

## 호우 영향예보를 위한 수문학적 정량강우예측(HQPF) 개선 연구

신윤후<sup>1\*</sup> · 김성민<sup>2</sup> · 지용근<sup>3</sup> · 이영미<sup>4</sup> · 김병식<sup>5</sup>

<sup>1</sup>주식회사 에코브레인 연구원, <sup>2</sup>주식회사 에코브레인 주임연구원, <sup>3</sup>주식회사 에코브레인 수석연구원/공학박사,  
<sup>4</sup>주식회사 에코브레인 대표이사/이학박사, <sup>5</sup>강원대학교 방재전문대학원 교수

## An Improvement Study on the Hydrological Quantitative Precipitation Forecast (HQPF) for Rainfall Impact Forecasting

Yoon Hu Shin<sup>1\*</sup>, Sung Min Kim<sup>2</sup>, Yong Keun Jee<sup>3</sup>, Young-Mi Lee<sup>4</sup>, and Byung-Sik Kim<sup>5</sup>

<sup>1</sup>Research Engineer, ECOBRAIN R&D Center

<sup>2</sup>Associate Research Manager, ECOBRAIN R&D Center

<sup>3</sup>Principal Research Engineer / Ph.D, ECOBRAIN R&D Center

<sup>4</sup>CEO / Ph.D, ECOBRAIN R&D Center

<sup>5</sup>Professor, Dept. of Urban Environment & Disaster Prevention School of Disaster Prevention, Kangwon National University

### 요약

최근 짧은 시간 동안 많은 강우가 내리는 국지성 집중호우가 빈번히 발생하고 이로 인한 침수피해가 증가하고 있다. 국지성 집중호우로 인한 피해를 예방하기 위하여 기상청이 제공하는 지역 앙상블 예측시스템(Local ENsemble prediction System, LENS)과 관측자료와 동네예보 자료를 활용한 기계학습과 확률 매칭(Probability Matching, PM) 기법을 이용하여 수문학적 정량강우예측정보(Hydrological Quantative Precipitation Forecast, HQPF)을 개발하였다. 국지성 집중호우로 인한 침수피해 대비를 위한 호우 영향정보로 HQPF를 생산하고 있지만, 낮은 강우강도에 대하여 과대예측하는 경향이 나타났다. 본 연구에서는 HQPF의 예측정확도 향상과 과대예측 성향을 개선하기 위하여 머신러닝 학습자료 기간확대, 앙상블 기법 분석 및 확률매칭(PM) 기법 프로세스 변경을 통하여 HQPF 개선하였다. 개선된 HQPF의 예측성능을 평가하기 위해 2021년 8월 27일 ~ 2021년 9월 3일 장마전선으로 인한 호우 사례를 대상으로 예측성능 검증을 수행하였다. 10 mm 이하의 강우에 대하여 예측정확도가 크게 향상되었고, 관측과 유사한 발생가능성 및 강우영역을 예측하는 등 과대예측 성향이 개선되었음을 확인하였다.

**핵심용어:** 집중호우, Local ENsemble prediction System (LENS), 앙상블 기법, 확률매칭, Hydrological Quantative Precipitation Forecast (HQPF)

### ABSTRACT

In recent years, frequent localized heavy rainfalls, which have a lot of rainfall in a short period of time, have been increasingly causing flooding damages. To prevent damage caused by localized heavy rainfalls, Hydrological Quantitative Precipitation Forecast (HQPF) was developed using the Local ENsemble prediction System (LENS) provided by the Korea Meteorological Administration (KMA) and Machine Learning and Probability Matching (PM) techniques using Digital forecast data. HQPF is produced as information on the impact of heavy rainfall to prepare for flooding damage caused by localized heavy rainfalls, but there is a tendency to overestimate the low rainfall intensity. In this study, we improved HQPF by expanding the period of machine learning data, analyzing ensemble techniques, and changing the process of Probability Matching (PM) techniques to improve predictive accuracy and over-predictive propensity of HQPF. In order to evaluate the predictive performance of the improved HQPF, we performed the predictive performance verification on

\*Corresponding author: Yoon Hu Shin, syh@ecobrain.net

Received: 8 December 2022, Revised: 27 December 2022, Accepted: 28 December 2022



heavy rainfall cases caused by the Changma front from August 27, 2021 to September 3, 2021. We found that the improved HQPF showed a significantly improved prediction accuracy for rainfall below 10 mm, as well as the over-prediction tendency, such as predicting the likelihood of occurrence and rainfall area similar to observation.

**Keywords:** Heavy rainfall, Local ENsemble prediction System (LENS), Ensemble techniques, Probability Matching, Hydrological Quantitative Precipitation Forecast (HQPF)

## 1. 서론

최근 짧은 시간 동안 많은 강우가 내리는 국지성 집중호우가 빈번히 발생하고 있으며, 이로 인한 침수피해 또한 증가하고 있다. 기상청(Korea Meteorological Administration, KMA)에서 제공하는 5 km 공간해상도의 1시간 간격 기상학적 정량적 강우 예보(Quantitative Precipitation Forecast, QPF)인 동네예보(Digital forecast)는 짧은 시간에 좁은 지역에 많은 강우가 발생하는 집중호우와 같은 시·공간변화가 큰 기상자료를 침수피해와 같은 수문학적 이벤트를 해석하는 데 한계가 있다.

