

# SVM 회귀 모형을 활용한 격자 강우량 상세화 기법

## Spatial Downscaling of Grid Precipitation Using Support Vector Machine Regression

문희원\* / 백종진\*\* / 황석환\*\*\* / 최민하\*\*\*\*

Moon, Heewon / Baik, Jongjin / Hwang, Sukhwan / Choi, Minha

### Abstract

A spatial downscaling method using the Support Vector Machine (SVM) Regression for 25 km Tropical Rainfall Measuring Mission (TRMM) Monthly precipitation is proposed. The nonlinear relationship among hydrometeorological variables and precipitation was effectively depicted by the SVM for predicting downscaled grid precipitation. The accuracy of spatially downscaled precipitation was estimated by comparing with rain gauge data from sixty-four stations and found to be improved than the original TRMM data in overall. Especially the positive bias of the original TRMM data was effectively removed after the downscaling procedure. The spatial distributions of 25 km and 1 km grid precipitation were generally similar, while the local spatial trend was better detected by 1 km grid precipitation. The downscaled grid data derived from the proposed method can be applied in hydrological modelling for higher accuracy and further be studied for developing optimized downscaling method incorporation other regression methods.

**keywords** : TRMM 3B43 V7, SVM, downscaling, ASOS

### 요 지

본 연구에서는 Tropical Rainfall Measuring Mission (TRMM) 3B43 V7 (25 km)의 월 누적 격자 강우량을 1 km 해상도로 상세화하기 위해 Support Vector Machine (SVM) 회귀를 활용한 상세화 기법을 제안하였다. 비선형 예측모델인 SVM은 상세화의 기반이 되는 다양한 수문기상인자와 강우 발생간의 월별 상관성 구축에 효율적으로 활용되었다. 상세화된 격자 강우는 전국에 고루 분포한 64개 지점 관측 강우와의 비교 분석을 통해 상세화 이전의 격자 강우 보다 다소 개선된 정확도를 지니는 것으로 확인되었다. 특히, 상세화 이전 격자 강우가 지니는 양의 Bias가 효과적으로 개선되었다. 상세화 전후의 공간분포 비교에서 두 분포는 평균적으로 유사했으나, 상세화 이전 강우의 공간분포에서 나타나지 않았던 강우의 국지적 특성이 상세화된 공간분포를 통해 잘 표현되는 것을 확인할 수 있었다. 특히, 일부 지점의 과소 및 과대산정이 상세화를 통해 개선되어 전반적인 정확도 향상에 기여하였음을 확인했다. 본 연구에서 제안된 상세화 기법이 적용된 격자 강우는 모델의 정확도 향상을 위한 고해상도 입력자료로 활용될 수 있으며, 추후 연구에서는 SVM 외에 다른 회귀 방식을 활용하여 최적의 강우 상세화 기법 개발에 기여할 수 있을 것으로 보인다.

**핵심용어** : TRMM 3B43 V7, SVM, 상세화 기법, ASOS

\* 성균관대학교 건설환경시스템공학과 석사과정 (e-mail: hwmooon@skku.edu)

Graduate Student, Dept. of Civil, Architectural and Environmental System Engineering, Sungkyunkwan Univ., Suwon 440-746, Korea

\*\* 성균관대학교 건설환경시스템공학과 박사수료 (e-mail: jjbaek@skku.edu)

Ph.D. Candidate, Dept. of Civil, Architectural and Environmental System Engineering, Sungkyunkwan Univ., Suwon 440-746, Korea

\*\*\* 한국건설기술연구원 수자원연구실 수석연구원 (e-mail: sukany@kict.re.kr)

Senior Researcher, Water Resources Research Division, Korea Institute of Civil Engineering and Building Technology

\*\*\*\* 교신저자, 성균관대학교 수자원대학원 수자원학과 부교수 (e-mail: mhchoi@skku.edu, Tel: 82-31-290-7527)

Corresponding Author, Associate Professor, Dept. of Water Resources, Graduate School of Water Resources, Sungkyunkwan Univ., 2066 Seobu-ro, Jangan-gu, Suwon, Gyeonggi-do 440-746, Korea

## 1. 서 론

강우는 지구규모 수문순환과정의 핵심적 구성 인자로, 기상현상 자체로서 지니는 중요성 외에도 다양한 수문 모델 및 대기 모델의 정확도를 결정하는 입력 인자로 사용되기 때문에 다양한 연구분야에서 그 가치와 활용도가 매우 크다(Creutin et al., 1988; Su et al., 2008). 현재까지 강우관측은 주로 육상에서 실시되었으며, 가장 널리 사용되는 관측 장비는 강우 우량계이다. 그러나 우량계는 한 지점의 일정 반경 내 강우량만을 대표하기 때문에 공간적인 강우의 분포를 파악하기에는 한계가 있다(Jia et al., 2011). 지점 관측의 한계를 보완하기 위하여 사용되는 기상레이더를 통한 강우 추정에는 강우의 시공간적 분포를 효과적으로 감지하는 데에 효율적이다. 그러나 강우를 직접 관측하는 것이 아니라 반사파를 통해 강우를 간접적으로 추정하는 데에서 발생하는 부정확성에 대한 단점이 있고, 우량계와 마찬가지로 일정 반경 내의 강우 분포에 대해서만 대표성을 나타내는 한계를 가진다. 이러한 문제들을 개선하기 위해 1960년대 들어 인공위성자료를 이용한 강우 관측에 대한 연구가 본격적으로 시작되었다. 인공위성을 통해 관측된 강우는 지점 관측 자료와 비교해 공간분포 파악에 용이하나, 국지적 강우 분포를 파악하기에는 다소 낮은 공간해상도로 인하여 선행연구들을 통해 이를 개선하기 위한 다양한 상세화 기법들이 연구된 바 있다. 이 기법들은 주로 강우와 상관성을 지니는 기상인자 및 지형적 특성과 강우 발생 간 회귀분석을 통해 고해상도의 강우를 계산하는 방식으로 사용되었으며, 다변량회귀, 인공신경망, 정준상관관계 분석, 주성분분석, SVM 회귀 등이 적용되었다(Chen et al., 2010).

