{r}{R_p}\right)^\alpha\right] e^{-\alpha}, & r \leq R_p \end{cases} \quad (5)$$

where w_r is the distance weighting with radial distance r from point p and R_p is the radius of influence circle at a point p and α is a non-dimensional shape parameter and this study employed the value of 6.25 as suggested in Thornton *et al.* (1994). KNU/RM used a fixed radius of influence that guarantees a specified minimum number of stations and weighting is obtained by the inverse distance method. However, a constant R_p may induce a large disparity of station number at regions with sparse and dense density of stations. In other words, a data-poor region has the least number of stations with non-zero weighting and vice versa. We suggest a modified iterative station density algorithm that R_p can be increased for a data-poor region and decreased for a data-rich region (Thornton *et al.*, 1994). R_p is a function of the local station density to achieve a regular station density users specify over a domain. The procedure of a modified iterative station density algorithm to obtain R_p is as following;

1. An initial R_p is set over a domain. The radius derived from KNU/RM (= 150 km) can be used as the initial value of R_p .
2. The local station density, D_p , is determined by Eq. (6) where w_i is obtained from Eq. (5) and $\bar{w}$ is the average weight and defined by Eq. (6).

$$D_p = \frac{1}{\pi R_p^2} \sum_{i=1}^n \frac{w_i}{\bar{w}} \quad (6)$$

$$\bar{w} = \frac{1}{\pi R_p^2} \int_0^{R_p} w_r dr = \left(\frac{1-e^{-\alpha}}{\alpha} \right) e^{-\alpha} \quad (7)$$

Using the local station density, a new R_p is obtained with the average number of stations, N , by Eq. (8). This study set the minimum number of stations to five.

$$R_p = \sqrt{\frac{N}{\pi D_p}} \quad (8)$$

Substituting R_p calculated by Eq. (8) into Eq. (6), steps (2) and (3) are iterated until stop conditions are satisfied. Thornton *et al.* (1994) used only a stop condition until maximum number of iteration is reached while we newly introduced one more stop condition; one is a tolerance ($= 0.0001$) for the absolute difference value of R_p between the previous and current iterations and the other is the maximum number of iteration, 50 in this study.

As a pre-process, once R_p at each grid cell is specified, the values are used at that point in time to calculate station weighting.

Elevation weighting, w_z , represents the influence of stations by altitudinal range of climate data. That is, the elevation weighting expressed by Eq. (9) decreases with the vertical distance between a target grid cell and station.

$$w_z = \begin{cases} \frac{1}{\Delta z_{min}}, & \text{if } \Delta z \leq z_{min} \\ \frac{1}{\Delta z}, & \text{if } z_{min} < \Delta z < z_{max} \\ 0, & \text{if } z \geq z_{max} \end{cases} \quad (9)$$

where Δz represents the vertical distance from a target cell to station; Δz_{min} and Δz_{max} are minimum and maximum vertical distances which are set to 100 m and 500 m as suggested in Daly *et al.* (2002), respectively.

In addition, weights of topographic facet and coastal proximity, w_f and w_c are calculated by Eq. (10) and Eq. (11);

$$w_f = \begin{cases} 1, & \text{if } f \leq 1 \\ \frac{1}{\Delta f}, & \text{if } f > 1 \end{cases} \quad (10)$$

$$w_c = \begin{cases} 1, & \text{if } c = 0 \\ \frac{1}{\Delta c}, & \text{if } 0 < c < c_{max} \\ 0, & \text{if } c \geq c_{max} \end{cases} \quad (11)$$

where Δf is absolute orientation difference between a target cell and a station within R_p . When $f = 1$, the orientation difference is either same or 180° . Δc is the difference of coastal proximity between a target cell and a station. The coastal proximity has been indexed from 1 to 6 corresponding to the distance from a coastline and elevation as shown in Table 1 as suggested by Daly *et al.* (2002).

Table 1. Indexing Coastal Proximity Corresponding to the Elevation and Horizontal Distance from the Coastline

Coastal proximity	Elevation (z)	Distance from the coastline (d)
1	$z \leq 200$ m	$d \leq 10$ km
2	$z > 200$ m	$d \leq 10$ km
3	$z \leq 200$ m	$10 < d \leq 20$ km
4	$z > 200$ m	$10 < d \leq 20$ km
5	$z \leq 200$ m	$d > 20$ km
6	$z > 200$ m	$d > 20$ km

3.3 Precipitation and Temperature Predictions

Precipitation predictions have a complex system that needs to predict both the precipitation occurrence and the total amount of precipitation. Given an assumption that the patterns of precipitation occurrence are spatially coherent at daily time scale, precipitation occurrence is defined by a function of weighted precipitation occurrence at the stations within the radius of influence as shown in Eq. (12).

$$POP = \sum_{i=1}^n w_i PO_i / \sum_{i=1}^n w_i, \quad PO_i = \begin{cases} 0, & \text{if } P_i = 0 \\ 1, & \text{if } P_i > 0 \end{cases} \quad (12)$$

where POP represents the probability of precipitation occurrence; PO_i is a binomial variable that is defined as follow; 0 when the amount of precipitation at i th station is 0 and 1 when the amount of precipitation at i th station is above 0. With such a simple binomial predictor, this study defines the precipitation occurrence when POP is above 0.5.

Prediction methods for (minimum and maximum) air temperature are identical to the procedure for precipitation except incorporating POP . That is, the same linear equation expressed and combined weighting expressed Eq. (1) and Eq. (4) are used with the same radius of influence, R_p , to estimate the minimum and maximum temperature in this study.

3.4 Selection of Spatial Resolution and Performance Measures

Water resources management in South Korea is generally planned per each city or county, which requires at least 5 km spatial resolution. Using 60 ASOS stations over South Korea, therefore, we generate a long-term gridded datasets with IGISM at various levels of spatial resolution (i.e. 1 km, 2 km, 3 km, 4 km, and 5 km) to select a suitable resolution based on both performance measures and computational efficiency. The assessment of model performance is conducted by CC, Bias, MAE, and RMSE as shown in Eq.(13) to Eq.(16).

$$CC = \frac{\sum (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum (P_i - \bar{P})^2} \sqrt{\sum (O_i - \bar{O})^2}} \quad (13)$$

$$Bias = \frac{\sum (P_i - O_i)}{\sum O_i} \quad (14)$$

$$MAE = \frac{1}{N} \sum |P_i - O_i| \quad (15)$$

$$RMSE = \sqrt{\frac{\sum (P_i - O_i)^2}{N}} \quad (16)$$

where P_i and O_i are predicted and observed values, respectively; N is total number of observation at the stations.