정확한 기상예측정보 생산을 위해 방대한 기상 및 기후와 같은 빅데이터 기반의 다양한 방법론이 개발되어 예측정보 정확도를 향상을 연구가 진행되었고(Ghada et al., 2019; Valipour et al., 2019; Zamami et al., 2019), 수치예보자료의 정량적 강우량 추정 개선을 위해 인공지능망을 활용한 연구뿐만 아니라 집중호우와 같이 예측오차가 크고 단기간에 정확한 예측이 필요한 현상에 대해 기존의 예측한계를 보완하기 위한 방법으로 기계학습을 적용한 연구도 진행되었다(Hong, 2008; Sumi et al., 2012). 또한, 국지성 집중호우로 인한 피해 예방 즉, 수문학적 해석을 위한 예측정보 활용을 위하여 기상청이 제공하는 지역 앙상블 예측시스템(Local ENsemble prediction System, LENS)과 관측자료와 동네예보 자료를 활용한 기계학습과 확률매칭(Probability Matching, PM) 기법을 이용하여 수문학적 정량강우예측정보(Hydrological Quantative Precipitation Forecast, HQPF)를 개발하였다(Ko et al., 2020; Lee et al., 2021).

HQPF는 머신러닝의 XGBoost 기법을 통해 LENS 예보자료와 동네예보자료를 입력하여 3시간 간격의 앙상블 멤버별 보정강우를 생산한 후, 앙상블 멤버별 보정 강우를 각 멤버별로 PM 기법을 적용하여 3시간 누적 강우를 생산하는 예측강우 정보이다. 그러나 현재 HQPF는 낮은 강우강도의 강수에 대하여 과대예측하는 경향을 보이는 성능의 한계를 가지고 있다.

본 연구에서는 기존 연구(Lee et al., 2021)에서 개발한 HQPF의 예측정확도 향상과 과대예측 성향을 개선하기 위하여 머신러닝 학습자료의 기간확대, 앙상블 기법 분석 및 확률매칭(PM) 기법 프로세스 변경을 통하여 HQPF 개선하였다. 기존 연구(Lee et al., 2021)에서는 RMSE (Root Mean Square Error)와 공간장 제시로 HQPF의 정확도를 검증하였지만, 개선된 HQPF의 예측성능을 평가하기 위해 2021년 8월 27일~2021년 9월 3일 장마전선으로 인한 호우 사례를 대상으로 발생가능성 검증, 수치적 공간장 검증을 추가하여 예측성능 검증을 수행하였다.

## 2. 연구자료 및 방법

### 2.1 수문학적 정량강우예측정보(HQPF) 생산 프로세스

수문학적 정량강우예측정보(Hydrological Quantative Precipitation Forecast, 이하 HQPF)는 기상청에서 제공하는 지역 앙상블 예측시스템(Local ENsemble prediction System, LENS)과 동네예보를 사용한다. LENS 자료는 한반도 위험기상에 대한 확률예측을 위해 13개의 앙상블 섭동 멤버로 72시간까지 예측정보를 제공하는 UM (Unified Model) 기반 국지 앙상블 예측시스템으로 3 km의 공간 해상도와 12시간 주기의 시간 해상도로 제공된다. HQPF는 관측자료와 수치예보를 사용하여 전처리, 머신러닝, 확률매칭 기법, 후처리의 과정을 통해 생산된다. 첫 번째, 전처리 과정은 수치예보자료에서 머신러닝의 입

력자료로 사용되는 기상예측 인자들을 추출하기 위한 과정으로 기상관측자료와 기상예측자료에서 기상예측 인자들을 추출 후, 격자로 내삽되는 과정이다. 두 번째, 머신러닝 과정은 2016년부터 2018년까지 수집된 기상예측 인자를 활용하여 학습하였다. 기상예측 인자들은 머신러닝의 XGBoost 기법을 통해 LENS 예보자료와 동네예보자료를 입력하여 3시간 간격의 앙상블 멤버별 보정강우를 생산한다. 머신러닝 과정을 통해 생산된 앙상블 멤버별 보정 강우를 각 멤버별로 확률매칭(PM) 기법을 적용하여 생산하고 마지막 후처리 과정은 확률매칭(PM) 기법이 적용된 멤버별 예측강우를 평균하여 최종으로 3시간 누적 강우의 HQPF를 생산한다(Ko et al., 2020; Lee et al., 2021). 또한, 머신러닝 과정을 통해 생산된 13개 앙상블 멤버에 대하여 침수피해 대상체(보행, 교통, 시설)별 영향한계강우량(Lee et al., 2018) 초과 여부에 따라 확률값으로 발생가능성을 생산한다.

