기계학습 모형과 잔차 크리깅을 이용한 저해상도 위성기반 강수 자료의 공간 상세화

김예슬 · 박노옥

인하대학교 공간정보공학과

(2016년 5월 27일 접수, 2016년 6월 8일 수정, 2016년 6월 28일 게재 확정)

Spatial Disaggregation of Coarse Scale Satellite-based Precipitation Data using Machine Learning Model and Residual Kriging

Yeseul Kim · No-Wook Park

Department of Geoinformatic Engineering, Inha University

(Received 27 May 2016, Revised 8 June 2016, Accepted 28 June 2016)

Abstract : A novel disaggregation model that combines a machine learning model and kriging of residuals is presented to map precipitation at a fine scale from coarse scale precipitation data. Random forest (RF) and fine scale auxiliary variables are used to estimate trend components at a fine scale. Residual components are then estimated by area-to-point residual kriging. A case study of spatial disaggregation of TRMM monthly precipitation data acquired over the Korean peninsula is carried out to illustrate the potential of the presented disaggregation method. From the evaluation results, the presented method outperformed the RF-based disaggregation method that only considers trend components and ignores residual components, in terms of accuracy statistics and the ability of coherent predictions. This case study indicates that accounting for residual components by applying a proper spatial prediction method such as area-to-point kriging is very important in spatial disaggregation of coarse scale spatial data, even though advanced regression models such as RF could have high goodness of fit for the quantification of relationships between a target attribute and auxiliary variables.

Key Words : disaggregation, precipitation, machine learning, kriging

주요어 : 상세화, 강수, 기계학습, 크리깅

1. 서론

최근 고온과 한파, 해수면 증가 및 폭우 등 전 세계적으로 이상 기후 현상이 발생하고 있는데(Lati-

fovic and Pouliot, 2007; Yuan *et al.*, 2011; Yang *et al.*, 2013), 이러한 기후 변화는 생태계에 직간접적으로 영향을 미칠 수 있다(Brown *et al.*, 2012; Atzberger, 2013). 앞으로 이러한 추세가 가속화될 것

을 고려할 때, 기후 시스템의 모니터링과 변화 분석은 매우 중요할 것으로 예상된다.

기온, 강수 등의 기후 인자들은 공간적으로 산재된 기상관측소의 관측 자료를 이용하여 주제로 제작되고, 다양한 모델링의 입력 자료로 사용되고 있다. 그러나 지상 관측 자료를 이용한 주제도 제작은 다량의 관측 자료를 필요로 하며, 관측소가 공간적으로 균질하게 분포해야 한다(Back and Jang, 2011; Park and Jang, 2013; Jang *et al.*, 2015). 이러한 한계와 관련하여 기온, 강수 주제도 작성을 위해 광역 정보를 제공하는 원격탐사 자료가 활용될 수 있다. 주기성과 광역성을 갖는 위성 기반 원격탐사 자료는 기후 인자의 시공간적 변이 관측이 용이하기 때문에 기온, 강수, 토양 수분 등의 모니터링과 분석에 많이 활용되고 있다(Iguchi *et al.*, 2000; Huffman *et al.*, 2007).

최근 센서 기술의 발달로 다양한 해상도의 위성 자료의 이용이 가능해졌는데, 위성 기반 기후 인자 정보는 상대적으로 높은 시간 해상도에 비해 낮은 공간 해상도로 제공되고 있다. 강수 정보를 제공하는 Tropical Rainfall Measuring Mission(TRMM) 자료 등은 0.25°의 공간 해상도로 제공된다. 이러한 낮은 공간 해상도는 전 지구 규모나 광역 규모의 모니터링에는 적합하지만, 우리나라와 같이 국지적 기후 특성을 나타내는 지역의 상세 분석에는 적용에 한계가 있다.

다수의 지상 관측 자료 이용이 불가능하여 지역 분석에 저해상도 위성기반 기후 인자 자료를 이용하는 경우, 저해상도 자료로부터 고해상도에서의 속성값을 추정하는 상세화(disaggregation, downscaling)를 적용할 필요가 있다(Bechtel *et al.*, 2012; Atkinson, 2013). 저해상도 자료의 상세화 과정에 고해상도에서의 세부 변동성을 고려하기 위해 기후 인자와 관련성이 있는 고해상도 환경 변수를 사용할 수 있다(Agam *et al.*, 2007; Guan *et*

al., 2009; Rahman *et al.*, 2009). 기존 연구에서는 저해상도 강수 자료의 상세화를 목적으로 상대적으로 고해상도에서 획득된 식생지수와 고도 등의 부가 자료를 이용하여 상세 변동성을 나타내는 고해상도 강수 주제도를 작성해왔다(Jia *et al.*, 2011; Park, 2013; Chen *et al.*, 2014).