Because KNU/RM uses a fixed radius of influence circle at all grid cells, a longer radius can be obtained at a poor-density region to guarantee the minimum number of stations within the radius. Spatial continuity of precipitation is much less than that of other variables (e.g. air temperature and humidity), particularly in a complex topographic region. In a case with the longer radius of influence, therefore, it might often happen that there are many wet days with tiny amounts of precipitation at many grid cells, named as “drizzle effect”. In other words, KNU/RM may generate too many wet days in a year compared with observational station data. Therefore, we evaluate five performance measures that focus on the correspondence of precipitation occurrence between predictions and observations (Murphy 1997); Frequency bias of occurrence [FBias, Donaldson *et al.* (1975)], hit rate [H, Swets (1986)], false-alarm ratio [FAR, Donaldson *et al.* (1975)], proportion correct [PC, Finley (1884)], and critical success index [CSI, Gilbert (1884)]. The definition of the five performance measures are described in Eq. (17) to Eq. (21) which are functions of the number of success and/or failure of prediction based on the four possible combinations of predictions and observations as shown in a contingency table (Table 2).

$$FBias = \frac{a+b}{a+c} \quad (17)$$

$$H = \frac{a}{a+c} \quad (18)$$

$$FAR = \frac{b}{a+b} \quad (19)$$

$$PC = \frac{a+d}{a+b+c+d} \quad (20)$$

$$CSI = \frac{a}{a+b+c} \quad (21)$$

FBias represents the ratio of the number of predictions of occurrence to the number of observed occurrences. The perfect score of FBias is equal to 1 and above 1 represents a higher number of wet days from predictions than that from the observations.

Table 2. Contingency Table to Categorize a Binary Event of Prediction. The Sum of All Events are the Number of Observation; $N = a + b + c + d$

Predicted condition	True condition		
	Positive	Negative	Total
Positive	a	b	$a+b$
Negative	c	d	$c+d$
Total	$a+c$	$b+d$	$N = a + b + c + d$

On the contrary, less 1 of FBias indicates less number of wet days in predictions. H represents the ratio of correct predictions for precipitation occurrences, which is also termed by a probability of detection (Donaldson *et al.*, 1975). Sometimes, the hit rate is termed to represent Proportion Correct (PC) (Wilks, 1995). However, this study distinguishes the two measures to identify how IGISRM predicts well rainy or no-rainy days at all grid cells. FAR is an estimate of conditional probability of a false alarm given that rainy days are predicted, consequently 0 is the perfect ratio. CSI represents an estimate of the conditional probability of a hit given that the rainy days are either predicted or observed, or both. The perfect score of CSI, therefore, is 1.

3.5 Assessment for Extreme Events

Severe socio-economic impacts on natural and infrastructural systems are induced by extreme events such as flood, drought, and heat and cold waves (Butt *et al.*, 2011; Wernberg *et al.*, 2012). A requirement to analyze the impacts related to extreme climate events has been increasing at regional and global scales (Tebaldi *et al.*, 2006; Kharin *et al.*, 2007; Min *et al.*, 2011; Zhang *et al.*, 2005; Burgers *et al.*, 2012). We evaluate the ability of IGISRM and KNU/RM to reproduce extreme climate indices referred to CLIMDEX (www.climdex.org, Zhang *et al.*, 2011; Sillmann *et al.*, 2013). All CLIMDEX indices are calculated yearly-based and the numbers of heavy precipitation (R_{nnmm}), which is equal to or greater than different thresholds, are increasing by 10 mm from 10 mm to 80 mm to investigate the change of performance in IGISRM and KNU/RM corresponding to different levels of heavy rainfall. Using 236 AWSs' data from 2000 to 2010, we inter-compare the skill of the two GIS-based regression models (IGISRM and KNU/RM) to reproduce extreme events using correlation coefficient and Kolmogorov-Smirnov D ($K-S D$) statistic (Murdock *et al.*, 2013) that is the supremum (maximum) distances of empirical distributions between observations and generated data.

4. Results and Discussion

4.1 Decision of Influence Radius (R_p)

In the IGISRM procedure, all grid cells in a domain have different radii of influence, R_p , depending on the station density within R_p . Fig. 3 represents the spatial distribution of R_p decided by the iterative station density algorithm. The estimated R_p values range from 41 km to 268 km, where station-adjacent regions (represented by green dots in Fig. 3) are occupied by blue-color while regions far from stations are represented by yellow or red colors, in particular the northern and north-eastern areas that are data-poor region showed longer R_p than those at the other regions. Such results indicate that R_p calculated by the modified iterative station density algorithm properly captures spatial variability of station density at each grid cell. We also decided the radius of influence circle to implement the KNU/RM, 150 km to include the minimum number of stations over all grid cells in South Korea.

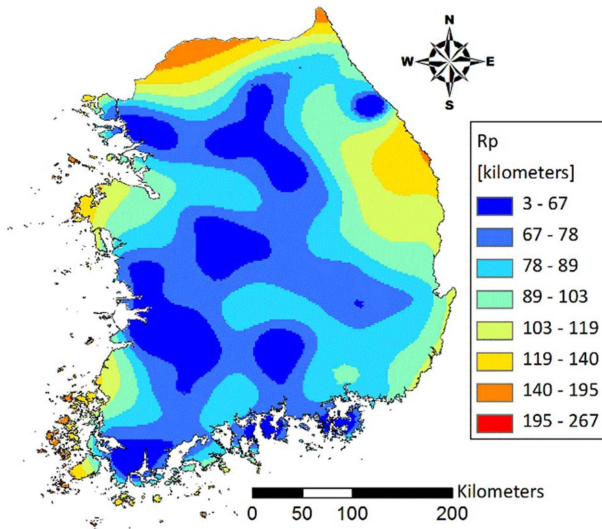


Fig. 3. Spatial Distribution of Radius of Influence Circle (Rp) Obtained by the Iterative Station Density Algorithm in IGISM

4.2 Selection of Spatial Resolution of IGISM's Gridded Data

Deciding a spatial resolution is an important step before generating gridded data with IGISM. Very high resolutions such as 0.5 km or 1 km require more computational burden and data storage capacity while a coarse resolution is generally

difficult to capture the impacts of physiographical characteristics in climate variables due to the smoothing effect by averaging topographic features at each grid cell (Cherubini *et al.*, 2002; Petrik *et al.*, 2011). Therefore, we tested the various levels of spatial resolution to identify a threshold grid-spacing that may show a sharply-shift level of accuracy. Because water resources managers in South Korea require at least 5 km resolution to build a long-term plan per city or county, therefore, we tested IGISM with five different spatial resolutions from 1 km to 5 km to investigate which resolution is the most suitable with regard to the accuracy and computational efficiency.

Four performance measures (CC, RMSE, MAE, and Bias) corresponding to each spatial resolution are tabulated in Table 4. All numbers are averaged over all grid cells located in South Korea. Each gridded dataset is generated for 38 years from 1973 to 2010 and then evaluated the four measures by comparing AWS station data from 2000 to 2010 for validation. As expected, 1 km-resolution provides the best performance on most measures for all variables; precipitation (PRCP), maximum temperature (TMAX), and minimum temperature (TMIN). Interestingly, we found a threshold (3 km) of spatial resolution that is a changing point in performance for TMAX and TMIN although the performance for PRCP is similar to each resolution. To clearly identify the change point of spatial resolution in performance, all measures are normalized and summed to avoid the scale effect, where only CC takes the reverse form. Fig. 4 presents the