### 2.2 수문학적 정량강우예측정보(HQPF) 개선 프로세스

HQPF의 과대모의를 보완하기 위해 2가지의 프로세스를 개선하였다. 첫 번째는 정확도 향상을 위해 머신러닝의 학습자료를 기존 2016년부터 2018년에서 2019년과 2020년을 추가하였다. 두 번째는 과대모의의 직접적인 영향을 준 것으로 판단되는 확률매칭(PM) 기법의 프로세스를 변경하였다. 기존 HQPF의 확률매칭 프로세스의 경우 앙상블 각 멤버별로 확률매칭 기법이 적용되어 멤버들의 최대값의 평균으로 HQPF가 생산되었으며, 멤버들의 최대값이 적용되는 과정에서 낮은 강우에 대하여 과대예측하는 경향이 나타났다. 이에 본 연구에서는 각 멤버별로 적용되었던 확률매칭 기법을 앙상블 멤버 전체에 적용하여 개선하였다. 멤버 전체에 적용한 확률매칭 기법의 경우 강우 영역을 앙상블 평균 강우 영역으로 선정하였고, 강우량의 경우 앙상블 중간값을 적용하여 HQPF를 생산하였다. 기존 HQPF를 HQPF1이라 명명하였고, 학습기간 확대 및 확률매칭 기법 프로세스 개선을 통해 생산된 HQPF를 HQPF2로 명명하였다.

Fig. 1은 HQPF의 프로세스 개선 전과 후를 도식화한 것이다.

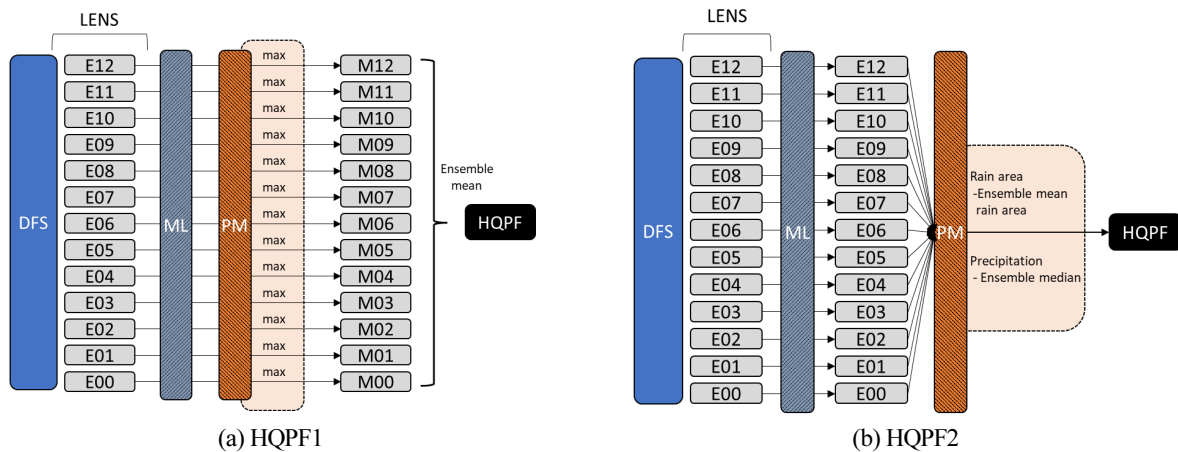


Fig. 1. Process of HQPF (a:HQPF1, b:HQPF2)

### 2.3 호우 사례 선정 및 검증방법

개선된 HQPF의 예측성능 검증을 위해 2021년 8월 27일 ~ 2021년 9월 3일 장마전선의 이동으로 인해 남부지방과 중부지방을 중심으로 집중호우가 발생한 호우 사례로 선정하였다.

선정된 호우 사례에 대하여 검증을 수행하기 위하여 기상청에서 제공하는 종관기상관측시스템(Automated Surface Observing System, ASOS)과 방재기상관측시스템(Automatic Weather Station, AWS)의 전국 1시간 간격의 강우 관측자료를 3시간 누적 관측자료로 생산하였고, 예측시간 1시간 간격으로 생산되는 5 km 격자의 동네예보 강우 자료를 3시간 누적 강우 자료로 생산하였다. 또한, 기존 연구에서 개발된 프로세스를 통해 생산된 HQPF와 앙상블 기법 변경과 확률매칭(PM)기

법 프로세스 변경을 통해 개선된 HQPF를 생산하였다. 한편, 관측자료는 지점 값으로 격자 단위로 생산되는 동네예보와 HQPF를 비교하기에는 한계가 있다. 따라서, 본 연구에서는 격자 자료인 동네예보와 HQPF의 예측강우량을 관측지점 주변 9개 격자의 평균값으로 계산하여 관측지점의 강우량에 대하여 동네예보와 HQPF1, HQPF2의 예측강우량 검증과 공간장수 치적 검증을 위하여 FSS (Fraction Skill Score)을 통해 검증을 실시하였다. 추가로 발생가능성을 검증하기 위해 신뢰도 다이어그램과 RPS (Ranked Probability Score)를 이용하여 HQPF2의 개선 효과를 확인하였다.