Immerzeel et al. (2009)은 이베리아 반도를 대상으로 해당 지역에서 연 평균 식생과 연 누적강우 사이에 존재하는 지수적 관계에 기반한 회귀 분석을 통해 25 km Tropical Rainfall Measuring Mission (TRMM) 강우를 1 km로 상세화하였다. Jia et al. (2011)도 이와 유사한 방식으로 강우 자료 상세화를 수행하였는데, 티벳 고원 내 Qaidam 유역을 대상으로 식생 외에 고도를 고려하였으며, 강우를 포함한 세 개 인자 간 다중 선형 회귀 모델을 구축하여 25 km TRMM 강우를 1 km 해상도로 상세화 한 바 있다. Park (2013)은 한국을 대상으로 정규식생지수(Normalized Difference Vegetation Index, NDVI)와 DEM(Digital Elevation Model)을 사용하여 Residual Kriging과 지구통계학적 상세화 기법을 적용해 2005년 10월의 25 km TRMM 강우를 1 km로 상세화하는 실험적 성격의 연구를 수행하였으

며 지상관측 강우와의 비교를 통해 상세화 이후의 강우분포가 향상된 정확도를 가지는 것을 확인하였다. 이처럼 유역의 지형적 특성과 기후적 특성에 의해 강우자료 상세화에 사용되는 인자는 달라질 수 있으나, 식생은 강우와 강한 상관관계를 갖는 인자로 월 단위 이상의 시간적 규모를 가지는 강우 자료의 상세화에 관한 연구에서 공통적으로 고려되었으며, DEM을 사용한 지형적 특성의 고려는 주로 최고해발고도 4,000 m의 높은 산지지형에 대한 연구에서 수행되었다(Yin et al., 2008; Jia et al., 2011). 연구지역의 고도가 크게 높지 않은 경우 강우에 대한 지형적 특성의 고려는 배제된 경우가 많으며, 이는 습도 및 대기 중 가능강수 수분량 등 강우와 보다 직접적 상관관계를 가지는 인자들에 대한 고려로 보완되었다(Beckmann and Buishand, 2002; Charles et al., 1999; Cavazos, 1999). Chen et al. (2010)은 일 강우에 대한 상세화를 위해 평균 해수면 기압, 상대습도, 비습도, 대기온도, 풍향, 와도 등의 인자와 강우 사이의 관계성을 SVM 회귀, 다변량 회귀 그리고 인공신경망 모형을 통해 각각 구축하여 상세화된 강우 자료의 정확성을 검증하였다. 세 방법 모두 상세화 과정에 적용되었을 때 유효함을 확인하였으나, 전반적으로 SVM 회귀가 보다 정확한 결과를 나타냈으며 특히 극한 강우사상의 예측에 대한 정확도가 상대적으로 높은 것을 확인하였다. 반대로 적은 양의 강우 사상 상세화에는 다변량 회귀에 의한 상세화 결과가 다른 두 방법보다 정확했다. 이외에도 다양한 기상인자와 강우 간 관계성을 토대로 앞서 언급한 회귀분석 방법을 적용해 강우의 시공간적 상세화에 대한 많은 연구들이 수행된 바 있다(Bardossy, 1997; Anandhi et al., 2008; Chu et al., 2008; Olsson et al., 2001). 특히 SVM 회귀 방법을 다른 회귀 방법과 함께 적용하여 비교한 연구의 경우 공통적으로 SVM을 사용한 예측 값이 다른 방법에 비해 전반적으로 우수한 것으로 나타났다. 이와 같은 SVM 회귀 모형의 우수성은 강우의 상세화 연구 뿐 아니라 이를 활용하여 다른 인자의 공간적 상세화 및 시계열 예측 등에 적용한 분야에서도 확인된 바 있다(Keramitoglou et al., 2013; Osuna et al., 1997; Gunn, 1998; Kwon et al., 2012). 학습 및 일반화 능력이 SVM의 입력변수 및 매개변수의 선택에 지나치게 큰 영향을 받게되는 단점에도 불구하고(Wang, 2012; Ben-David and Lindenbaum, 1997; Cho, 2013), 근래에는 변수 선택의 최적화를 돕는 도구의 개발로 보다 다양한 분야에의 적용이 활발해지고 있다.

이와 같은 다양한 선행연구를 토대로 본 연구에서는 월 단위 강우와 유효한 상관성이 있는 것으로 판단되는 식생, 지표온도, 이슬점 온도 및 가능강수를 입력변수로 선정하

였으며, 다양한 인자 간의 상관관계를 기반으로 종속변수를 다른 회귀 방법에 비해 효과적으로 예측하는 SVM-Regression 방법을 적용하여 강우에 대한 상세화를 수행하였다. 25 km의 공간해상도를 지니는 TRMM 3B43 Version 7 (V7)의 2009년 1월에서 12월까지의 월 누적 강우 자료를 1 km 공간해상도로 상세화 하였으며 상세화 된 격자 강우의 유효성을 검증하기 위해 전국 64개 Automated Synoptic Observing System (ASOS) 강우 자료와의 비교·분석을 수행하여 상세화 기법의 적용성 및 적합성을 확인하였다.

## 2. 연구 지역 및 데이터

### 2.1 연구 지역

한반도는 아시아 대륙의 북반구, 33°N~38°N 및 124°E~131°E 사이에 위치하고 있으며 사계절이 뚜렷한 계절적 특성을 나타낸다. 우리나라의 우기기간은 장마가 시작되는 6월 하순부터 가을이 시작되는 9월이며, 이 기간 동안의 강수량은 연간 강수량 약 1,300 mm 중 65% 이상을 차지한다. 이러한 강우특성을 보이는 대표적인 이유는 한반도가 여름철 태평양에서 불어오는 남동 및 남서 계절풍의 영향 때문이다. 본 연구에서는 2009년 1월에서 12월까지 월별 누적강우량을 사용하였다.

기상청에서는 고층기상장비로 5개 지점의 GPS (Global Positioning System) 레원존데, 9개 지점의 Radiometer와 윈드프로파일러를 운영하고 있으며, 지상에서의 기상관측 업무수행을 위해 2개의 지상기상관측망인 자동기상관측장비(Automatic Weather System, AWS)와 중관기상관측장비(ASOS)를 사용하고 있다. 한반도전역에는 총

600개 이상 지점의 지상기상관측망을 운영하고 있으며, 각각 AWS는 550개 이상, ASOS 81개를 현재 운용 중에 있다 (<http://www.kma.go.kr>). 본 연구에서는 TRMM 3B43 V7 월 누적강우량 및 상세화 된 자료와의 유효성에 대한 비교·검증을 실시하기 위해 ASOS의 81개 지점 중, 자료의 활용도가 높은 64개 지점을 선정하였다(Fig. 1). Table 1에 ASOS 64개 지점의 자세한 정보를 나타내었다.