Table 3. CLIDEX Indices Employed in This Study and Their Description

ID	Definition	Description	Unit
CDD	Maximum length of dry spell	Maximum number of consecutive days with daily PRCP < 1 mm	Day
CWD	Maximum length of wet spell	Maximum number of consecutive days with daily PRCP	Day
PRCPTOT	Annual wet-day total PRCP	Annual total PRCP in wet days (daily PRCP ≥ 1 mm)	Mm
R01mm	Number of wet days	Annual count of days when PRCP ≥ 1 mm	Day
Rnnmm	Number of PRCP days	Annual count of days when PRCP ≥ nmm (10 mm ~ 80 mm in this study)	Day
R95pTOT	Very wet days	Annual total PRCP when daily PRCP > 95p	mm
R99pTOT	Extremely wet days	Annual total PRCP when daily PRCP > 99p	mm
Rx1day	Max 1-day PRCP	Annual maximum 1-day precipitation (PRCP)	mm
SDII	Simple precipitation intensity index	Annual precipitation divided by the number of wet days	mm/day
CSDI	Cold spell duration index	Annual count of days with at least 6 consecutive days when daily minimum temperature (TMIN) < 10 th percentile	Day
DTR	Daily temperature range	Annual mean difference between daily maximum temperature (TMAX) and TMIN	°C
FD	Frost days	Annual count of days when TMIN < 0°C	Day
ID	Icing days	Annual count of days when TMAX < 0°C	Day
SU	Summer days	Annual count of days when TMAX > 25°C	Day
TN10p	Cool night	Percentage of days when TMIN < 10 th percentile	%
TN90p	Warm night	Percentage of days when TMIN > 90 th percentile	%
TN _n	Min TN	Annual minimum value of TMIN	°C
TN _k	Max TN	Annual maximum value of TMIN	°C
TR	Tropical nights	Annual count of days when TMIN > 20°C	Day
TX10p	Cool days	Percentage of days when TMAX < 10 th percentile	%
TX90p	Warm days	Percentage of days when TMAX > 90 th percentile	%
TX _n	Min TX	Annual minimum value of TMAX	°C
TX _k	Max TX	Annual maximum value of TMAX	°C
WSDI	Warm spell duration index	Annual count of days with at least 6 consecutive days when TMAX > 90 th percentile	Day

Table 4. Evaluation of Performance Measures of IGISM for Precipitation, Minimum and Maximum Air Temperature Corresponding to the Five Spatial Resolutions

Climate variable	Measures	Resolution/Number of grid cells				
		1 km/391,476 (646 × 606)	2 km/98,496 (324 × 304)	3 km/44,268 (217 × 204)	4 km/25,256 (164 × 154)	5 km/16,368 (132 × 124)
Precipitation (PRCP)	CC	0.85	0.86	0.86	0.86	0.85
	RMSE	7.47	7.43	7.40	7.31	7.71
	MAE	1.83	1.84	1.84	1.85	1.92
	Bias	0.02	0.06	0.06	0.07	0.13
Maximum temperature (TMAX)	CC	0.99	0.99	0.99	0.99	0.99
	RMSE	1.54	1.55	1.58	1.77	1.91
	MAE	1.19	1.21	1.24	1.41	1.52
	Bias	-0.11	-0.17	-0.05	-0.23	-0.20
Minimum temperature (TMIN)	CC	0.98	0.98	0.98	0.98	0.98
	RMSE	2.28	2.48	2.39	2.52	2.74
	MAE	1.62	1.81	1.72	1.84	2.04
	Bias	0.08	-0.02	0.12	-0.17	0.28

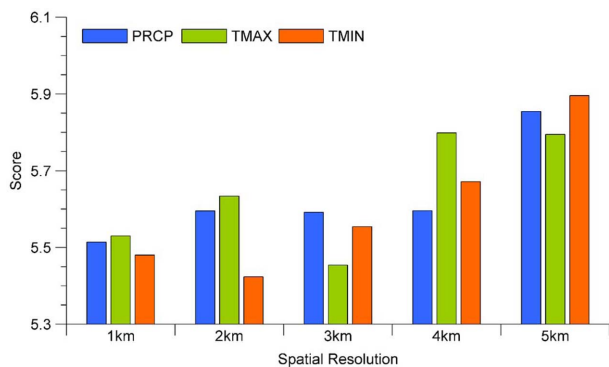


Fig. 4. Summation of Normalized Measures for PRCP, TMAX, and TMIN Corresponding to the Five Spatial Resolutions Tested in This Study

summation scores for PRCP, TMAX, and TMIN corresponding to each spatial resolution, where a lower value indicates better performance. The scores of coarser resolutions (i.e. 4 km and 5 km) are visually higher than those of finer resolutions. While the performance varies with each variable, the sum of scores at 1 km to 3 km for the three variables are 16.5, 16.7, and 16.6, respectively, showing very compatible performances. However, the difference of grid cell number between 1 km and 3 km is 347,208 cells and 1-km resolution also needs approximately 13 hours computational time more to complete the generation of gridded data for 38 years compared with that of the resolution of 3 km. Such results indicate that the suitable spatial resolution of gridded data might be 3 km with regard to similar accuracy and less computational time compared to those of 1-km resolution. Therefore, we

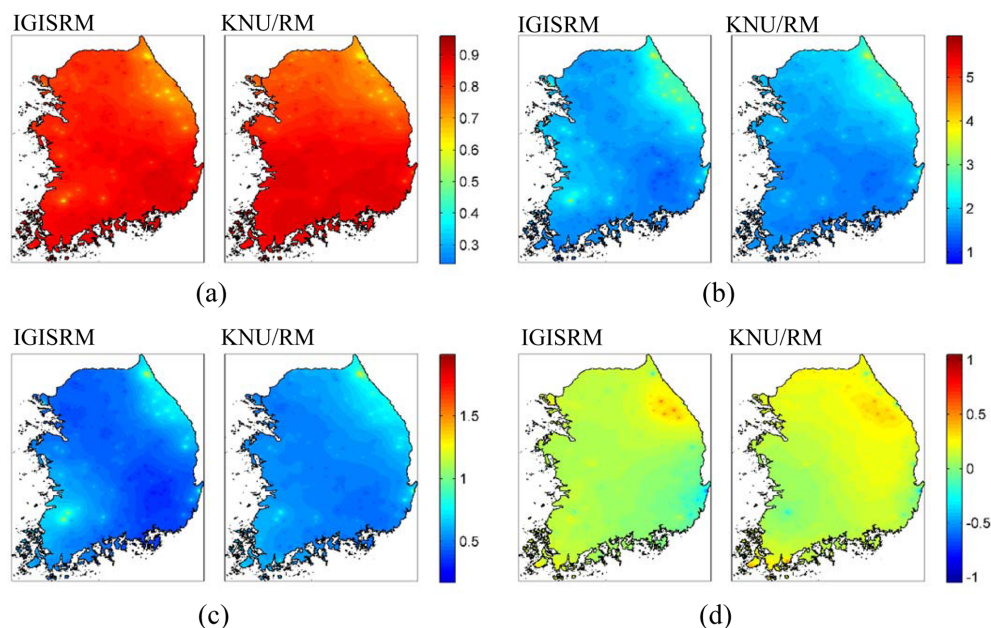


Fig. 5. Spatial Map of Performance Measures of IGISM and KNU/RM for PRCP: (a) CC, (b) RMSE, (c) MAE, (d) Bias

selected 3-km resolution as the most suitable grid-spacing.

4.3 IGISRM vs. KNU/RM

We inter-compare the performances between IGISRM and KNU/RM at the 3 km-resolution using both the four performance measures [Eq. (13) to (16)] used in selection of spatial resolution and additional five performance measures related to precipitation occurrence [Eq. (17) to (21)].

