빅데이터 분석 및 예측을 위한 멀티모델 태풍 시뮬레이션

Multi-model Typhoon Simulation for Big Data Analysis and Prediction

강 지 순, 육 진 희, 조 민 수
한국과학기술정보연구원 융합기술연구본부
재난대응HPC연구센터

Kang Ji-Sun, Yuk Jin-Hee, Joh Minsu
Disaster Management HPC Technology Research Center,
Korea Institute of Science and Technology Information

요약

한국과학기술정보연구원 융합기술연구본부 재난대응HPC연구센터에서는 초고성능컴퓨팅 기반의 풍수해 예측 및 피해 정보 생산 기술을 연구개발하여 재난·재해에 대한 국가현안 대응 의사결정지원 시스템을 구축 중에 있다. HPC 기반의 풍수해 예측 시스템과 빅데이터 분석 기반의 피해 예측 시스템에 대한 연구를 독자적으로 진행하는 가운데, 최근 여러 분야에 적용되고 있는 빅데이터 분석 기술을 HPC 기반의 풍수해 예측 시스템에 적목시켜 더 정확하고 신속한 풍수해 예측 정보 생산에 기여하고자 한다. 본 연구는 빅데이터 분석을 위한 학습 데이터 생산을 목적으로 HPC 기반 태풍 예측의 주요 기상 인자들을 조정하여 서로 다른 성능의 예측 모델을 구축하고, 각 모델 별 태풍 시뮬레이션의 성능을 진단하였다. 향후 빅데이터 분석을 통한 예측 성능의 검증을 위해 HPC 기반 풍수해 예측 및 검증 데이터를 최대한 생산하고자 한다.

I. 서론

재난대응HPC연구센터의 풍수해 대응 의사결정지원 시스템에서 주요한 예측 변수는 태풍의 진로와 강도이다. 이를 정확하게 예측하기 위해서 HPC기반 기상예측모델을 다양하게 개발 중에 있으며, 모델 별 특성에 따라 특정 모델 하나만을 예측 정보 생산에 쓰는 것이 아니라, 각각의 특성을 가진 모델의 장점을 극대화하고 단점을 최소화하는 방안인 멀티모델 앙상블 예측 시스템을 운영하여 최종 예측정보를 생산한다. 이는 수치예측모델을 자체 운영 중인 대부분의 현업 예보 기관에서도 도입 중이거나 도입 예정인 시스템으로서, 예보 불확실성을 줄이는 방안이며 앙상블 수만큼 계산 비용을 요구한다. 고성능컴퓨팅 기반 예측기술을 개발하는 본 기관에서 본 연구를 통하여 이와 같은 선도적 기술을 도입하고, 동시에 멀티모델 앙상블 예측의 결과를 빅데이터 분석에도 활용하고자 한다.

Ⅱ. 멀티모델 구성 및 태풍 사례 선정

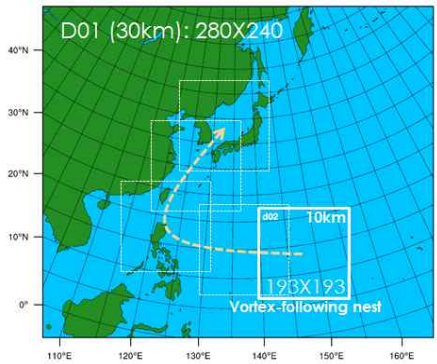
1. WRF 모델을 이용한 멀티모델

본 연구에서는 지역규모 날씨 예측에 가장 보편적으로 사용되는 수치모델인 WRF-ARW (이하 WRF)을 채택하였다[1]. 태풍의 진로와 강도 예측에 지대한 영향을 끼칠 것으로 판단되는 요소를 선별하여 5개의 서로 다른 WRF 모델을 설정하였다 (표 1). 태풍의 강한 대류현상을 모수화하는 cumulus 기법과 태풍의 진로와 강도에 영향을 줄 수 있는 해수면 온도 설정 (SST update)을 달리하고, 추가로 지역규모 모델의 초기 및 경계 조건으로 활용되는 전구모델을 미국 기상청 자료 (GFS FNL)과 유럽 중기예보청 (ECMWF) 자료 두 가지로 달리 설정해줌으로써 전구자료에 대한 변별력을 실험하였다. 두 자료 모두 해상도는 1° (중위도에서 약 110km)이다.

표 1. 멀티모델 앙상블 구성

Ensemble	Cumulus	SST update	Global data
M01	KF	X	GFS FNL
M02	KF	O	GFS FNL
M03	Tiedtke	X	GFS FNL
M04	Tiedtke	O	GFS FNL
M05	KF	X	ECMWF

여기서 WRF 모델은 그림 1과 같이 2개의 영역으로 바깥 영역은 30km 해상도로 고정된 영역을 시뮬레이션하나, 안쪽 영역은 태풍을 따라 이동하는 moving nest이며 10km 해상도로 시뮬레이션하였다. 연직층은 30개이다.



▶▶ 그림 1. WRF 모델 시뮬레이션 영역

2. 태풍 사례 선정

전통적 방식의 통계 분석과 마찬가지로 빅데이터 분석 역시 학습할 사례가 많으면 많을수록 좋으나, 현재까지 총 10개 사례에 대해 5개 앙상블 예측을 완료하였고 (표 2), 지속적으로 사례를 추가 중에 있다.