### 2.2 TRMM 3B43 월 누적 강우

TRMM은 미국의 National Aeronautics and Space Administration (NASA)와 일본의 Japan Aerospace Exploration Agency (JAXA)의 공동 프로젝트로 강우와 관련된 기상에 관한 연구를 목적으로 계획되었다. 열대 강우량 관측 위성인 TRMM은 1998년부터 지금까지 적도 주변 영역(50°S~50°N)의 열대지방과 아열대지방의 강우량을 파악한다. TRMM Multisatellite Precipitation Analysis (TMPA)를 통해 제공되는 TRMM 산출물들은 다양한 위성(SSM/I, AMSR, AMSU) 및 TRMM 위성 자료(2B31, 2A12), 지상 관측 장비로 측정된 강우 자료를 통합하여 매 관측 시간마다 격자별로 가장 적합한 강우량을 산출한다(Huffman et al., 2007).

본 연구에서 사용된 TRMM 3B43 V7 자료는 기존의 TRMM 3B43 V6 자료와 비교하여 Global Precipitation Climatology Centre 등의 자료와 방법들이 추가되어 개선되었다. TRMM 3B43 V7의 월 강우자료는 25 × 25 km의 공간해상도를 가지고 있으며, 관측 범위는 경도 180°W-180°E에서 위도 50°S-50°N로 전 세계의 대부분을 포함하고 있다. TRMM 3B43 자료는 위경도 좌표계(Geographic

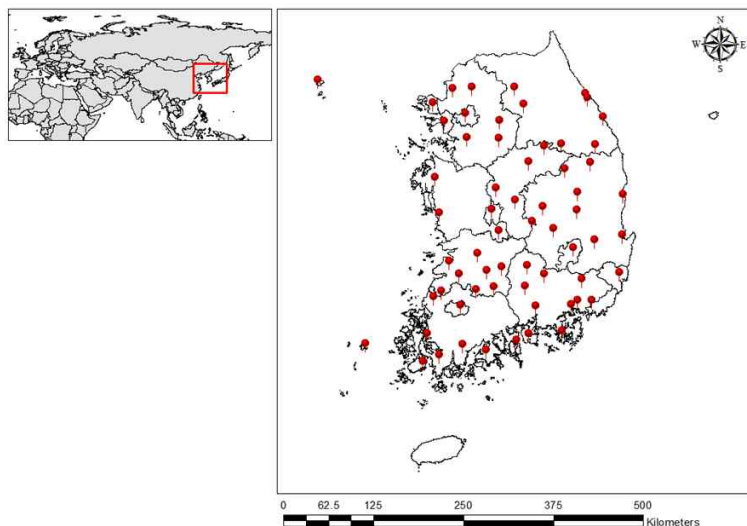


Fig. 1. Study Area

Table 1. Description of 64 Rain Gauge Stations

No.	Station	Coordinate		No.	Station	Coordinate	
		Long.	Lat.			Long.	Lat.
1	Baengnyeongdo	124.630	37.966	33	Geumsan	127.482	36.106
2	Chuncheon	127.736	37.903	34	Pohang	129.380	36.033
3	Dongducheon	127.061	37.902	35	Yeongcheon	128.951	35.977
4	Paju	126.766	37.886	36	Daegu	128.619	35.885
5	Bukgangneung	128.855	37.805	37	Jeonju	127.155	35.822
6	Gangneung	128.891	37.751	38	Buan	126.717	35.730
7	Ganghwa	126.446	37.707	39	Geochang	127.909	35.667
8	Hongcheon	127.880	37.684	40	Jangsu	127.520	35.657
9	Daegwallyeong	128.718	37.677	41	Imsil	127.286	35.612
10	Seoul	126.966	37.571	42	Hapcheon	128.170	35.565
11	Donghae	129.124	37.507	43	Jeongeup	126.866	35.563
12	Yangpyeong	127.494	37.489	44	Ulsan	129.320	35.560
13	Incheon	126.625	37.478	45	Miryang	128.744	35.491
14	Suwon	126.985	37.272	46	Sancheong	127.879	35.413
15	Icheon	127.484	37.264	47	Namwon	127.397	35.402
16	Yeongwol	128.457	37.181	48	Sancheong	127.129	35.371
17	Taebaek	128.989	37.170	49	Gochang	126.599	35.349
18	Jecheon	128.194	37.159	50	Younggwang	126.478	35.284
19	Chungju	127.953	36.970	51	Gimhae	128.893	35.227
20	Bongwhoa	128.915	36.944	52	Bukchangwon	128.673	35.226
21	Yeongju	128.517	36.872	53	Gwangju	126.892	35.173
22	Seosan	126.494	36.777	54	Changwon	128.573	35.170
23	Cheongju	127.441	36.639	55	Jinju	128.040	35.164
24	Andong	128.707	36.573	56	Tongyeong	128.436	34.845
25	Yeongdeok	129.409	36.533	57	Mokpo	126.381	34.817
26	Boeun	127.734	36.488	58	Namhae	127.926	34.817
27	Sangju	128.157	36.408	59	Yeosu	127.741	34.739
28	Daejeon	127.372	36.372	60	Jangheung	126.919	34.689
29	Uiseong	128.689	36.356	61	Heuksando	125.451	34.687
30	Boryeong	126.557	36.327	62	Goheung	127.276	34.618
31	Chupungnyeong	127.995	36.220	63	Haenam	126.569	34.554
32	Gumi	128.321	36.131	64	Jindo	126.324	34.472

latitude and longitude) 형태로 되어 있으며, 연구지역에 해당하는 적합한 강우 자료만을 추출하기 위해 한반도 지역에 대하여 Subset 하였다. TRMM 3B43 V7 자료는 Goddard Earth Sciences Data and Information Services Center (GES DISC)에서 HDF 및 Netcdf 파일로 제공되고 있다(<http://disc.sci.gsfc.nasa.gov/>). TRMM 3B43 V7 자료에 대한 자세한 정보 및 특징은 Table 2에서 확인 할 수 있다.