### 2.3.1 강우량 예측 검증(RMSE (Root Mean Square Error))

HQPF1이 낮은 강우강도에 대하여 과대예측하는 경향이 존재하였다. 이에 본 연구에서는 전체 강우강도, 10 mm 이하, 10~30 mm의 강우강도에 대하여 전국 ASOS 및 AWS의 관측강우량과 동네예보, HQPF1, HQPF2 예측강우량의 오차 정도 및 예측정확도를 파악할 수 있는 RMSE(Root Mean Square Error)를 활용하였다(식 (1)).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O_i - P_i)^2} \quad (1)$$

여기서, N은 자료 개수,  $P_i$ 는 예측강우량,  $O_i$ 는 관측강우량을 의미한다.

### 2.3.2 발생가능성 검증(신뢰도 다이어그램, RPS (Ranked Probability Score))

발생가능성의 검증을 위하여 확률예측 검증방법인 신뢰도 다이어그램과 RPS (Ranked Probability Score)를 활용하였다. 신뢰도 다이어그램과 RPS값 계산을 위하여 관측 강우의 영향한계강우량 초과 여부에 따라 0 또는 1로 변환된 값과 HQPF1 과 HQPF2의 침수피해 대상체별 발생가능성을 통해 계산하였다.

신뢰도 다이어그램은 예측 확률에 대한 관측된 빈도를 그래프를 통해 예측값의 과소, 과대예측에 대하여 표현하며, 예측 확률의 범위는 발생가능성에 따라 0.0~1.0 구간에서 0.1 단위로 표현된다. 발생가능성은 표현된 곡선과 대각선의 근접성을 이용하여 대각선으로부터 곡선이 대각선 아래에 있으면 과대예측을 나타내고 대각선 위의 점은 과소예측을 나타낸다.

RPS는 다중 범주의 확률예측에 대한 누적 확률값의 차이 제곱의 합으로, 예측 확률이 관측값과 차이가 클수록 큰 값을 나타내며, 0.25를 기준으로 예측성능을 판단한다. 범주가 2일 경우 BS (Bier Score)와 동일하다. 본 연구의 범주의 경우 침수피해 대상체별 영향한계강우량에 따라 관측유무(0 또는 1)이므로 BS (Bier Score)와 같이 계산된다(식 (2)).

$$RPS = \frac{1}{M-1} \sum_{i=1}^M [(\sum_{k=1}^i P_k) - (\sum_{k=1}^i O_k)]^2 = \frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2 \quad (2)$$

여기서, M은 확률예보의 범주,  $P_k$ 는 예보범주 k에서의 예측확률,  $O_k$ 는 범주 k에서 관측유무(0 또는 1),  $P_i$ 는 예보범주 i에서의 예측 확률,  $O_i$ 는 범주 i에서 관측유무(0 또는 1)을 의미한다.

### 2.3.3 공간장 검증(FSS (Fraction Skill Score))

FSS (Fraction Skill Score)는 한 지점을 중심으로 이웃한 지점에 대한 강수예보를 검증하는 방법이다. FSS 값은 관측값과 예측값이 각 격자점에서의 영향한계강우량(임계값)의 초과 여부에 따라 0 또는 1로 변환 후, 변환된 값으로 중심 격자에서 이웃한 지역에 대한 확률값을 계산한다. 관측값에 대한 확률값과 예측값에 대한 확률값의 비율로 FSS가 계산되며, FSS는 0에

서 1 사이의 값을 가지며 1에 가까울수록 예측모델의 강우 현상의 규모에 따른 예측성능이 높다(Roberts et al., 2008; Kim et al., 2015).

본 연구에서는 관측지점을 중심으로 5 km, 25 km, 51 km에 대하여 동네예보, HQPF1, HQPF2의 FSS값 비교를 통해 공간장 검증 실시하였다.