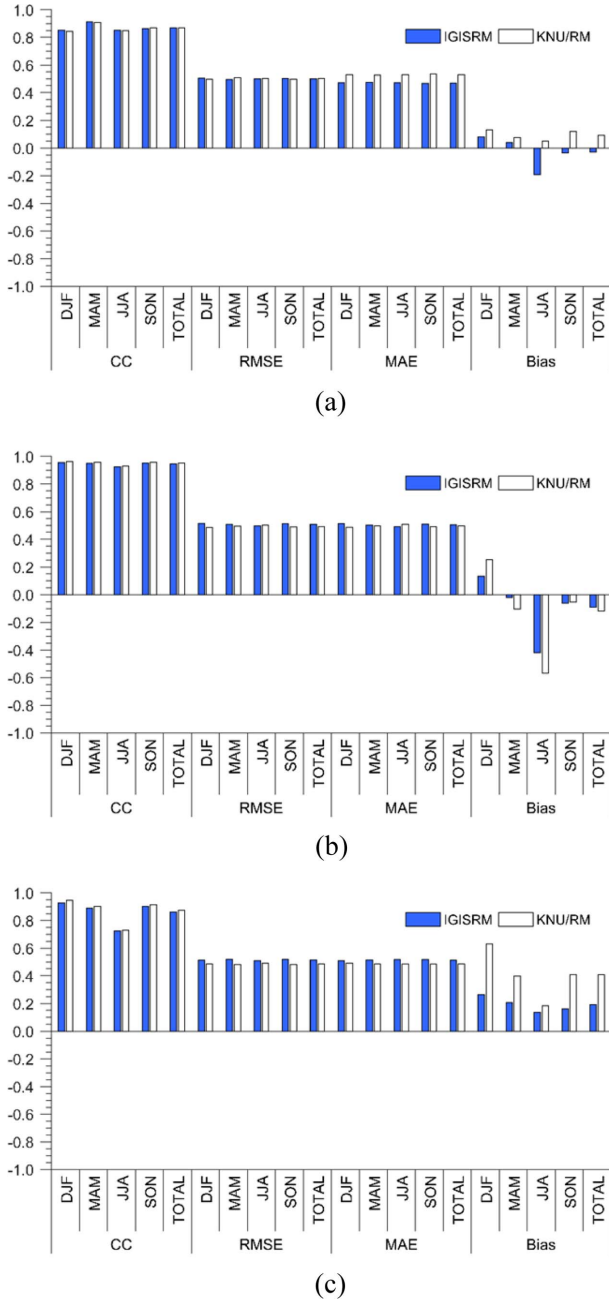


Fig. 6. Inter-comparison of Performance Measures for PRCP, TMAX, and TMIN; DJF (winter), MAM (spring), JJA (summer), SON (autumn), and TOTAL (all time period). RMSE is Normalized for Visualizing the Graph: (a) PRCP, (b) TMAX, (c) TMIN

Figure 5 shows the spatial distribution of the four measures for only PRCP that typically show higher spatial variability. Overall, two schemes generate less accurate climate data at the east area where sparse station network exists due to mountainous topography. In addition, the areas showing less skill (represented by light or red colors) by IGISRM are smaller than those by KNU/RM. Around the center of South Korea where the station density is relatively higher, IGISRM produces better results for RMSE [Fig. 5(b)], MAE [Fig. 5(c)], and Bias [Fig. 5(d)]. In addition, performance measures of IGISRM and KNU/RM for PRCP, TMAX, and TMIN during each season and total period are presented in Fig. 6. IGISRM is slightly better than KNU/RM in MAE and Bias for PRCP. Bias of KNU/RM is lower than that of IGISRM for the summer season (JJA) in which high precipitations occur due to the monsoon and typhoons. However, Bias of IGISRM during the total period is better due to lower Bias during other seasons. For TMAX, RMSE and MAE of KNU/RM are lower than those of IGISRM. For TMIN, IGISRM shows better in Bias for all seasons and total period. Overall, the superiority between two schemes is not clearly detected because large differences in CC, RMSE, and MAE between the two schemes are not observed for all variables while IGISRM shows considerable improvement in Bias for both TMIN and TMAX.

Therefore, five performance measures related to precipitation occurrence are evaluated to examine as to which technique provides more reliable and accurate climate information. The five measures averaged over all grid cells for all time period are presented in Fig. 7. FBias of KNU/RM is above 1, which indicates that it predicts rainy points more than those of observation (AWS points) while FBias of IGISRM is close to the perfect score 1. Such result is associated with the excessively longer radius of influence of KNU/RM that is applied to all grid cells in the domain because more stations far from a target grid cell may be included in calculation of POP. Such over-prediction of wet days affects the H score, showing KNU/RM has a higher score. For other measures (i.e. FAR, PC, and CSI), however, IGISRM outperforms KNU/RM. In particular, FAR shows the largest difference between IGISRM and KNU/RM, which may be induced by difference of the over-prediction of wet days between

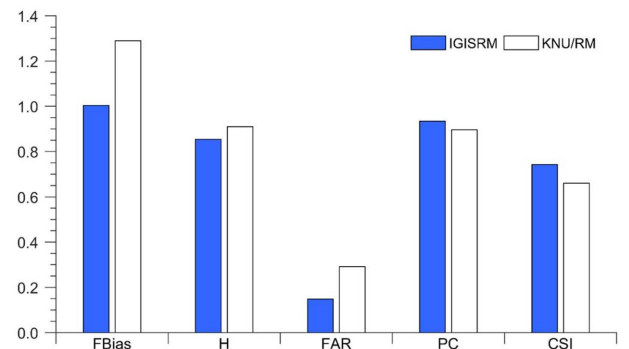


Fig. 7. Inter-comparison of Performance Measures Related to Precipitation Occurrence between IGISRM and KNU/RM: (a) FBias, (b) H, (c) FAR, (d) PC, (e) CSI

Table 5. Ten Precipitation Events Selected Based on Percentages of Rainfall-detected AWSs and Representative Statistics of Each Event

Case	Date (yyyy/mm/dd)	Number of station $\geq 1 \text{ mm d}^{-1}$	Rate (Col.3 / total number of stations $\times 100$) (%)	Maximum rainfall (mm d^{-1})	Averaged rainfall (mm d^{-1})
1	2006/04/18	23	9.4	46.5	1.3
2	2008/07/14	47	19.2	33.5	1.6
3	2008/08/08	71	29.0	74.0	3.1
4	2000/08/08	97	39.6	90.0	4.9
5	2002/07/21	118	48.2	202.5	12.0
6	2009/07/15	144	58.8	211.0	33.2
7	2006/07/27	164	66.9	203.0	53.1
8	2010/06/12	190	77.6	87.5	19.7
9	2009/07/07	215	87.8	274.0	53.2
10	2009/07/09	236	96.3	246.0	89.2