### 2.3 MODerate-resolution Imaging Spectro-radiometer (MODIS) product

MODIS는 NASA에서 육상 및 대기 자료를 관측하기 위한 Terra 위성과 해양 자료를 관측하기 위한 Aqua 위성에 탑재된 분광계이다. 이 센서는 0.4~14.4 $\mu$ m의 파장영역을 가진 36개의 밴드로 구성되어 있고, 250 m, 500 m 및 1,000 m의 다양한 공간해상도를 갖는 자료를 제공 한다(Barnes

Table 2. Characteristics of Version 7 TRMM 3B43 Monthly Data

Tropical Rainfall Measuring Mission 3B43 V7	
Geographic Coverage	Global tropics (50°N-50°S, 180°W-180°E) nlat=400, nlon=1440
Time Period	December 1997-Present
Temporal and Spatial Resolution	0.25 × 0.25 degree, monthly
File format	HDF and Netcdf

et al., 1998). 본 연구에서는 SVM 방법을 이용하여 상세화 하기 위해 MODIS 산출물 중 MOD11에서 제공하는 지표 먼온도(Land Surface Temperature, LST), MOD13에서 제공하는 NDVI, MOD07의 가능강수수분량(Precipitable Water Content, PWC)과 이슬점온도(Dew Point Temperature,  $T_d$ )를 사용하였다. LST는 MOD11A1 자료에 포함되어 있으며 1 km의 공간해상도를 가진 MODIS의 31번 밴드(10.8  $\mu$ m)와 32번 밴드(12.0  $\mu$ m)를 Split-window 방법을 활용하여 산정된다(Baek et al., 2012). NDVI의 경우, MOD13A2 자료에 포함되어 있으며 1 km의 공간해상도를 가진 16일 단위의 NDVI를 사용하였다. PWC와  $T_d$  자료는 MOD07L2에 포함되어 있으며 5 km의 공간해상도를 가지고 있다.

TRMM 3B43 V7 자료와 동일한 공간해상도를 가지게 하기 위하여 LST와 NDVI자료는 1 × 1 km에서 25 × 25 km로 upscaling 되었으며, 5 km의 해상도를 가지는 PWC와  $T_d$ 도 25 × 25 km로 upscaling을 실시하였다. 또한, 두 인공위성이 다른 좌표체계를 가지고 있으므로 기하보정 및 re-projection을 통하여 오차가 없도록 TRMM 3B43 V7 자료와 같은 좌표체계로 변경하였다.

### 3. 연구방법

#### 3.1 SVM-Regression

SVM은 여러 분류(class)에 속하는 데이터들을 최대한의 거리(margin)로 분리하는 최적의 초평면(hyperplane)을 찾는 방식으로 구동되며, 이는 Vapnik and Chervonenkis (1974), Vapnik (1982, 1955) 등에 의해 구체적으로 정립되었다(Smola and Scholkopf, 2004; Foody, 2004). 주어진 자료 집합  $\{(x_i, y_i)\}_i^n$ 에서 SVM-regression은

$$f(x) = \langle w, x \rangle + b \quad \text{with } w \in X, b \in R \quad (1)$$

를 통해 주어진  $y$ 값을 최적으로 근사하는 회귀함수  $f(x)$ 를 찾는다. Eq. (1)의  $w$ 와  $b$ 는 Eq. (2)와 같은 최적화 문제의 변환에 의해 추정될 수 있다.

$$\text{minimize } \frac{1}{2} \| w \|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*)$$

$$\text{subject to } \begin{cases} y_i - \langle w, x_i \rangle + b \leq \epsilon + \xi_i \\ \langle w, x_i \rangle - b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (2)$$

여기서,  $C$ 는 경험적 오차와 일반화 항인  $\frac{1}{2} \| w \|^2$  사이의 절충점을 결정하며, 경험적 오차 추정에 사용되는 일반화 상수를 의미하며  $C$ 값의 증가는 전체 오차 내에 경험오차의 상대적 비중이 증가하는 것을 나타낸다.  $\epsilon$ 은 Epsilon tube로 오차의 허용 범위를 의미하며, 너무 작은  $\epsilon$ 은 회귀 모형의 과적합을 유발하게 된다.

Eq. (3)은 Lagrange multipliers와 최적 제약조건을 이용하여 표현된 Eq. (1)의 회귀함수의 최종적인 형태이다.

$$f(x, a_i, a_i^*) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b \quad (3)$$

여기서,  $K(x, x_i)$ 는 커널함수(Kernel Function)로 정의되며, 훈련자료(Training Data)의 표본들을 내적 형태로 표현하여 표본 간의 비선형적 상호관계를 효과적으로 규명한다. SVM의 이론적 배경은 Vapnik (1995), Smola and Scholkopf (2003), Kwon et al. (2012), Hwang et al. (2007) 등의 기존연구에서 보다 자세히 다루어졌다.

본 연구에 사용된 SVM은 libSVM으로 Chang and Lin (2011)에 의해 구축되었다 (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>). libSVM 내에 Radial Basis Function (RBF) 커널함수를 사용한 epsilon-Regression 통해 회귀를 수행하였으며, 해당 회귀 모델을 결정하는 변수인  $C, \gamma, \epsilon$ 의 조정을 통해 월별 회귀모형을 최적화하였다.  $\gamma$ 는 RBF 커널에 사용되는 변수이다.

회귀 모델 구축에 독립변수로 사용된 인자들은 NDVI, LST, PWC,  $T_d$ 과 각 픽셀의 2차원적 위치로 총 6개로 구성되어 있으며 종속변수인 강우에 대해 예측하도록 한다 (Table 3). SVM-regression의 정확도를 최대화시키기 위해 회귀 모델의 구축은 데이터 집합 각각에 대해 월별로 시행되었으며 training data와 predicting data 모두 predicting data의 인자별 값의 범위에 따라 [0, 1]로 표준화하여

각 인자의 양적 단위 차이에서 발생할 수 있는 오류를 최소화 하였다. 이에 따라 표준화 된 회귀모델의 예측 값 또한 역으로 복원하는 과정을 거쳐 유효한 강우 값을 갖도록 계산하였다.