고려하고 있는 저해상도 속성과 부가 자료와의 연관성을 정량적으로 추정하기 위해 선형 회귀 모형이 주로 적용되었는데, 다양한 비선형 회귀 모형도 적용이 가능하다. 저해상도 자료와 부가 자료 사이에 선형 관계성이 뚜렷하게 나타나지 않을 경우, 복잡한 비선형 관계를 고려하기 위해 최근 기계학습 모형을 이용한 상세화 연구가 진행되고 있다(Najafi *et al.*, 2011; Rodriguez-Galiano *et al.*, 2012; Srivastava *et al.*, 2013; Hutengs and Vohland, 2016).

고해상도 부가 자료와 기계학습 모형을 이용하는 기존 상세화 연구에서는 기계학습의 높은 변수 설명 능력을 이용하여 고해상도에서의 상세한 변동성을 예측한다. 그러나 기계학습 모형의 변수 설명 능력이 높더라도 오차를 포함하는 저해상도 자료에 과적합(over-fitting)하는 문제가 나타날 수 있고, 이러한 문제는 국소적으로 차이를 보일 수 있다. 그 결과 부가 자료의 설명 능력이 충분하게 반영되지 않은 지역에서 상세 변동성 예측 능력이 저하될 수 있으며, 이는 전체 상세화 결과의 예측 정확도에 영향을 줄 수 있다. 따라서 상세화 과정에서 부가 자료로부터 상세 변동성이 충분하게 설명되지 않는 지역을 고려할 필요가 있다. 부가 자료와 저해상도 자료의 통합을 위하여 기계학습 모형을 포함한 회귀 모형을 적용하면, 부가 자료로 설명되지 않는 성분인 잔차(residual)를 계산할 수 있다. 저해상도 자료의 상세 변동성을 설명하는 과정에 잔차를 고려할 경우, 고해상도 부가 자료로 설명되지 않는 부분을 보완함으로써 예측 정확

도를 향상시킬 가능성이 높아진다.

저해상도에서 적용된 회귀 모형의 결과로부터 얻어지는 잔차는 저해상도에서만 값을 얻을 수 있기 때문에 상세화 과정에서 고해상도에서의 잔차를 추정해야 한다. 이를 위해 저해상도에서 계산된 잔차를 대상으로 다양한 공간 예측 기법을 적용할 수 있다. 여러 공간 예측 기법 중에서 지구통계학적 크리깅은 잔차의 공간 상관성 정보를 고려할 수 있고, 크리깅에 사용하는 입력 자료와 예측 지점의 해상도 차이를 고려할 수 있는 장점이 있다 (Goovaerts, 1997). 특히 Kyriakidis(2004)는 다양한 형태의 저해상도 공간 자료로부터 상대적으로 고해상도에서의 값을 추정할 수 있는 영역-점 변환 크리깅(area-to-point kriging)을 제안하였다. 기존 기계학습 모형을 이용한 상세화에서 잔차의 공간 예측 결과와 회귀 모형의 결과를 결합하는 연구가 일부 진행되었지만(Shi *et al.*, 2015; Hutengs and Vohland, 2016), 잔차의 공간 상관성 정보와 해상도 차이를 직접적으로 고려한 연구는 드문 상황이다. 최근 Park(2013)이 선형 회귀 모형과 잔차 크리깅을 결합하였는데, 상대적으로 모형 적합도가 높은 기계학습 모형 등을 고려하지는 않았다.

이 연구에서는 기계학습 모형 기반의 상세화에

서 영역-점 변환 잔차 크리깅을 결합하는 상세화 방법론을 제안하였다. 제안 방법론은 고해상도 위치에서의 세부 변동성을 고려하기 위해 고해상도 부가 자료와 기계학습 모형을 통해 경향 성분을 추정하고, 모형으로 설명되지 않는 잔차의 추정에 영역-점 변환 크리깅을 적용한다. 제안 방법론의 적용 과정과 성능을 예시하기 위해 우리나라를 대상으로 저해상도 위성기반 강수 자료의 공간 상세화 사례 연구를 수행하였다.