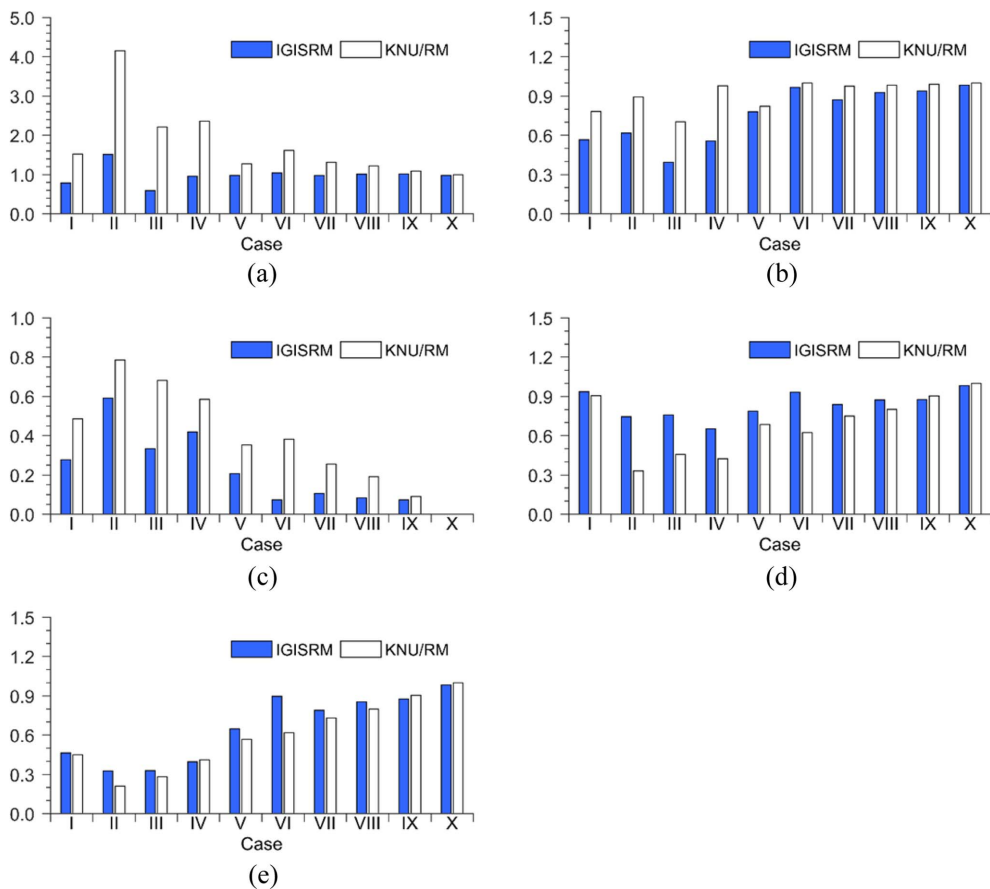


Fig. 8. Performance Measures of IGISRM and KNU/RM for Ten Cases Selected by Percentages of Rainfall-detected AWS Stations in South Korea: (a) Observation, (b) IGISRM, (c) KNU/RM

IGISRM and KNU/RM.

We also tested the skill related to predicting precipitation occurrence by evaluating the performance for 10 precipitation cases that have different percentages of rainfall-detected AWS station number as shown in Table 5. A lower percentage of rainfall-detected AWS station number represents the lower continuity of precipitation occurrence, which indicates the higher spatial variability of precipitation events. Five performance measures for the 10 cases are shown in Fig. 8. It is prominent that

KNU/RM shows the over-prediction of wet days, $\text{FBias} > 1$, in all cases while IGISRM predicts precipitation occurrence properly, near 1 from Case 4 to Case 10. In particular, KNU/RM extremely over-predict wet days in the Case 2 (20% of rainfall-detected AWS stations). Overall, IGISRM performs better than KNU/RM for all measures except the H score while the two schemes show compatible performances for high percentages of rainy stations (i.e., Case 8 to Case 10). Interestingly, the performance of KNU/RM is improved as the percentage increases from Case 2 to Case

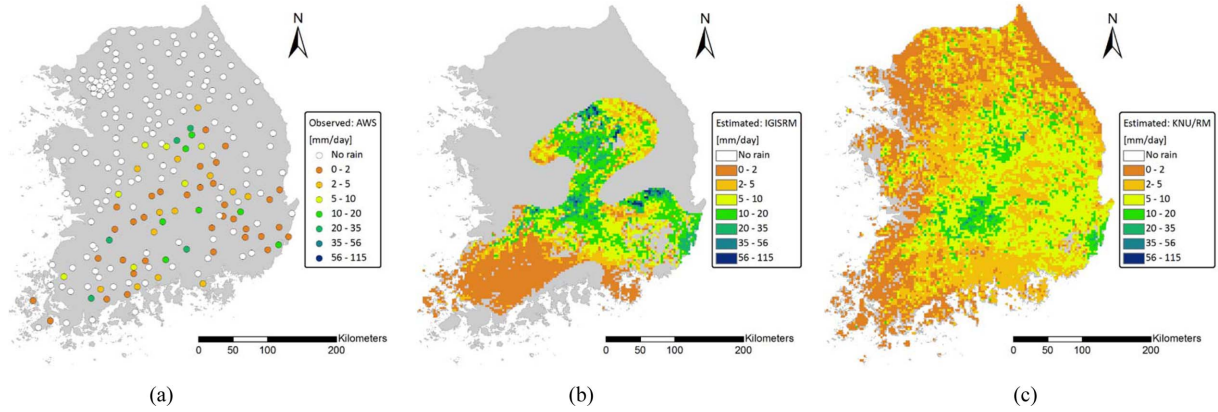


Fig. 9. Spatial Distributions of PRCP Derived from Observation, IGISRM, and KNU/RM on 14 July 2008 (Case 2). Observed Precipitations are Represented by Point Data at AWS Stations

10, clearly found in performances for FBias, FAR, and CSI. Therefore, it is evident that the differences in all performance measures between IGISRM and KNU/RM for the Case 1 to Case 4 are much larger than other cases, which might be induced by the higher spatial variability of precipitation occurrences on those cases. In particular, IGISRM prominently outperforms KNU/RM in the Case 2 that notably high spatial variability in precipitation occurrence exists. Fig. 9 shows the spatial distribution of precipitation for observation, IGISRM and KNU/RM on July 18, 2008 (Case 2). As expected, KNU/RM predicts overly wet areas, mostly colored by orange and yellow, while IGISRM shows good agreement with observation. Such results are strongly related to the difference of distance of R_p between IGISRM and KNU/RM. In other words, KNU/RM used the longer R_p that may induce higher POP by excessively including many rainy AWS stations far from a target grid cell, consequently producing the drizzle effect. Such results indicate that IGISRM with the radius of influence varying with station density has a better skill to capture spatial heterogeneity in precipitation occurrence on wet days.

4.4 Inter-comparison of IGISRM and KNU/RM for CLIMDEX

Precipitation and temperature-related extreme climate indices listed in Table 3 are evaluated to examine which scheme produces more accurate information on extreme events. Fig. 10 presents the correlation coefficient and the $K-S D$ statistics between observation and generated data at 236 AWS stations from 2000 to 2010. Note that lower $K-S D$ statistics [Fig. 10(b)] represents better skill and the values in Fig. 10 represent the average over all AWS stations. In addition, the first 16 indices, i.e. from CDD to SDII, indicate precipitation-related indices and other indices after SDII represent temperature-related indices. Notable differences of CLIMDEX between the two schemes in correlation coefficient are observed in the CDD and CWD indices, representing maximum length of dry and wet spell, respectively, those are related to the skill to predict precipitation occurrence. Because KNU/RM predicts overly many wet days due to the longer R_p , it shows the less skill in reproducing CDD