#### 4. 연구결과 및 토의

##### 4.1 상세화 강우와 지점자료 간 비교

Table 4는 SVM-Regression 모델을 결정하는 변수들의 값을 월별로 나타낸 것이다. 모델 결정 변수들은 4월, 9월, 12월을 제외한 달에 대해서 대체적으로 비슷한 범위를 나타냈다. 비슷한 범위의 모델 결정 변수들은 해당 월에 각 인자들과 강우 간 관계성이 유사한 것을 의미한다. 세부적으로는 모델 결정 변수들이 계절적 특성에 의해 분류되는 것을 확인할 수 있다(Table 4). 연중 강수량이 집

중적으로 분포하고 식생 활력이 높게 유지되는 5월에서 8월의 모델과 강수량이 상대적으로 적고 기온과 식생이 낮은 1월에서 3월의 모델은 각각 유사성을 띤다. 따라서 앞서 언급한 4월, 9월, 12월은 식생 활력에 대한 급격한 변화 등의 계절적 변동이 발생하는 시기이기 때문에 다른 달과 유사하지 않은 모델 특성을 가지는 것으로 판단된다 (Fig. 2).

Table 5는 SVM 회귀를 통한 강우자료 상세화의 적합성을 파악하기 위해 상세화 이전 강우 및 이후 강우와 지점 관측 강우와의 상관성을 월별로 분석한 결과이다. 지점 관측 강우와의 오차를 나타내기 위한 지표로는 Root Mean Square Error (RMSE), Bias, Index Of Agreement (IOA)를 사용하였다. IOA는 모형의 오차를 0에서 1 사이의 값으로 표준화하여 나타내며 1은 가장 큰 오차를 의미한다(Wilmott, 1981). TRMM 강우와 지점 관측 자료 간

Table 3. Characteristics of Input Variables of SVM-Regression Model

Dependent Variable y	Independent Variable x						Spatial Resolution	
	NDVI	LST	PWC	T <sub>d</sub>	longitudinal position	latitudinal position	Training Data	Predicting Data
Monthly cumulative rainfall (mm/hr)	Monthly average (Unitless)	Monthly average (°K)	Monthly cumulative (mm/hr)	Monthly average (°K)	Per pixel (Unitless)	Per pixel (Unitless)	25 km	1 km

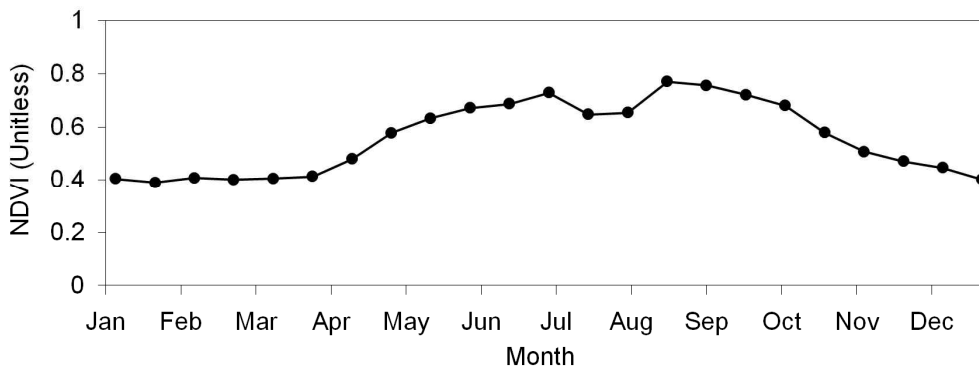


Fig. 2. Mean Value of 16-day NDVI

Table 4. Parameters of SVM-Regression Models

Month	SVM-Regression Parameters			Month	SVM-Regression Parameters		
	C	$\gamma$	$\epsilon$		C	$\gamma$	$\epsilon$
1	186	0.2	0.09	7	193	0.21	0.09
2	188	0.21	0.09	8	193	0.21	0.09
3	189	0.2	0.09	9	320	0.75	0.07
4	800	0.5	0.07	10	175	0.35	0.09
5	191	0.22	0.09	11	105	0.21	0.09
6	191	0.22	0.09	12	970	0.72	0.09

Table 5. Comparison of Statistics between Satellite Measurements and Rain Gauge Measurements ( $y= ax+b$ ,  $y$ : TRMM Precipitation,  $x$ : Rain gauge Precipitation)

Month	TRMM 3B43 V7 (25 km)						TRMM Downscaled (1 km)					
	Bias (mm/hr)	RMSE (mm/hr)	a	b	R	IOA	Bias (mm/hr)	RMSE (mm/hr)	a	b	R	IOA
1	0.008	0.010	0.779	0.011	0.750	0.762	-0.004	0.009	0.667	0.010	0.681	0.783
2	0.014	0.022	0.719	0.030	0.748	0.801	-0.005	0.014	0.576	0.035	0.787	0.805
3	0.012	0.024	0.286	0.062	0.414	0.563	-0.018	0.023	0.467	0.033	0.658	0.765
4	0.008	0.032	0.503	0.049	0.678	0.783	0.005	0.028	0.758	0.035	0.785	0.850
5	0.011	0.036	0.699	0.065	0.654	0.780	0.000	0.022	0.507	0.099	0.558	0.710
6	0.030	0.056	0.544	0.112	0.696	0.741	-0.007	0.035	0.540	0.107	0.669	0.763
7	0.052	0.165	0.542	0.357	0.637	0.770	-0.030	0.115	0.419	0.424	0.609	0.720
8	0.021	0.073	0.766	0.073	0.862	0.913	0.044	0.076	0.638	0.147	0.763	0.780
9	0.000	0.021	0.502	0.034	0.583	0.719	-0.002	0.023	0.422	0.039	0.546	0.576
10	0.027	0.043	0.536	0.048	0.500	0.565	-0.013	0.042	0.571	0.036	0.658	0.657
11	-0.002	0.044	0.360	0.046	0.566	0.702	0.020	0.036	0.514	0.057	0.624	0.742
12	0.018	0.021	0.989	0.019	0.748	0.625	-0.008	0.025	0.996	0.002	0.588	0.700
mean	0.016	0.046	0.602	0.076	0.644	0.727	-0.002	0.037	0.589	0.085	0.667	0.738