2. 방법론

고해상도 부가 자료를 이용하는 상세화는 성분 분해를 기반으로 하는데, 원 저해상도 자료의 속성을 부가 자료로부터 추정한 경향 성분과 부가 자료로 설명될 수 없는 잔차 성분으로 분해한다. 이러한 성분 분해를 기반으로 이 연구에서 제안한 상세화 방법론도 경향 성분 추정과 잔차 성분 추정의 두 단계로 구성되며, 경향 성분 추정에는 기계학습 모형을 적용하고 잔차 성분 추정에는 영역-점 변환 크리깅을 적용한다(Figure 1).

우선 경향 성분 추정에는 여러 기계학습 모형 중

Figure 1. Work flows for the disaggregation procedure proposed in this study.

에서 랜덤 포레스트(Random Forest: RF)를 적용하였다. RF는 결정 트리(decision tree)의 앙상블 기법으로, 입력 자료를 임의로 분할하여 서로 상관관계가 없는 다중 결정 트리를 생성하여 분류 혹은 예측을 수행하는 기계학습 모형이다(Breiman, 2001). RF는 적용의 간편성, 다중 변수 사용의 용이성 등의 장점으로 저해상도 자료의 상세화에 많이 적용되고 있다(Shi *et al.*, 2015; Hutengs and Vohland, 2016). RF를 이용한 경향 성분 추정을 위해 고해상도 부가 자료를 저해상도로 업스케일링한 후에, 저해상도에서 RF 기반의 회귀 모형을 추정한다. 추정된 모형을 고해상도 부가 자료에 적용함으로써 고해상도에서의 경향 성분을 추정할 수 있다. 기존 연구에서는 잔차 성분을 고려하지 않고 고해상도 경향 성분 추정 결과를 상세화 결과로 제시한 경우가 많았다. 그래서 이 연구에서는 제안 기법과의 비교를 목적으로 고해상도 경향 성분을 상세화 결과로 간주하는 경우를 또 다른 상세화 방법론으로 고려하였다.

다음 과정으로 저해상도 잔차 성분으로부터 고해상도 잔차 성분을 추정한다. 저해상도에서 추정된 경향 성분과 원 자료로부터 잔차 성분을 계산한 후에, 영역-점 변환 크리깅을 적용하여 고해상도에서의 잔차 성분을 추정한다. 영역-점 변환 크리깅은 저해상도 자료의 가중 선형 결합을 통해 고해상도에서의 속성값을 예측한다(Kyriakidis, 2004). 대상 연구지역내 저해상도 픽셀(v_i)에서의 잔차 성분을 $R(v_i)$ 로 나타냈을 때, 고해상도 픽셀(u_j)에서의 잔차 성분($R(u_j)$)는 식 (1)의 영역-점 변환 정규 크리깅에 의해 예측할 수 있다.

$$R(u_j) = \sum_{k=1}^K \lambda_k(u_j) R(v_k) \quad (1)$$

여기서 $\lambda_k(u_j)$ 는 u_j 주변에 위치한 k 번째 저해상도 잔차 성분에 부여되는 정규 크리깅 가중치를 나타

낸다.

영역-점 변환 정규 크리깅 가중치는 식 (2)의 블록 정규 크리깅 시스템을 통해 계산된다(Kyriakidis, 2004; Goovaerts, 2008; Park, 2013).

$$\begin{cases} \sum_{k=1}^K \lambda_k(u_j) \bar{c}(v_k, v_i) = \bar{c}(v_k, u_j) \\ \sum_{k=1}^K \lambda_k(u_j) = 1 \end{cases} \quad (2)$$

여기서 $\bar{c}(v_k, v_i)$ 는 주변 저해상도 자료 사이의 공분산을 나타내며, $\bar{c}(v_k, u_j)$ 는 저해상도 자료와 예측 위치 사이의 공분산을 나타낸다.