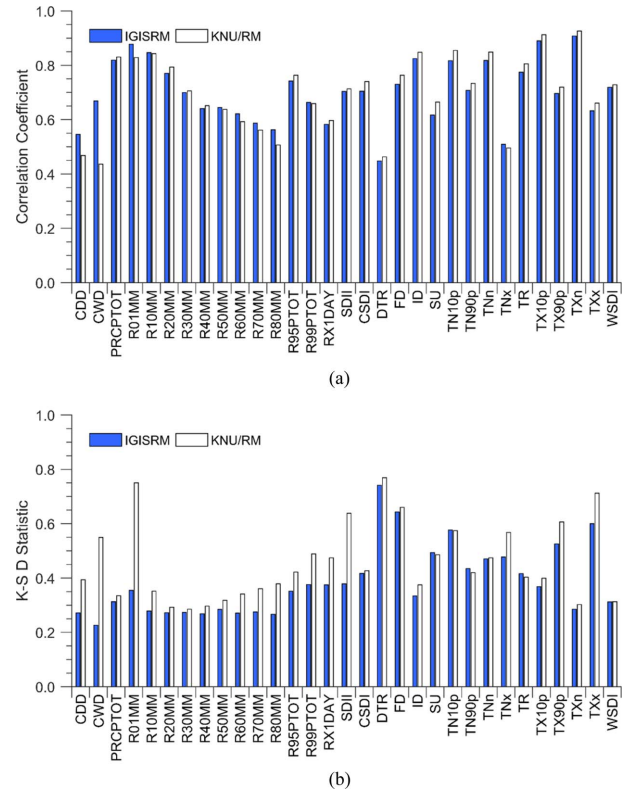


Fig. 10. Inter-comparison of Performance on CLIMDEX Indices between IGISRM and KNU/RM Based on Correlation Coefficient and K-S D Statistic: (a) Correlation Coefficient, (b) K-S D Statistics

and CWD with regard to correlation coefficient compared with those of KNU/RM. However, any significant differences are not detected for other CLIMDEX indices in the correlation coefficient. Regarding the $K-S D$ statistics, it is found that the two schemes show compatible skill for the temperature-related indices. However, IGISRM shows better skill (lower $K-S D$ statistics) for all precipitation-related indices. In particular, significant differences are found in CDD, CWD, R01MM, and SDII. Such results may

be also closely related to the drizzle effect of KNU/RM. For example, the R01MM index, representing the annual count of days when $PRCP \geq 1$ mm, may be a much higher number by the drizzle effect in KNU/RM compared with the observation. In addition, SDII of KNU/RM shows relatively low due to a too much higher number of wet days in a year. On the contrary, IGISRM introducing varying influence radius by the iterative station density algorithm shows better skills in the CLIMDEX indices related to precipitation occurrence.

6. Conclusions

We suggested an improved GIS-based regression model, named of IGISRM, by introducing the iterative station density algorithm into the structure of a prior model KNU/RM to incorporate radius of influence circle (R_p) varying with the station density near a target grid. We evaluated the performances of IGISRM at various spatial resolutions to select a suitable spatial resolution with regard to accuracy and computational efficiency. In addition, we inter-compared the performances between IGISRM and KNU/RM by evaluating not only performance measures related to the amount of each variable and precipitation occurrence but also CLIMDEX indices to examine the ability to reproduce extreme climate events. The results showed

1. The iterative station density algorithm in IGISRM properly reflected the station density near each grid cell in calculation of R_p (41 km ~ 268 km).
2. 3-km resolution properly produced a long-term gridded climate dataset with regard to the performance measures and total number of grid cells to be interpolated. The selected grid spacing might be changed when employing more observational stations into the IGISRM procedure.
3. Both IGISRM and KNU/RM showed compatible performances in the performance measures related to the amount of each variable, i.e. CC, Bias, MAE and RMSE except that IGISRM showed better skill in Bias for TMIN.
4. IGISRM outperformed KNU/RM for all performance measures related to precipitation occurrence (FBias, H, FAR, PC and CSI). In particular, IGISRM produced better spatial distribution of precipitation when spatial variability in precipitation occurrence is higher, for example Case 2 to Case 4. Such results indicate that IGISRM can provide better skill to capture spatial heterogeneity in precipitation occurrence.
5. While the two schemes showed the compatible skill for the temperature-related the CLIMDEX indices, IGISRM showed the higher skill in reproducing CDD, CWE, R01MM, and SDII that are closely related to the drizzle effect of KNU/RM by the longer R_d inducing overly wet days over the domain.

A poor station density, only 60 ASOS stations in this study, was employed in this study to produce a long-term gridded climate dataset and to inter-compare the performances between IGISRM and KNU/RM. Originally, KNU/RM has been developed with 450 AWS and ASOS stations (~12 km station-spacing on average) over South Korea. An optimal station density might be

suggested by investigating how the performance of IGISRM is improved by increasing the station density using all 450 stations, which might be economically important to optimize expenses for maintenance and installation of stations.

Acknowledgements

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 18AWMP-B083066-05). The authors would also like to thank APCC colleagues for their valuable comments and suggestions to improve the quality of this manuscript.

References

- Abatzoglou, T. J. and Brown, J. T. (2012). "A comparison of statistical downscaling methods suited for wildfire applications." *Int. J. Climatol.*, Vol. 32, pp. 772-780, DOI: 10.1002/joc.2312.
- Barsugli, J. J. and Coauthors. (2013). "The Practitioner's Dilemma: How to assess the credibility of downscaled climate projections." *EOS, Trans. Am. Geophys. Union*, Vol. 94, No. 46, pp. 424-425, DOI: 10.1002/2013EO460005.
- Barry, R. G and Chorley, R. J. (1987). *Atmosphere, weather, and climate*, 5th ed. Routledge, London, United Kingdom, 460p.
- Bennett, K. E., Werner, A. T., and Schnorbus, M. (2012). "Uncertainties in hydrologic and climate change impact analyses in headwater basins of british Columbia." *Journal of Climate*, Vol. 25, No. 17, pp. 5711-5730, DOI: 10.1175/JCLI-D-11-00417.1.
- Boer, E. P., de Beurs, K. M., and Hartkamp, A. D. (2001). "Kriging and thin plate splines for mapping climate variables." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 3, No. 2, pp. 146-154, DOI: 10.1016/S0303-2434(01)85006-6.
- Brekke, L. D., Kiang, J. E., Olsen, J. R., Pulwarty, R. S., Raff, D. A., Turnipseed, D. P., Webb, R. S., and White, K. D. (2009). *Climate change and water resources management—A federal perspective*, U.S. Geological Survey Circular, 1331, 65p. [Available at <http://pubs.usgs.gov/circ/1331/>].
- Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., and Cannon, A. J. (2012). "Downscaling extremes—an intercomparison of multiple statistical methods for present climate." *Journal of Climate*, Vol. 25, No. 12, pp. 4366-4388, DOI: 10.1175/JCLI-D-11-00408.1.
- Butt, N., De Oliveira, P. A., and Costa, M. H. (2011). "Evidence that deforestation affects the onset of the rainy season in Rondonia, Brazil." *J. Geophys. Res.*, Vol. 116, No. 11, DOI: 10.1029/2010JD015174.
- Chen, J., Brissette, F. P., Poulin, A., and Leconte, R. (2011). "Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed." *Water Resour. Res.*, Vol. 47, W12509, DOI: 10.1029/2011WR010602.
- Cherubini, T., Ghelli, A., and Lalaurette, F. (2002). "Verification of precipitation forecasts over the Alpine region using a high-density observing network." *Weather Forecast.*, Vol. 17, pp. 238-249, DOI: 10.1175/1520-0434(2002)017<0238:VOPFOT>2.0.CO;2.
- Choi, W., Kim, S. J., Rasmussen, P. F., and Moore, A. R. (2009). "Use of the North American regional reanalysis for hydrological modelling in Manitoba." *Can. Water Resour. J.*, Vol. 34, No. 1, pp. 17-36, DOI: 10.4296/cwrj3401017.
- Christensen, N. and Lettenmaier, D. P. (2007). "A multimodel ensemble approach to assessment of climate change impacts on the hydrology