상관계수는 모든 달에 대해 유의수준 0.05 내에서 유의하게 나타났다(Table 5). 상세화 된 강우는 지점 관측 강우와 상세화 이전의 강우에 비해 평균적으로 약간 높은 상관성을 나타냈다. 월별로는 3월과 10월의 상관성이 특히 향상된 것을 확인할 수 있으나 일부 달에서는 상관성이 다소 감소한 것을 확인할 수 있다. 지점 관측 자료와의 평균적 오차를 나타내는 RMSE의 경우 모든 달에서 상세화된 강우가 25 km 강우에 비해 감소하는 경향을 보였으며 모형에 대한 표준화된 평가 지수인 IOA는 월별로 차이는 있으나 평균적으로 상세화된 강우에서 높게 나타났다(Fig. 3). TRMM 강우는 평균적으로 양의 Bias를 나타낸 것에 비해, 상세화 된 강우는 매우 작은 음의 Bias를 나타냈다. TRMM 자료가 가지는 전반적인 양의 Bias는 주로 강우가 많이 발생하는 6월에서 8월 사이에서 다른 달 보다 비교적 크게 나타나며 이는 TRMM 자료와 지점 관측 강우 비교에 관한 연구에 나타난 양상과 부분적으로 일치한다(Duan et al., 2012). 상세화 기법의 적용을 통해 이와 같은 과대산정의 양상은 6월과 7월에 대해 작은 Bias를 가지는 값으로 보정되었다. 한편 같은 기간에 대한 다른 통계 지표인 상관계수는 상세화 이후 오히려 다소 감소하는 양상을 보였다. 이는 TRMM이 많은 양의 강우에 대해 상당히 높은 상관성을 가지는 측정값을 제공하며, 다른 기상인자의 고려가 강우가 지배적인 인자인 해당시기에는 다소 불확실성을 증가시킨 것으로 판단된다.

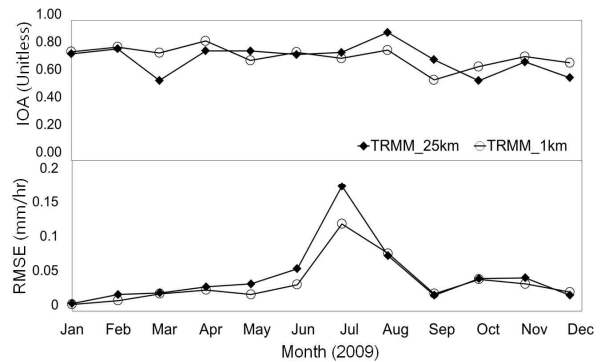


Fig. 3. Time Series of Monthly RMSE and IOA

#### 4.2 상세화 강우의 공간분포

SVM 회귀를 통해 상세화 된 TRMM 3B43 자료에 대한 공간분포 Mapping은 월별로 구축되었으며, 상세화 전과 후의 공간분포에 대한 비교 분석을 위해 Fig. 4에 월별 공간 분포를 도시하였다. SVM-Regression을 이용한 상세화 기법의 적용을 통해 강우에 대한 공간분포는 시각적으로 확인 가능할 만큼 세분화 되었으며 전반적으로 상세화 이전의 격자 강우 공간 분포와 비슷한 양상을 보였다. 상세화 이후 연구지역 내에서 최저와 최고 누적 강우 분포가 나타난 위치가 다소 변한 달이 있으나 상세화 전과 후의 공간적 분포 양상은 유사하게 나타났다.