식 (2)에서 $\bar{c}(v_k, v_i)$ 은 저해상도 자료를 이산화시킨 점들 사이의 공분산의 평균이며, $\bar{c}(v_k, u_j)$ 는 저해상도 자료를 이산화시킨 점들과 예측 위치 사이의 공분산의 평균이다(Goovaerts, 2008). 공분산은 2차 정상성의 가정하에 베리오그램(variogram)으로부터 계산될 수 있는데(Goovaerts, 1997), 식 (2)의 공분산을 계산하기 위해서는 점 기반 베리오그램이 필요하다. 그러나 잔차 성분은 저해상도에서만 계산되기 때문에 직접적으로 점 기반 베리오그램을 계산할 수 없다. 이러한 점 기반 베리오그램을 추정하기 위해 이 연구에서는 Goovaerts(2008)가 제안한 베리오그램 디컨볼루션(deconvolution)을 적용하였다. 일종의 역산에 해당하는 베리오그램 디컨볼루션은 우선 저해상도 자료의 베리오그램 모델링을 수행한 후에, 초기 점 기반 베리오그램 모델로부터 정규화시킨 베리오그램 값을 계산하여 이 값과 저해상도 자료의 베리오그램 값과 비교한다. 이러한 값 차이의 비교를 초기 점 기반 베리오그램 모델을 바꿔가면서 반복적으로 수행하여 최종적으로 값 차이가 가장 적은 점 기반 베리오그램 모델을 추정한다(Goovaerts, 2008).

최종 단계에서는 RF로부터 추정한 고해상도에서의 경향 성분과 영역-점 변환 크리깅을 통해 추정한 잔차 성분을 결합하여 상세화 결과를 얻는다.

3. 사례 연구

1) 사용 자료

이 연구에서 제안한 상세화 방법론의 적용 과정을 예시하기 위해 저해상도 TRMM 강수 자료의 상세화 사례 연구를 수행하였다. TRMM은 전 지구의 강수 형태와 변화의 모니터링을 목적으로 미국 National Aeronautics and Space Administration(NASA)와 일본 Japan Aerospace Exploration Agency(JAXA)의 공동 프로젝트의 일환으로 발사한 위성으로, 약 96분 주기로 강수 자료를 제공한다(Kummerow *et al.*, 1998). 이 연구에서는 시범적으로 2013년 4월과 7월의 TRMM 3B43 자료를 이용하였는데, 이 자료의 공간 해상도는 0.25° 로 3시간 주기로 관측된 강우 강도의 월 평균 자료로 제공된다. 정확도 검증 자료와의 직접 비교를 위하여 TRMM 강우 강도 자료를 TM 좌표계의 25km 해상도로 변환하였으며, 월 평균 강우 강도를 월 누적 자료로 변환하였다. 이러한 변환 과정을 거

친 TRMM 자료는 Figure 2에 제시하였는데, 비교적 건조한 날이 많은 4월에는 상대적으로 월 누적 강우량이 적고, 여름철 집중 호우로 인하여 7월에는 월 누적 강우량이 많게 나타난다.

강우는 식생 분포나 지형 형태 등의 환경 인자에 영향을 받는 것으로 알려져 있으며(Badas *et al.*, 2005; Onema and Taigbenu, 2009), 기존 TRMM 자료의 상세화 연구에서는 고해상도 부가 자료로 식생지수 자료나 고도 자료를 많이 이용해 왔다(Fang *et al.*, 2013; Shi *et al.*, 2015). 기존 연구를 참고하여 이 연구에서도 고해상도 부가 자료로 MODIS(Moderate Resolution Imaging Spectroradiometer) 정규식생지수(Normalized Difference Vegetation Index: NDVI) 자료와 SRTM(Shuttle Radar Topography Mission) 수치표고모형(Digital Elevation Model: DEM) 자료를 이용하였다(Figure 3). 식생지수로는 1km 공간 해상도의 MODIS 월 합성 NDVI 자료(MOD13A2)를 이용하였다. SRTM DEM의 원 공간 해상도는 90m인데, 정규

Figure 2. TRMM precipitation data used for spatial disaggregation: (a) April, 2013 and (b) July, 2013. ASOS rain gauge data are denoted as circles.

Figure 3. Auxiliary data used for spatial disaggregation of TRMM precipitation: (a) MODIS NDVI in April, 2013, (b) MODIS NDVI in July, 2013, and (c) SRTM DEM.

식생지수 자료의 공간 해상도를 고려하여 1km 공간 해상도로 변환하였다. 해양과 육상의 NDVI와 고도 특성이 다르기 때문에 상세화 방법론은 육상 지역에서만 적용하였다. 육상과 해양의 서로 다른 물리적 특성이 혼재된 연안 지역에서는 원 TRMM 자료의 오차나 불확실성이 크게 나타날 수 있다. 그래서 육상 지역에서의 상세화 결과를

생성할 때, Figure 2에 제시된 것처럼 해양 지역의 TRMM 자료의 이용은 최소화하였다.