### 3. 연구결과

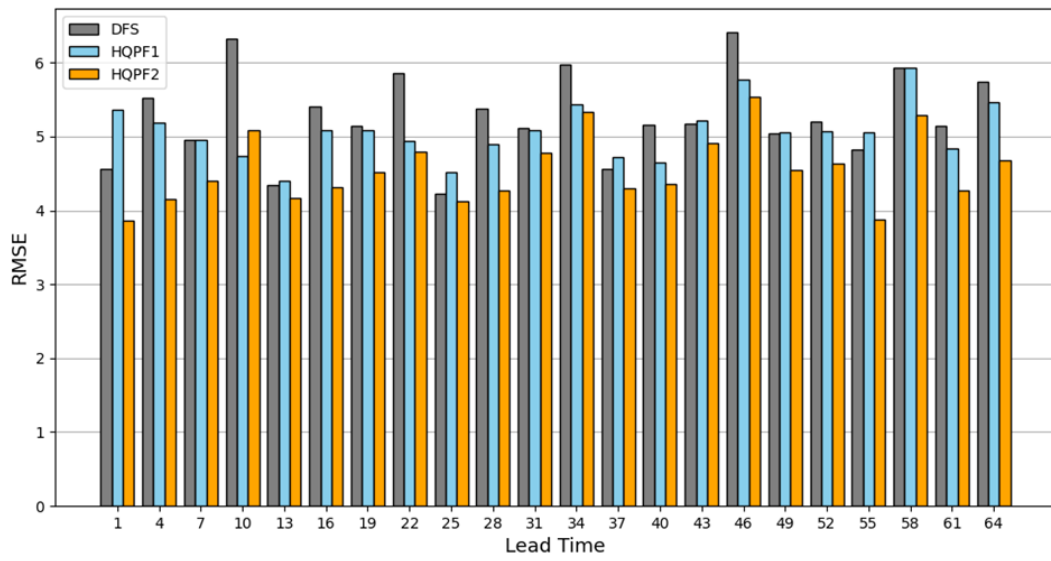
#### 3.1 강우량 검증결과

전국 관측지점을 중심으로 동네예보와 HQPF1, HQPF2의 선행시간별 강우량에 대한 예측성능을 평가하였다. Fig. 2는 호우 사례에 대한 선행시간별, 강우강도별 관측강우량과 예측강우량에 대한 3시간 누적강우량의 검증결과이다. Fig. 2(a) 전체 강우강도에 대하여 선행시간 +13 hr일 때를 제외하고 HQPF2의 RMSE가 동네예보와 HQPF1 보다 약 11.2% 예측정확도가 향상되었다. Fig. 2(b) 10 mm 이하의 강수에 대하여 HQPF2의 RMSE가 동네예보와 HQPF1 보다 모든 선행시간에서 가장 낮게 나타났으며, HQPF1 보다 약 34.2% 예측정확도가 향상되었음을 확인할 수 있었다. 또한, Fig. 2(c) 10 mm 초과 30 mm 이하의 강수에 대한 HQPF2의 RMSE가 일부 선행시간에서 HQPF1보다 높게 나타났지만, 약 7.8% 예측정확도가 향상되었다.

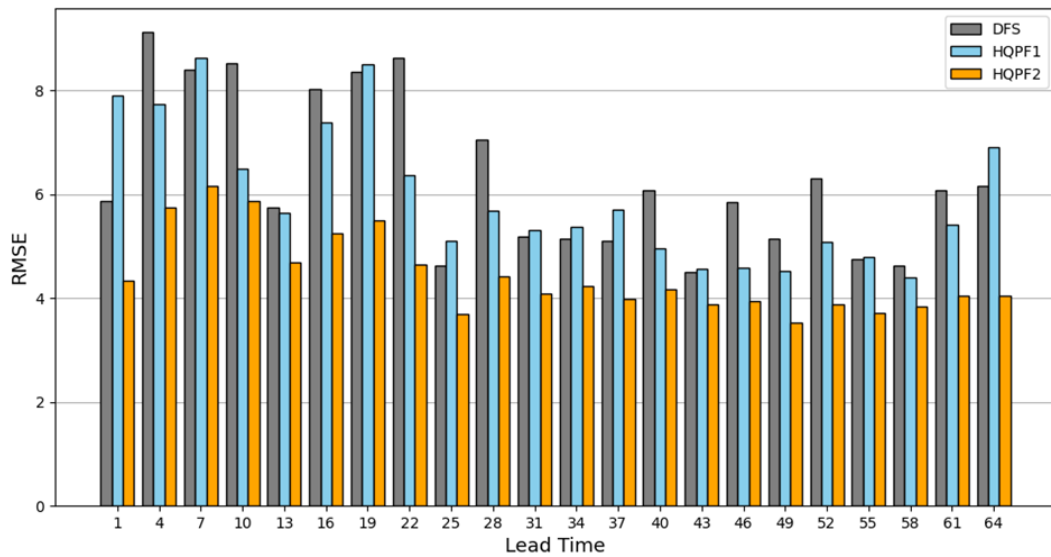
Table 1은 해당 호우 사례의 선행시간별 강우강도별 동네예보, HQPF1, HQPF2의 RMSE를 결과이다.