- and water resources of the Colorado River basin." *Hydrol. Earth Syst. Sc.*, Vol. 11, pp. 1417-1434, DOI: 10.5194/hess-11-1417-2007.
- Crozier, L., Zabel, R. W., and Hamlet, A. H. (2007). "Predicting differential effects of climate change at the population level with life-cycle models of spring Chinook salmon." *Global Change Biol.*, Vol. 14, pp. 236-249, DOI: 10.1111/j.1365-2486.2007.01497.x.
- Daly, C. (2006). "Guidelines for assessing the suitability of spatial climate data sets." *Int. J. Climatol.*, Vol. 26, pp. 707-721, DOI: 10.1002/joc.1322.
- Daly, C., Gibson, W. P., Taylor, G. H., Johnson, G. L., and Pasteris, P. (2002). "A knowledge-based approach to the statistical mapping of climate." *Climate Research*, Vol. 22, No. 2, pp. 99-113, DOI: 10.3354/cr022099.
- Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., and Pasteris, P. P. (2008). "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States." *International Journal of Climatology*, Vol. 28, No. 15, pp. 2031-2064, DOI: 10.1002/joc.1688.
- Daly, C., Helmer, E. H., and Quiñones, M. (2003). "Mapping the climate of puerto rico, vieques and Culebra: CLIMATE MAPPING." *International Journal of Climatology*, Vol. 23, No. 11, pp. 1359-1381, DOI: 10.1002/joc.937.
- Daly, C., Neilson, R. P., and Phillips, D. L. (1994). "A statistical-topographic model for mapping climatological precipitation over mountainous terrain." *Journal of Applied Meteorology*, Vol. 33, pp. 140-158, DOI: 10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.
- Dingman, S. L., Seely-Reynolds, D. M., and Reynolds, R. C. (1988). "Application of Kriging to estimating mean annual precipitation in a region of orographic influence." *Journal of the American Water Resources Association*, Vol. 24, No. 2, pp. 329-339, DOI: 10.1111/j.1752-1688.1988.tb02991.x.
- Donaldson, R. J., Dyer, R. M., and Kraus, M. J. (1975). *An objective evaluator of techniques for predicting severe weather events*, In: Preprints, 9th Conference on Severe Local Storms, October 21-23, Norman, Oklahoma.
- Eum, H.-I., Dibike, Y., Prowse, T., and Bonsal, B. (2014). "Inter-comparison of high-resolution gridded climate data sets and their implication on hydrological model simulation over the Athabasca Watershed, Canada." *Hydrol. Process.*, Vol. 28, pp. 4250-4271, DOI: 10.1002/hyp.10236.
- Eum, H.-I., Gachon, P., Laprise, R., and Ouara, T. (2012). "Evaluation of regional climate model simulations versus gridded observed and regional reanalysis products using a combined weighting scheme." *Clim. Dynam.*, Vol. 38, pp. 1433-1457, DOI: 10.1007/s00382-011-1149-3.
- Eum, H.-I. and Simonovic, S. P. (2012). "Assessment on variability of extreme climate events for the Upper Thames River basin in Canada." *Hydrol. Process.*, Vol. 26, pp. 485-499, DOI: 10.1002/hyp.8145.
- Finley, J. P. (1884). "Tornado predictions." *American Meteorological Journal*, Vol. 1, pp. 85-88.
- Gilbert, G. K. (1884). "Finley's tornado predictions." *Bull. Amer. Meteor. Soc.*, Vol. 1, pp. 166-172.
- Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M. (2014). "An intercomparison of statistical downscaling methods used for water resources assessments in the United States." *Water Resour. Res.*, Vol. 50, pp. 7167-7186, DOI: 10.1002/2014WR015559.
- Hanson, C. L. (1982). "Distribution and stochastic generation of annual and monthly precipitation on a mountainous watershed in southwest Idaho." *J. Amer. Water Resour. As.*, Vol. 18, pp. 875-883, DOI: 10.1111/j.1752-1688.1982.tb00085.x.
- Hanson, R. T., Flint, L. E., Flint, A. L., Dettinger, M. D., Faunt, C. C., Cayan, D., and Schmid, W. (2012). "A method for physically based model analysis of conjunctive use in response to potential climate changes." *Water Resour. Res.*, Vol. 48, W00L08, DOI: 10.1029/2011WR010774.
- Hay, L., LaFontaine, J., and Markstrom, S. (2014). "Evaluation of statistically downscaled GCM output as input for hydrological and stream temperature simulation in the Apalachicola-Chattahoochee-Flint River Basin (1961-1999)." *Earth Interact.*, Vol. 18, pp. 1-32, DOI: 10.1175/2013EI000554.1.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). "Very high resolution interpolated climate surfaces for global land areas." *International Journal of Climatology*, Vol. 25, No. 15, pp. 1965-1978, DOI: 10.1002/joc.1276.
- Hong, K. O., Suh, M. S., Rha, D. K., Chang, D. H., Kim, C., and Kim, M. K. (2007). "Estimation of high resolution gridded temperature using GIS and PRISM." *Atmosphere*, Vol. 17, pp. 255-268 (in Korean).
- Hopkinson, R. F., McKenney, D. W., Milewska, E. J., Hutchinson, M. F., Papadopol, P., and Vincent, L. A. (2011). "Impact of aligning climatological day on gridding daily maximum-minimum Temperature and precipitation over Canada." *Journal of Applied Meteorology and Climatology*, Vol. 50, No. 8, pp. 1654-1665, DOI: 10.1175/2011JAMC2684.1.
- Hutchinson, M. F. (2004). *ANUSPLIN version 4.3. Centre for Resource and Environmental Studies*, Australian National University Canberra. [Available online at <http://fennerschool.anu.edu.au/publications/software/anusplin.php>.]
- Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E., and Papadopol, P. (2009). "Development and testing of Canada-Wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961-2003." *Journal of Applied Meteorology and Climatology*, Vol. 48, No. 4, pp. 725-741, DOI: 10.1175/2008JAMC1979.1.
- IPCC (2007). *Climate change 2007: Impacts, adaptation and vulnerability working group II contribution to the Intergovernmental Panel on Climate Change Fourth Assessment Report*, Intergovernmental Panel on Climate Change, Geneva, Switzerland, ISBN 978 0521 88010-7 hardback.
- Johansson, B. and Chen, D. (2005). "Estimation of areal precipitation for runoff modelling using wind data: A case study in Sweden." *Climate Research*, Vol. 29, No. 1, pp. 53-61, DOI: 10.3354/cr029053.
- Karl, T. R., Nicholls, N., and Ghazi, A. (1999). "CLIVAR/GCOS/WMO workshop on indices and indicators for climate extremes. Workshop summary." *Clim. Chang.*, Vol. 42, pp. 3-7, DOI: 10.1007/978-94-015-9265-9_2.
- Kay, A. L., Davies, H. N., Bell, V. A., and Jones, R. G. (2009). "Comparison of uncertainty sources for climate change impacts: Flood frequency in England." *Clim. Chang.*, Vol. 92, pp. 41-63, DOI: 10.1007/s10584-008-9471-4.
- Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C. (2007). "Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations." *Journal of Climate*, Vol. 20, No. 8, pp. 1419-1444, DOI: 10.1175/JCLI4066.1.
- Lee, W.-S., Chun, J. A., and Kang, K. (2014). "Development and application of GIS-based PRISM integration through a plugin approach." *Journal of Hydrology*, Vol. 513, pp. 58-67, DOI: 10.1016/j.jhydrol.2014.03.029.
- Maurer, E. P., Hidalgo, H. G., and Das, T. (2010). "The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California." *Hydrol. Earth Syst. Sc.*, Vol. 14,