1월, 3월, 6월, 9월, 12월의 경우 다른 달에 비해 모든 지역에 대하여 과대산정 되었던 강우가 상세화 과정을 통해

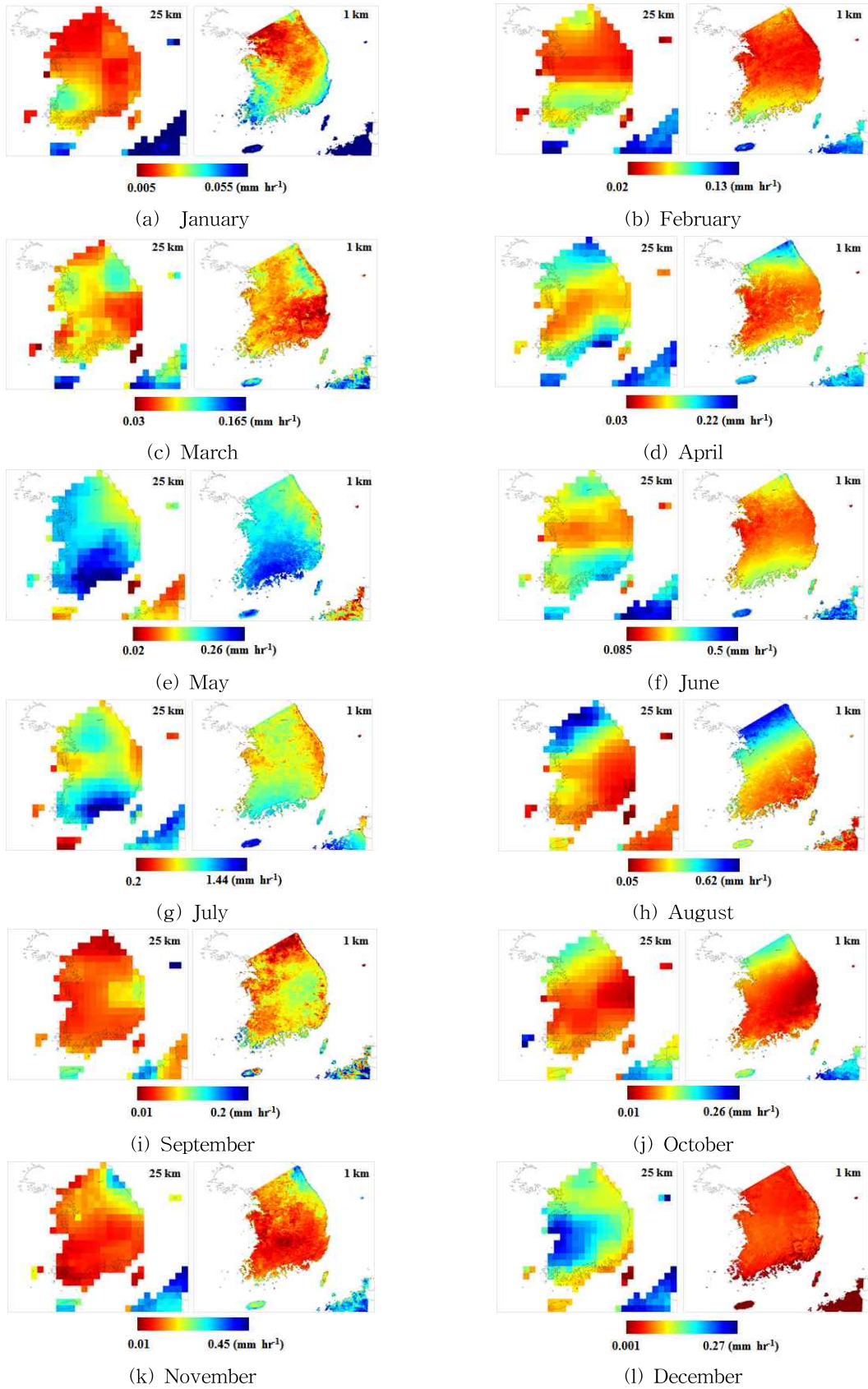


Fig. 4. Spatial Distribution of TRMM 3B43 V7 Monthly Precipitation and Downscaled Monthly Precipitation



전반적으로 감소하였으며, 이는 앞선 4.1절에 서술된 통계적 분석을 통해 확인된 바 있다. 2월의 강우분포는 상세화 전후에 같은 범위 내에 분포하였으나, 상세화 이전 남부 지방에 강우가 전반적으로 높게 분포한 것에 비해 상세화 이후에는 보다 낮은 위도에서 높은 강우분포가 나타났다. 또한, 중부지방의 강우분포 역시 상세화 이후 평균적으로 더 적은 값을 갖게 되었다. 이와 같은 공간분포의 변화가 일어난 지역(동두천, 파주, 부안, 거창 등)의 Bias를 상세화를 통하여 감소시켜 2월 강우의 정확성 향상에 기여하였다. 4월에는 강우의 공간분포에 나타난 최대치가 증가하여 보다 넓은 범위 내에 강우가 분포하는 양상을 보였다. 이 같은 변화로 인해 지점강우와의 상관분석에서 다소 낮았던  $a$ 값이 개선되었으며 상관계수 역시 증가하였다(Table 5). 11월의 평균적인 공간분포 양상은 상세화 전과 후가 비슷한 가운데 주로 중부지방의 강우량이 세분화되었다. 특히 파주, 동해, 동두천 등 지점에서 과소산정되었던 강우가 개선된 것을 확인할 수 있다. 5월, 7월, 8월, 10월에는 상세화 전후의 격자 강우 공간분포의 범위와 양상이 매우 비슷하게 유지되었으며, 지상 관측 강우와의 상관성 및 평균적 오차에도 큰 변동이 발견되지 않았다. 이와 같이 월별 격자 강우의 상세화로 인한 공간분포의 변화는 지점 관측 강우와의 상관성 개선에 기여하며 25 km 격자 강우가 가지는 일부 지점에 대한 과대 및 과소 산정의 오차를 효과적으로 감소시키는 것을 확인할 수 있었다.

## 5. 결 론

본 연구에서는 TRMM 3B43 V7 강우자료의 공간해상도 개선을 위해 강우 발생과 높은 상관성을 지니는 인자들과 강우 간의 관계성을 기반으로 한 상세화 모형을 제안하였다. 해당 과정에 사용된 인자들은 NDVI, LST, PWC, Td이며 각 픽셀의 위치도 회귀 모델 구축을 위하여 함께 고려되었다. 여기서 다양한 변수들과 강우 간 관계성을 효과적으로 고려하여 상세화된 공간해상도의 강우를 예측하기 위해 비선형 예측모형인 SVM을 활용하였다.

식생 활력이 다른 달에 비해 크게 변화하는 달의 강우와 입력변수 간의 상관성은 다른 달과 확연한 차이를 나타냈으며, 나머지 기간은 계절별로 유사한 특성을 가졌다. 상세화 된 격자 강우는 지점 관측자료와의 비교 분석에서 상세화 이전의 25 km 격자 강우에 비해 상관계수 향상과 RMSE, IOA 및 Bias의 개선을 보였으며 전반적으로 높은 정확도를 지니는 것으로 나타났다. 특히 상세화 이전

TRMM 3B43 V7 (25 km) 강우가 나타냈던 양의 Bias는 상세화 이후 매우 작은 음의 Bias로 전환되어 전반적인 과대 산정의 경향이 효과적으로 제거된 것을 확인하였다. 상세화 전후의 공간분포 비교에서 두 공간분포는 전반적으로는 유사한 양상을 나타냈다. 대부분의 월별분포에서 강우 분포의 범위가 전반적으로 감소함으로써 실제 지점 관측 강우에 대해 개선된 정확성을 보였으며, 1 km 격자 강우가 25 km 격자 강우에서는 감지되지 않았던 강우의 국지적 특성을 묘사하여 일부 25 km 픽셀에서 나타난 과대 및 과소 산정에 의한 오차를 감소시켰다.

본 연구를 통해 제안된 격자강우의 상세화 기법은 수문 모형의 정확도 개선을 위한 고해상도의 입력자료 구축에 적용될 수 있으며, 강우의 국지적 분포를 보다 효과적으로 파악할 수 있는 강우분포 Mapping을 가능하게 한다. 향후 연구에서는 SVM-Regression 외에 다른 회귀분석 방법을 함께 적용하여 보다 개선된 상세화 기법을 개발할 수 있을 것으로 판단된다.

## 감사의 글

본 연구에 사용된 ASOS 강우량 자료를 제공해주신 기상청, TRMM 강우자료를 제공해주신 NASA Goddard Space Flight Center 및 MODIS 자료를 제공해주신 NASA Land Processes Distributed Active Archive Center에 감사의 말을 전합니다. 본 연구는 한국건설기술연구원 주요(시드) 사업(지역특성을 반영한 상세 격자강우량 생산기술 개발)의 연구비 지원에 의해 수행되었습니다.