**Table 1.** RMSE for DFS rainfall and HQPF1,2 rainfall

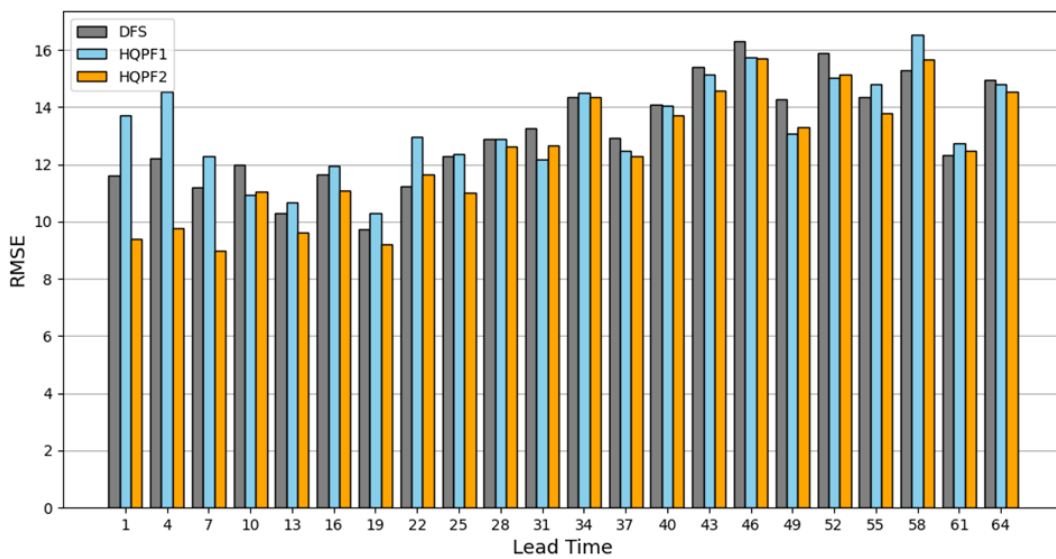
Lead Time (hr)	DFS RMSE			HQPF1 RMSE			HQPF2 RMSE		
	total	≤10 mm	10 < x ≤30	total	≤10 mm	10 < x ≤30	total	≤10 mm	10 < x ≤30
1	4.56	5.88	11.62	5.36	7.91	13.72	3.86	4.33	9.38
4	5.53	9.13	12.21	5.19	7.75	14.54	4.16	5.75	9.75
7	4.95	8.40	11.19	4.96	8.63	12.29	4.40	6.15	8.96
10	6.33	8.53	11.98	4.74	6.50	10.94	5.09	5.86	11.04
13	4.35	5.74	10.30	4.40	5.65	10.67	4.16	4.69	9.61
16	5.40	8.04	11.65	5.09	7.38	11.96	4.32	5.25	11.08
19	5.15	8.37	9.73	5.08	8.51	10.30	4.51	5.50	9.20
22	5.86	8.63	11.25	4.95	6.36	12.98	4.79	4.65	11.65
25	4.23	4.63	12.29	4.52	5.11	12.35	4.13	3.69	11.00
28	5.38	7.06	12.88	4.90	5.69	12.90	4.28	4.43	12.61
31	5.11	5.19	13.25	5.09	5.31	12.18	4.78	4.09	12.65
34	5.98	5.14	14.36	5.44	5.37	14.49	5.34	4.24	14.37
37	4.56	5.10	12.94	4.72	5.71	12.48	4.30	3.98	12.29
40	5.15	6.07	14.10	4.65	4.96	14.05	4.36	4.16	13.70
43	5.18	4.51	15.41	5.22	4.56	15.15	4.92	3.89	14.56
46	6.41	5.86	16.32	5.78	4.59	15.74	5.54	3.94	15.69
49	5.05	5.16	14.27	5.06	4.52	13.09	4.55	3.53	13.31
52	5.20	6.31	15.89	5.07	5.07	15.03	4.63	3.89	15.14
55	4.82	4.76	14.36	5.06	4.80	14.80	3.88	3.71	13.78
58	5.93	4.62	15.31	5.93	4.40	16.53	5.30	3.83	15.67
61	5.14	6.08	12.34	4.85	5.42	12.74	4.27	4.05	12.48
64	5.74	6.16	14.97	5.47	6.91	14.82	4.68	4.06	14.53



(a) Total Precipitation of RMSE



(b) Under 10 mm Precipitation of RMSE



(c) Over 10 mm to less 30 mm Precipitation of RMSE

Fig. 2. Precipitation of RMSE for leadtime

### 3.2 발생가능성 검증결과

침수피해 대상체인 보행(Man), 교통(Car), 시설(Build)을 중심으로 HQPF1, HQPF2의 발생가능성을 평가하였다. Fig. 3 침수피해 대상체별 HQPF1과 HQPF2의 신뢰도 다이어그램으로 모든 대상체에 대하여 HQPF1과 HQPF2의 발생가능성이 관측값 대비 과대예측을 하고 있지만, HQPF2의 발생가능성이 HPQF1보다 대각선과 더욱 가까워 관측을 잘 예측하는 것으로 나타났다. Fig. 3(a)와 Fig. 3(b) 보행과 교통에서 전 구간에서의 발생가능성의 예측이 HQPF1보다 향상되었으며, Fig. 3(c) 시설에서 3.0 이하의 구간에서 HQPF2 발생가능성의 예측이 향상되었다. 시설의 경우 4.0 이상 구간에서는 시설에 대한 영향한계강우량을 넘는 HQPF2의 발생가능성이 존재하지 않았다.