상세화 결과의 정확도 검증에는 우리나라의 종관기상관측소(Automated Synoptic Observing System: ASOS)의 월 누적 강우량을 사용하였다. 79 개의 종관기상관측소에서 기상 관측이 중단된 관측소와 일부 섬 지역의 관측소를 제외하고, 총 69

개 관측소 자료를 정확도 검증에 이용하였다(Figure 2).

부가 자료의 공간 해상도를 고려하여 이 연구에서는 25km TRMM 자료를 1km 공간 해상도로 상세화를 수행하였는데, 기존 연구에서 많이 적용해온 RF를 이용한 경향 성분 추정 기반 상세화와 이 논문에서 제안한 RF와 잔차 크리깅을 결합한 방법론을 상호 비교하였다. ASOS 자료와 상세화 결과의 값 비교를 통해 오차 통계값으로 평균제곱근오차(Root Mean Square Error: RMSE)를 계산하여 검증에 사용하였다. 추가적으로 두 상세화 결과의 예측 성능을 상대적으로 비교하기 위해 RF를 이용한 경향 성분 추정 기반 상세화 방법론의 RMSE(RMSE₁)에 대한 제안 방법론의 RMSE(RMSE₂)의 상대적 향상지수(Relative Improvement Index: RI)를 식 (3)을 이용하여 계산하였다.

$$RI(\%) = \frac{(RMSE_1 - RMSE_2)}{RMSE_1} \times 100 \quad (3)$$

2) 상세화 결과

고해상도 NDVI와 DEM 자료를 이용한 저해상도 TRMM 자료의 경향 성분 추정에 RF를 적용하였다. 4월과 7월에 대하여 RF를 적용하고 모형 예측값과 TRMM 자료의 값 사이의 결정 계수를 계산한 결과, 4월과 7월에서 각각 83%와 88%의 높은 설명 능력을 나타내었다. 이 모형을 고해상도 NDVI와 DEM 자료에 적용하여 1km에서의 경향 성분을 추정하였다(Figure 4).

경향 성분 추정 후에는 25km 공간 해상도에서의 잔차 성분을 대상으로 베리오그램 디컨볼루션을 적용하여 점 기반 베리오그램 모델을 추정하였다. 그리고 추정된 모델과 영역-점 변환 크리깅을 이용하여 1km에서의 잔차 성분을 추정하였다(Figure 5). 월 누적 강우량이 상대적으로 적은 4월이 7월에 비해 잔차의 크기가 전반적으로 작으면서 변동성이 작게 나타났다. RF 모형의 설명 능력이 높게 나타났지만, 전반적으로 낮은 값을 과추정, 높은 값을 저추정하는 양상을 보였다. 특히 높은 값의 저추정 양상은 7월에 두드러지는 것으로

Figure 4. Trend components at 1km resolution estimated by RF: (a) April, 2013 and (b) July, 2013.

Figure 5. Residual components at 1km resolution estimated by area-to-point kriging: (a) April, 2013 and (b) July, 2013.

나타났으며, 이러한 양상은 Figure 5b의 북부 지역에 상대적으로 넓게 나타나고 있다. 따라서 이러한 양상은 잔차 성분 특성의 고려 여부에 따라 상세화 결과가 영향을 받을 수 있음을 지시한다.

Figure 4의 고해상도 경향 성분과 Figure 5의 영역-점 변환 크리깅으로 추정한 잔차 성분을 합하여 최종 상세화 결과를 얻었다(Figure 6). RF를 이용한 경향 성분 추정 기반 상세화 결과(Figure 4)와 RF와 잔차 크리깅을 결합하여 얻은 상세화 결과(Figure 6)는 모두 전반적으로는 TRMM 자료의 강우 패턴을 따르면서 부가 자료의 영향이 반영되어 있다. 그러나 이 두 방법론별 상세화 결과는 국소적으로 값의 차이를 보인다. 4월 상세화 결과에서 일부 지역을 살펴보면(Figure 7), RF만을 이용한 상세화 결과에서 상대적으로 값의 과추정 양상이 나타나지만, 잔차 크리깅을 결합한 상세화 결과에서는 이러한 과추정 양상이 상대적으로 완화되었다. 7월 일부 지역의 상세화 결과에서는 RF만을 이용한 결과에서 저추정 양상이 나타났지만, 제안 방법론을 이용한 상세화 결과에서는 이러한

양상이 완화된 것을 확인할 수 있다(Figure 7).