## References

- Anandhi, A., Srinivas, V.V., Nanjundiah, R.S., and Kumar, D.N. (2008). "Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine." *International Journal of Climatology*, Vol. 28, pp. 401-420.
- Baek, J.J., and Choi, M.H. (2012). "Availability of Land Surface Temperature from the COMS in the Korea Peninsula." *Journal of Korean Water Resources Association*, Vol. 45, No. 8, pp. 755-765.
- Bardossy, A. (1997). "Downscaling from GCMs to local climate through stochastic linkages." *Journal of Environmental Management*, Vol. 49, No. 1, pp. 7-17.
- Barnes, W.L., Pagano, T.S., and Salomonson, V.V. (1998)

- “Prelaunch Characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1.” *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 36, No. 4, pp. 1088-1100.
- Beckmann, B., and Buishand, A. (2002) “Statistical downscaling relationships for precipitation in the Netherlands and north Germany.” *International Journal of Climatology*, Vol. 22, pp. 15-32.
- Ben-David, S., and Lindenbaum, M. (1997). “Learning distributions by their density levels: a paradigm for learning without a teacher.” *Journal of Computer and System Science*, pp. 171-182.
- Cavazos, T. (1999) “Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas.” *Journal of Climate*, Vol. 12, pp. 1506-1523.
- Chang C.-C., and Lin, C.-J. (2011). “LIBSVM: a library for support vector machines.” *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 3, pp. 1-27.
- Charles, S.P., Bates, B.C., Whetton, P.H., and Hughes J.P. (1999) “Validation of downscaling models for changed climate conditions: case study for southwestern Australia.” *Climate Research*, Vol. 12, pp. 1-14.
- Chen, S.-T., Yu, P.-S., and Tang Y.-H. (2010). “Statistical downscaling of daily precipitation using support vector machines and multivariate analysis.” *Journal of Hydrology*, Vol. 385, pp. 13-22.
- Cho, K. (2013). *Flood runoff simulation using MIKE SHE and SVM in the Chungju Dam Basin*. Master’s Thesis, Kyunghee University, Seoul, Korea.
- Chu, J.-L., Kang, H., Tam, C.-Y., Park, C.-K., and Chen, C.-T. (2008). “Seasonal forecast for local precipitation over northern Taiwan using statistical downscaling.” *Journal of Geophysical Research*, Vol. 113, D12118, doi:10.1029/2007JD009424.
- Creutin, J.D., Delrieu, G., and Lebel, T. (1988). “Rain measurement by raingage-radar combination: a geostatistical approach.” *Journal of Atmospheric and Oceanic Technology*, Vol. 5, pp. 102-115.
- Duan, Z., Bastiaanssen, W.G.M., and Liu, J. (2012) “Monthly and annual validation of TRMM Multisatellite Precipitation Analysis (TMPA) products in the Caspian Sea Region for the period 1999-2003.” *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, Munich, Germany, pp. 3696-3699.
- Foody, G.M. (2004). “A relative evaluation of multiclass image classification by Support Vector Machines.” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 6, pp. 1335-1343.
- Gunn, S. (1998). *Support Vector Machines for classification and regression*. Image Speech and Intell. Syst. Group, Dept. Elect. Comput. Sci., Univ. Southampton, Southampton, U.K., Tech. Rep.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., and Wolff, D.B. (2007). “The TRMM Multisatellite Precipitation Analysis (TMPA): Qasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scale.” *Journal of Hydrometeorology*, Vol. 8, No. 1, pp. 38-55.
- Hwang, S.H., Kim, J.H., and Jung, S.W. (2007). *Hydrologic time series forecasting using SVM*. Korea Water Resources Association Annual Conference, 2007, Pyeongchang, Korea, pp. 1972-1976.
- Immerzeel, W.W., Rutten, M.M., and Droogers, P. (2009) “Spatial Downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula.” *Remote Sensing of Environment*, Vol. 113, pp. 362-370.
- Jia, Sh., Zhu, W., Ju, A., and Yan T. (2011) “A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in Qaidam Basin of China.” *Remote Sensing of Environment*, Vol. 114, pp. 3069-3079.
- Keramitsoglou, I., Kiranoudis, C.T., and Weng, Q. (2013). “Downscaling geostationary land surface temperature imagery for urban analysis.” *IEEE Geoscience and Remote Sensing Letters*, Vol. 10, No. 5, pp. 1253-1257.
- Kwon, H.H., Kim, M.J., and Kim, O.G. (2012) “A development of water demand forecasting model based on wavelet transform and support vector machine.” *Journal of Korean Water Resources Association*, Vol. 45, No. 11, pp. 1187-1199.
- Olsson, J., Uvo, C.B., and Jinno, K. (2001) “Statistical atmospheric downscaling of shortterm extreme rainfall

- by neural networks.” *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, Vol. 26, No. 9, pp. 695–700.
- Osuna, E., Freund, R., and Girosi, F. (1997). *An improved training algorithm for Support Vector Machines*. IEEE NNSP 1997, Amelia Island, FL, pp. 276–285.
- Park, N.W. (2013). “Spatial Downscaling of TRMM Precipitation Using Geostatistics and Fine Scale Environmental Variables.” *Advances in Meteorology*, Vol. 2013, Article ID 237126, 9 pages, doi:10.1155/2013/237126.
- Smola, A.J., and Scholkopf, B. (2004). “A tutorial on support vector regression.” *Statistics and Computing*, Vol. 14, pp. 199–222.
- Su, Fe., Hong, Y., and Lettenmaier, D.P. (2008). “Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin.” *Journal of Hydrometeorology*, Vol. 9, pp. 622–640.
- Vapnik, V.N. (1982). *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, U.S.A.
- Vapnik, V.N. (1995). *The nature of statistical learning theory*. Wiley, NY, U.S.A.
- Vapnik, V.N., and Chervonenkis, A.Y. (1971). “On the uniform convergence of relative frequencies of events to their probabilities.” *Theory of probability and its applications*, Vol. 16, No. 2, pp. 264–281.
- Wang, X. (2012). “Study on Genetic Algorithm Optimization for Support Vector Machine in Network Intrusion Detection.” *Advances in Information Sciences and Service Sciences*, Vol. 4, No. 2, pp. 282–288.
- Willmott, C.J. (1981). “On the validation of models.” *Physical Geography*, Vol. 2, pp. 184–194.
- Yin, Z.Y., Zhang, X., Liu, X., Colella, M., and Chen, X. (2008). “An Assessment of the biases of satellite rainfall estimates over the Tibetan Plateau and correction methods based on topographic analysis.” *Journal of Hydrometeorology*, Vol. 9, pp. 301–326.

paper number : 14-087

Received : 5 September 2014

Revised : 13 October 2014 / 3 November 2014

Accepted : 3 November 2014