- pp. 1125-1138, DOI: 10.5194/hess-14-1125-2010.
- Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B. (2002). "A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States." *Journal of Climate*, Vol. 15, No. 22, pp. 3237-3251, DOI: 10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.
- Miller, W. P., DeRosa, G. M., Gangopadhyay, S., and Valdes, J. B. (2013). "Predicting regime shifts in flow of the Gunnison River under changing climate conditions." *Water Resour. Res.*, Vol. 49, pp. 2966-2974, DOI: 10.1002/wrcr.20215.
- Min, S.-K., Zhang, X., Zwiers, F. W., and Hegerl, G. C. (2011). "Human contribution to more-intense precipitation extremes." *Nature*, Vol. 470, pp. 378-381, DOI: 10.1038/nature09763.
- Monk, W. A., Peters, D. L., Curry, A. R., and Baird, D. J. (2011). "Quantifying trends in indicator hydroecological variables for regime-based groups of Canadian rivers." *Hydrol. Process.*, Vol. 25, No. 19, pp. 3086-3100, DOI: 10.1002/hyp.8137.
- Murdock, T. Q., Cannon, A. J., and Sobie, S. R. (2013). *Statistical downscaling of future climate projections*, Pacific Climate Impacts Consortium (PCIC) Report (No.KM170-12-1236).
- Murphy, A. H. (1997). "The Economic value of weather and climate forecasts." *Cambridge University Press, Cambridge*, pp. 19-74, ISBN: 9780521434201.
- Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A. (2004). "Quantification of modelling uncertainties in a large ensemble of climate change simulations." *Nature*, Vol. 430, pp. 768-772, DOI: 10.1038/nature02771.
- Osborn, H. B. (1984). "Estimating precipitation in mountainous regions." *J. Hydraul. Eng.*, Vol. 110, pp. 1859-1863, DOI: 10.1061/(ASCE)0733-9429(1984)110:12(1859).
- Petrik, R., Baldauf, M., Schlunze, H., and Gassmann, A. (2011). "Validation of a mesoscale weather prediction model using subdomain budgets." *Tellus A*, Vol. 63, No. 4 pp. 707-726, DOI: 10.1111/j.1600-0870.2011.00528.x.
- Phillips, D. L., Dolph, J., and Marks, D. (1992). "A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain." *Agr. Forest Meteorol.*, Vol. 58, pp. 119-141, DOI: 10.1016/0168-1923(92)90114-J.
- Pouline, A., Brissette, F., Leconte, R., Arsenault, R., and Maloet, J.-S. (2011). "Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin." *J. Hydrol.*, Vol. 409, pp. 626-636, DOI: 10.1016/j.jhydrol.2011.08.057.
- Rinke, A., Marbaix, P., and Dethloff, K. (2004). "Internal variability in Arctic regional climate simulations: Case study for the SHEBA year." *Clim. Res.*, Vol. 27, pp. 197-209, DOI: 10.3354/cr027197.
- Roman, D. C., Vogel, R. M., and Schwarz, G. E. (2012). "Regional regression models of watershed suspended-sediment discharge for the eastern United States." *J. Hydrol.*, Vol. 472-473, pp. 53-62, DOI: 10.1016/j.jhydrol.2012.09.011.
- Salathe, E. P., Steed, R., Mass, C. F., and Pahn, Z. (2009). "A high-resolution climate model for the United States Pacific Northwest: Mesoscale feedbacks and local responses to climate change." *J. Clim.*, Vol. 21, pp. 5708-5726, DOI: 10.1175/2008JCLI2090.1.
- Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., and Bronaugh, D. (2013). "Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate." *J. Geophys. Res.*, Vol. 118, No. 4, pp. 1716-1733, DOI: 10.1002/jgrd.50203.
- Shin, S. C., Kim, M. K., Suh, M. S., Rha, D. K., Jang, D. H., Kim, C. S., Lee, W. S., and Kim, Y. H. (2008) "Estimation of high resolution gridded precipitation using GIS and PRISM." *Atmosphere*, Vol. 18, pp. 71-81 (in Korean).
- Swets, J. A. (1986). "Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance." *Psychol. Bull.*, Vol. 99, pp. 181-198, DOI: 10.1037/0033-2909.99.2.181.
- Tebaldi, C., Hayhoe, K., Arblaster, J. M., and Meehl, G. A. (2006). "Going to the extremes: An intercomparison of model-simulated historical and future changes in extreme events." *Climatic Change*, Vol. 79, Nos. 3-4, pp. 185-211, DOI: 10.1007/s10584-006-9051-4.
- Thornton, P., Running, S. W., and White, M. A. (1997). "Generating surfaces of daily meteorological variables over large regions of complex terrain." *J. Hydrol.*, Vol. 190, pp. 214-251, DOI: 10.1016/S0022-1694(96)03128-9.
- Vogel, R. M., Wilson, I., and Daly, C. (1999). "Regional regression models of annual streamflow for the United States." *J. Irrig. Drain. Eng.*, Vol. 125, pp. 148-157, DOI: 10.1061/(ASCE)0733-9437(1999)125:3(148).
- Wernberg, T., Smale, D. A., Tuya, F., Thomsen, M. S., Langlois, T. J., de Bettignies, T., Bennett, S., and Rousseaux, C. S. (2012). "An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot." *Nature Clim. Change*, Vol. 3, No. 1, pp. 78-82, DOI: 10.1038/nclimate1627.
- Werner, A. T., Schnorbus, M. A., Shrestha, R. R., and Eckstrand, H. D. (2013). "Spatial and temporal change in the hydro-climatology of the Canadian portion of the Columbia River basin under multiple emissions scenarios." *Atmos. Ocean*, Vol. 51, pp. 357-379, DOI: 10.1080/07055900.2013.821400.
- Wilby, R. L. and Harris, I. (2006). "A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK." *Water Resour. Res.*, Vol. 42, W02419, DOI: 10.1029/2005WR004065.
- Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences: An Introduction*, Academic Press, San Diego, ISBN: 9780080541723.
- Wu, T. and Li, Y. (2013). "Spatial interpolation of temperature in the United States using residual kriging." *Appl. Geogr.*, Vol. 44, pp. 112-120, DOI: 10.1016/j.apgeog.2013.07.012.
- Zhang, X., Hegerl, G., Zwiers, F. W., and Kenyon, J. (2005). "Avoiding inhomogeneity in percentile-based indices of temperature extremes." *Journal of Climate*, Vol. 18, No. 11, pp. 1641-1651, DOI: 10.1175/JCLI3366.1.
- Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T.C., Trewin, B., and Zwiers, F. W. (2011). "Indices for monitoring changes in extremes based on daily temperature and precipitation data." *Wiley Interdiscip. Rev. Clim. Change*, Vol. 2, No. 6, pp. 851-870, DOI: 10.1002/wcc.147.