상세화 결과의 특성을 좀 더 살펴보기 위해, 1km 해상도에서의 상세화 결과를 TRMM 자료의 원래 해상도인 25km로 업스케일링하여 원 자료와 비교하였다. 상세화 결과는 부가 자료의 분포 때문에 육상 지역에서만 생성하여 해양과 일부를 포함하는 픽셀을 제외하고 육상 지역에서만 업스케일을 수행하고 TRMM 자료와 비교하였다. 업스케일링된 상세화 결과와 원 TRMM 자료와의 산점도를 그려보면(Figure 8), 상세화 결과에서 나타난 것과 같이 4월과 7월 모두 RF 기반 상세화 결과에는 낮은 값의 과추정과 높은 값의 저추정 양상이 두드러지게 나타나고 있다. 반면 제안 방법론의 상세화 결과에서는 이러한 양상이 완화되어 상대적으로 많이 향상된 상관계수를 보이고 있다. 그러나 제안 방법론의 상세화 결과에서도 여전히 저추정 및 과추정 양상이 나타난다. 영역-점 변환 크리깅은 상세화 결과를 원 자료의 해상도로 업스케일링했을 때 원 자료의 값을 그대로 재생산하는 coherency 특성을 보이는데(Kyriakidis, 2004),

Figure 6. Spatial disaggregation results: (a) April, 2013 and (b) July, 2013.

Figure 7. Comparison of disaggregation results.

제안 방법론에서는 잔차 성분의 상세화에 영역-점 변환 크리깅을 적용했기 때문에 Figure 5의 잔차 성분의 상세화 결과를 25km로 업스케일링하면 25km에서의 저해상도 잔차 성분을 그대로 재생산한다. 그러나 최종 상세화 결과는 RF 모형으로부

터 추정된 경향 성분이 잔차 성분에 함께 더해 얻어진다. 따라서 경향 성분의 저추정 및 과추정 양상의 영향으로 업스케일링 했을 때 원 TRMM 자료값을 완벽하게 재생산할 수는 없다. 상세화는 저해상도 자료를 참값으로 간주하여 모델링을 수

행하기 때문에 이러한 coherency 특성의 재현 여부는 상세화 방법론의 비교에 중요한 특성으로 간주할 수 있다. 따라서 제안 기법이 완벽하게 coherency 특성을 만족하지는 않지만, 상세화 과정에 잔차 성분을 고려하지 않거나 혹은 고려하더라도 일반적인 공간 내삽 기법을 적용하는 기존 연구들과 차별화된다고 할 수 있다.

ASOS 자료와 상세화 결과의 비교를 통한 오차 통계값은 Table 1에 제시하였다. 4월에 비해 7월의 RMSE가 상대적으로 매우 높게 나타나는데, 해당 시기의 강우 특성에 기인한 것으로 보인다. 7월

에는 장마, 집중 호우, 국지성 호우 등이 나타나는데, 3시간마다 관측되는 TRMM 강우 강도를 누적 강우량으로 변환하는 과정에서 특정 시기의 강수 패턴을 반영하지 못해서 오차가 커진 것으로 판단된다. 4월과 7월 상세화 결과에서 모두 제안 방법론의 예측 정확도가 높게 나타났으며, 4월과 7월에 각각 13.53%와 17.66%의 RMSE 향상을 보였다. 이러한 결과는 RF의 적용을 통해 높은 설명능력을 보이는 경향 성분의 추정이 가능하지만, 일부 저추정과 과추정 양상으로 인해 예측 정확도가 상대적으로 떨어진 것으로 판단된다. 반면 제

Figure 8. Scatter plots between TRMM precipitation values and averages of disaggregation results within each TRMM pixel.

Table 1. Evaluation summary statistics of disaggregation methods.

Disaggregation method	April		July	
	RMSE (mm)	RI (%)	RMSE (mm)	RI (%)
RF only	19.58	—	199.95	—
RF with area-to-point residual kriging	16.93	13.53	164.63	17.66

안 기법은 잔차 성분을 고려함으로써 RF로 추정된 경향 성분의 저추정과 과추정 양상을 일부 보정할 수 있어서 예측 정확도가 향상되었다. 특히 경향 성분의 저추정 영향이 두드러졌던 7월의 경우, 잔차 성분을 함께 고려함으로써 4월에 비해 더 높은 RMSE의 향상을 보였다. 따라서 경향 성분 추정에 개선된 모형을 적용하더라도 잔차 성분의 영향을 반드시 상세화 과정에 반영해야 함을 확인할 수 있었다.