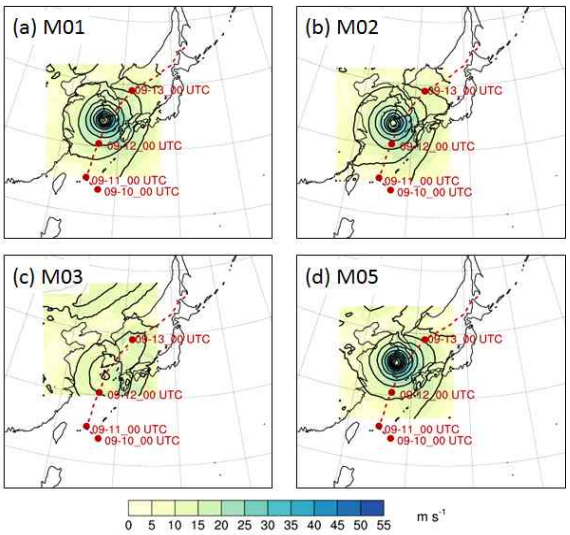
표 2. 선정된 10개 태풍 사례

발생연도	태풍 이름	발생연도	태풍 이름
2002	RUSA	2012	TEMBIN
2003	MAEMI	2012	BOLAVEN
2004	MEGI	2012	SANBA
2006	EWINIAR	2015	CHAN-HOM
2010	DIANMU	2017	CHABA

Ⅲ. 멀티모델 앙상블 시뮬레이션 분석

1. 적운 모수화 민감도

M01-M03 간의 비교와 M02-M04 간의 비교를 통하여 서로 다른 적운모수화 방안으로 인한 태풍 예측의 차이점을 살펴볼 수가 있는데, 10개 사례에 대해 평균적으로 KF 방안[2]이 Tiedtke방안[3]에 비해 태풍의 강도를 훨씬 강하게 모의하는 경향성을 보이며, 진로 예측 차이는 상대적으로 작다 (그림 2(a)&(c), 그림 3).



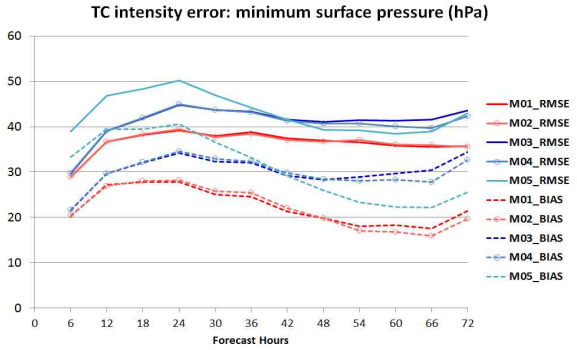
▶▶ 그림 2. WRF 이동영역에서의 M01, M02, M03, M05 모델 태풍 MAEMI 60시간 예측 결과. 컬러는 풍속(m/s)을, 곡선은 지표 기압(4hPa 간격)을, 붉은 점선은 태풍 베스트 트랙값을 나타냄.

2. 해수면 온도 설정 민감도

M01-M02 간의 비교와 M03-M04 간의 비교를 통하여 해수면 온도를 시뮬레이션 기간동안 업데이트를 하는 경우와 고정된 값을 쓰는 경우의 차이를 볼 때, 적운모수화나 전구자료 민감도에 비해 상대적으로 작은 민감도를 보였으나 향후 사례 추가를 통해 보다 면밀히 분석해볼 필요가 있다(그림 2(a)&(b), 그림 3).

3. 전구자료 민감도

M01-M05간의 비교를 통하여 살펴본 전구자료에 대한 민감도는 태풍의 진로와 강도 모두에 지대한 영향을 미치는 것으로 나타났다. 특히, GFS FNL 자료에 비해 ECMWF 자료가 단기 예측에 있어서는 열등하나 장기 예측으로 갈수록 성능이 다소 회복되는 패턴을 보였다(그림 3).



▶▶ 그림 3. 실험별 태풍강도(최저기압) 예측 오차 (RMSE는 실선, BIAS는 점선) 사례 평균 시계열

Ⅳ. 결론 및 향후 계획

태풍 예측 정확도 개선을 위해 고안된 멀티모델 앙상블 시스템을 활용하여 각 모델별 기상학적 이론을 바탕으로 한 설정 변화의 민감도를 살펴보았다. 태풍의 대류 현상을 모사하는 기법의 차이가 가장 큰 성능 차이를 야기한 반면, WRF 모델의 해수면온도 설정 차이에 의한 효과는 상대적으로 미미하게 나타났다. 물리적으로 해수면 온도 효과가 작다고 해석할 수 없고, WRF 모델의 해수면 온도 갱신 방안의 결과라고 보아야 타당할 것이며, 관련 연구가 더 필요하다. 전구 자료의 경우 ECMWF가 다고 열등한 성능을 보였으나, 진로에 대한 검증을 추가 진행해볼 계획이다. 앞으로 보다 신뢰도 높은 결과 도출을 위해 사례를 지속적으로 추가할 예정이며, HPC 기반의 분석 및 예측 결과를 빅데이터 분석이 얼마나 재현할 수 있는지, 혹은 능가할 수 있는지에 대한 연구가 뒤따를 예정이다.

사사

이 연구는 한국과학기술정보연구원 재난대응HPC연구센터가 수행하는 주요사업 '초고성능컴퓨팅기반 국가현안 대응체계 구축' 의 지원을 받아 수행되었습니다.

■ 참 고 문 헌 ■

[1] Skamarock, W. et al., "A Description of the Advanced Research WRF Version 3",NCAR Tech. Note, NCAR/TN-475+STR, 113pp.
[2] Kain, J.S., "The Kain-Fritsch convective parameterization: An update", J. Appl. Meteor., 43, 170-181, 2004.
[3] Tiedtke, M., "A comprehensive mass flux xcheme for cumulus parameterization in large-scale models", Mon. Wea. Rev., 117, 1779-1800, 1989.