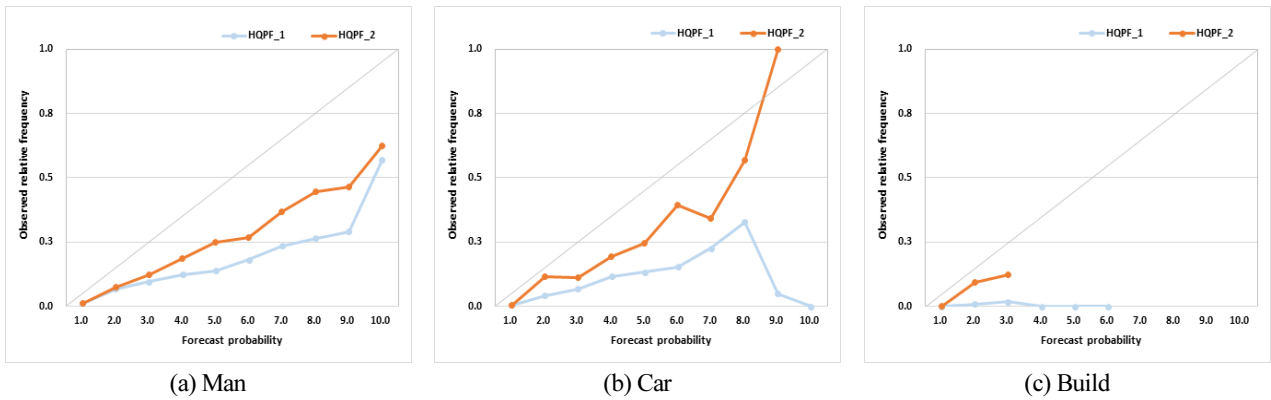


Fig. 3. Reliability diagram by Receptor

Fig. 4은 침수피해 대상체별 HQPF1과 HQPF2의 RPS 비교결과이다. 예측성능 판단기준인 0.25를 점선으로 나타내었으며, HQPF1과 HQPF2의 예측성능이 우수한 것으로 나타났으며, 모든 침수피해 대상체에 대하여 HQPF2의 예측성능이 향상되었다.

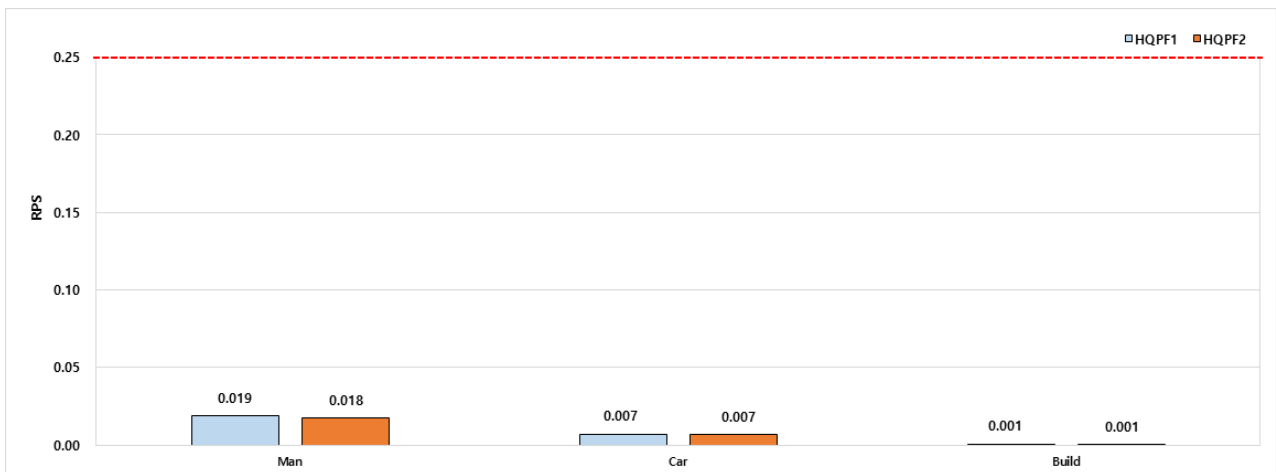


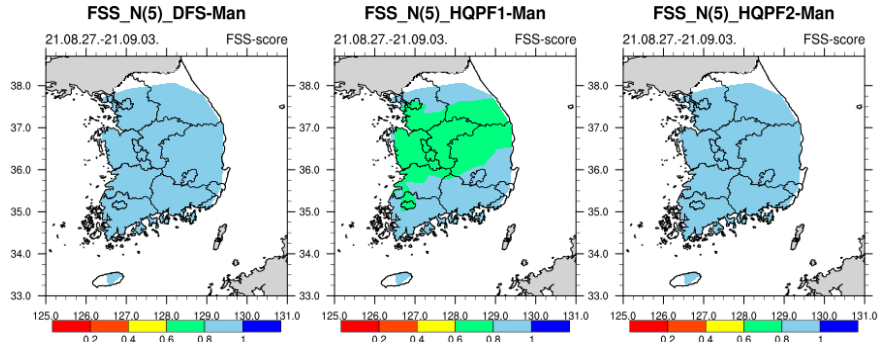
Fig. 4. Ranked Probability Score (RPS) by Receptor