4. 결론

이 연구에서는 저해상도 위성기반 강수 자료의 상세화를 목적으로 기계학습 모형과 잔차 크리깅을 결합한 상세화 방법론을 제안하였다. 기계학습 모형을 이용한 기존 상세화 연구의 대부분은 기계학습 모형을 고해상도 부가 자료에 적용한 경향 성분만을 고려하거나, 잔차 성분을 고려하더라도 저해상도 잔차 성분의 특성을 반영하지 못하는 공간 내삽 기법을 적용해왔다. 이 논문에서는 저해상도 강수 자료와 부가 자료와의 상호 연관성을 잘 반영하기 위해 고급 회귀 모형으로 기계학습 모형인 RF를 적용하여 고해상도 경향 성분을 추정하였다. 그리고 경향 성분으로 표현하지 못하는 잔차 성분의 상세화를 위해 저해상도 자료의 고해상도 변환에 적합한 영역-점 변환 크리깅을 적용하였고, 최종적으로 이 두 성분들을 합하여 상세화 결

과를 얻었다.

저해상도 TRMM 강수 자료를 대상으로 고해상도 NDVI와 DEM 자료를 이용한 상세화 사례 연구 결과, 제안 방법론이 기존 RF를 이용한 고해상도 경향 성분 추정 기반 상세화 방법론에 비해 향상된 예측 능력을 보였다. 또한 원 자료의 공간 해상도로 업스케일링 했을 때 제안 상세화 방법론의 원 자료 재현 능력이 상대적으로 높은 것을 확인하였다. 이러한 결과들은 고급 회귀 모형의 적용을 통해 고해상도 경향 성분을 추정하더라도 모형으로 설명할 수 없는 잔차 성분을 상세화 과정에 반드시 고려해야 함을 지시한다.

이 논문에서는 제안 기법을 저해상도 강수 자료의 상세화를 목적으로 적용하였지만, 고해상도 부가 자료의 이용이 가능하다면 범용적으로 상세화에 적용할 수 있다. 앞으로 강수 자료 이외의 다른 저해상도 기후 혹은 환경 인자들(지표면 온도, 토양 수분)에 대한 적용 연구를 통해 제안 방법론의 일반화 가능 여부를 확인할 예정이다. 또한 경향 성분 추정에 다른 기계학습 모형(support vector regression, artificial neural network)과 일반적으로 기계학습 모형에 비해 모형의 설명 능력이 떨어질 수 있는 다중 선형 회귀모형을 함께 비교하여 경향 성분의 추정 양상이나 설명 능력이 잔차 성분의 특성과 최종 상세화 결과에 미치는 영향을 정량적으로 검토할 예정이다.

사사

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행되었습니다(NRF-2015R1A1A1A05000966).

참고문헌

- Agam, N., Kustas, W. P., Anderson, M. C., Li, F., and Neale, C. M., 2007, A vegetation index based technique for spatial sharpening of thermal imagery, *Remote Sensing of Environment*, 107(4), 545-558.
- Atkinson, P. M., 2013, Downscaling in remote sensing, *International Journal of Applied Earth Observation and Geoinformation*, 22, 106-114.
- Atzberger, C., 2013, Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs, *Remote Sensing*, 5(2), 949-981.
- Badas, M. G., Deidda, R., and Piga, E., 2005, Orographic influences in rainfall downscaling, *Advances in Geosciences*, 2, 285-292.
- Baek, S.-G. and Jang, D.-H., 2011, Evaluation for applicability of cokriging for high resolution spatial mapping of temperature and rainfall, *Journal of Climate Research*, 6(3), 242-253 (in Korean).
- Bechtel, B., Zakšek, K., and Hoshyaripour, G., 2012, Downscaling land surface temperature in an urban area: a case study for Hamburg, Germany, *Remote Sensing*, 4(10), 3184-3200.
- Breiman, L., 2001, Random forests, *Machine Learning*, 45(1), 5-32.
- Brown, M., De Beurs, K., and Marshall, M., 2012, Global phenological response to climate change in crop areas using satellite remote sensing of vegetation, humidity and temperature over 26years, *Remote Sensing of Environment*, 126, 174-183.
- Chen, F., Liu, Y., Liu, Q., and Li, X., 2014, Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity, *International Journal of Remote Sensing*, 35(9), 3074-3093.
- Fang, J., Du, J., Xu, W., Shi, P., Li, M., and Ming, X., 2013, Spatial downscaling of TRMM precipitation data based on the orographical effect and meteorological conditions in a mountainous area, *Advances in Water Resources*, 61, 42-50.
- Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*, Oxford University Press.
- Goovaerts, P., 2008, Kriging and semivariogram deconvolution in the presence of irregular geographical units, *Mathematical Geosciences*, 40(1), 101-128.
- Guan, H., Wilson, J. L., and Xie, H., 2009, A cluster-optimizing regression-based approach for precipitation spatial downscaling in mountainous terrain, *Journal of Hydrology*, 375(3), 578-588.
- 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): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *Journal of Hydrometeorology*, 8(1), 38-55.
- Hutengs, C. and Vohland, M., 2016, Downscaling land surface temperatures at regional scales with random forest regression, *Remote Sensing of Environment*, 178, 127-141.
- Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., and Okamoto, K., 2000, Rain-profiling algorithm for the TRMM precipitation radar, *Journal of Applied Meteorology*, 39(12), 2038-2052.
- Jang, D.-H., Wi, N.-S., and Park, N.-W., 2015, High-resolution spatial mapping and evaluation of temperature and rainfall in South Korea using a simple kriging with local means, *Journal of Cli-*

- mate Research*, 10(2), 165-182 (in Korean).
- Jia, S., Zhu, W., Lu, A., and Yan, T., 2011, A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China, *Remote Sensing of Environment*, 115(2), 3069-3079.
- Kummerow, C. D., Barnes, W., Kozu, T., Shiue, J., and Simpson, J., 1998, The tropical rainfall measuring mission(TRMM) sensor package, *Journal of Atmospheric Ocean Technology*, 15(3), 809-817.
- Kyriakidis, P. C., 2004, A geostatistical framework for area to point spatial interpolation, *Geographical Analysis*, 36(3), 259-289.
- Latifovic, R. and Pouliot, D., 2007, Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record, *Remote Sensing of Environment*, 106(4), 492-507.
- Najafi, M. R., Moradkhani, H., and Wherry, S. A., 2011, Statistical downscaling of precipitation using machine learning with optimal predictor selection, *Journal of Hydrologic Engineering*, 16(8), 650-664.
- Onema, J. M. K. and Taigbenu, A., 2009, NDVI-rainfall relationship in the Semliki watershed of the equatorial Nile, *Physics and Chemistry of the Earth, Parts A/B/C*, 34(13), 711-721.
- Park, N.-W., 2013, Spatial downscaling of TRMM precipitation using geostatistics and fine scale environmental variables, *Advances in Meteorology*, 2013, 1-9.
- Park, N.-W. and Jang, D.-H., 2013, Geostatistical analysis of monthly temperature distributions using topographic data: a case study around Jirisan areas, *Journal of Climate Research*, 8(3), 217-235 (in Korean).
- Rahman, S., Bagtzoglou, A. C., Hossain, F., Tang, L., Yarbrough, L. D., and Easson, G., 2009, Investigating spatial downscaling of satellite rainfall data for streamflow simulation in a medium-sized basin, *Journal of Hydrometeorology*, 10(4), 1063-1079.
- Rodriguez-Galiano, V. F., Ghimire, B., Pardo-Igúzquiza, E., Chica-Olmo, M., and Congalton, R. G., 2012, Incorporating the downscaled Landsat TM thermal band in land-cover classification using random forest, *Photogrammetric Engineering and Remote Sensing*, 78(2), 129-137.
- Shi, Y., Song, L., Xia, Z., Lin, Y., Myneni, R. B., Choi, S., Wang, L., Ni, X., Lao, C., and Yang, F., 2015, Mapping annual precipitation across mainland China in the period 2001-2010 from TRMM3B43 product using spatial downscaling approach, *Remote Sensing*, 7(5), 5849-5878.
- Srivastava, P. K., Han, D., Ramirez, M. R., and Islam, T., 2013, Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application, *Water Resources Management*, 27(8), 3127-3144.
- Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., and Dickinson, R., 2013, The role of satellite remote sensing in climate change studies, *Nature climate change*, 3(10), 875-883.
- Yuan, H., Dai, Y., Xiao, Z., Ji, D., and Shangguan, W., 2011, Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling, *Remote Sensing of Environment*, 115(5), 1171-1187.

책임편집: 박지훈

영문교열: 조창현