### 3.3 공간장 검증결과

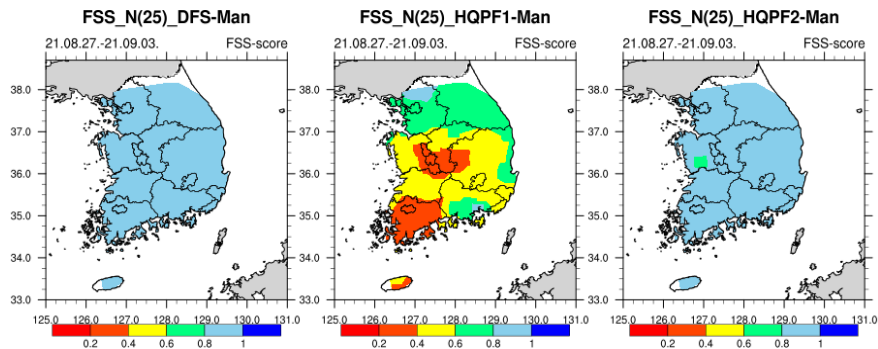
관측지점을 중심으로 5 km, 25 km, 51 km 영역에 따른 보행, 교통, 시설 대상체별 FSS를 통해 비교하였다.

Fig. 5는 보행(Man)의 FSS값에 대한 공간장과 영역별 FSS값을 비교한 결과이며, 관측지점에서 영역이 멀어질수록 예측

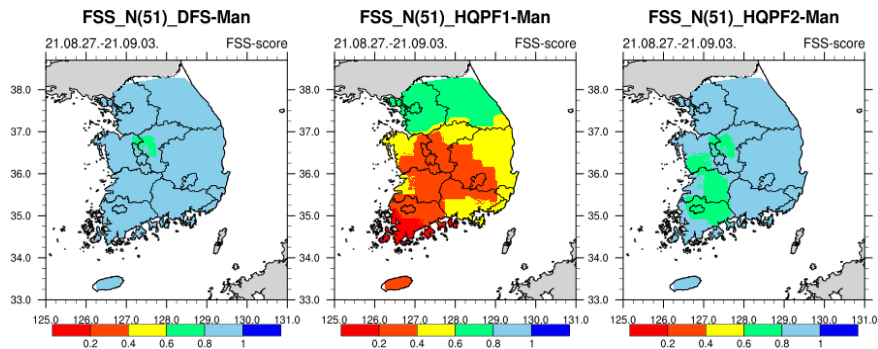
정확도가 낮아지고 있다. 영역별 비교결과, 5 km 영역에서 HQPF2가 관측과 가장 유사한 강우 영역이 나타났고, 25 km와 51 km에서는 동네예보가 관측과 가장 유사한 강우 영역 나타났다. 보행의 모든 영역에 대하여 HQPF2가 HQPF1보다 관측과 유사한 강우 영역을 보여, HQPF2의 관측 강우 영역에 대한 예측성능이 향상되었음을 확인할 수 있다.



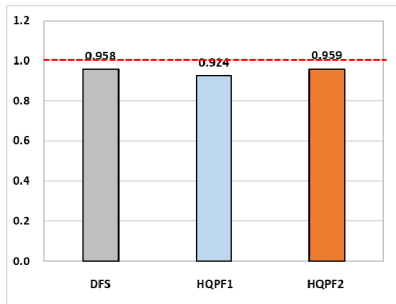
(a) Spatial of Man FSS (N=5)



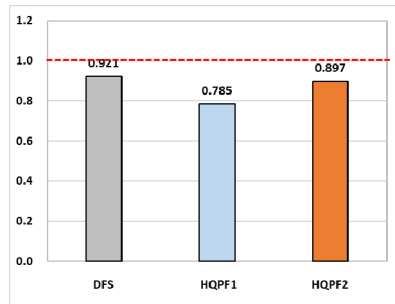
(b) Spatial of Man FSS (N=25)



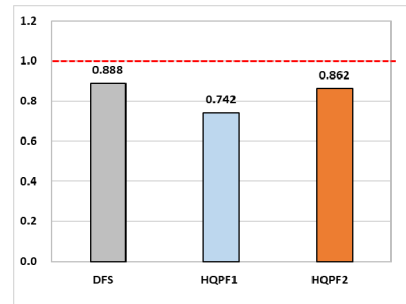
(c) Spatial of Man FSS (N=51)



(d) FSS of Man (N=5)



(e) FSS of Man (N=25)

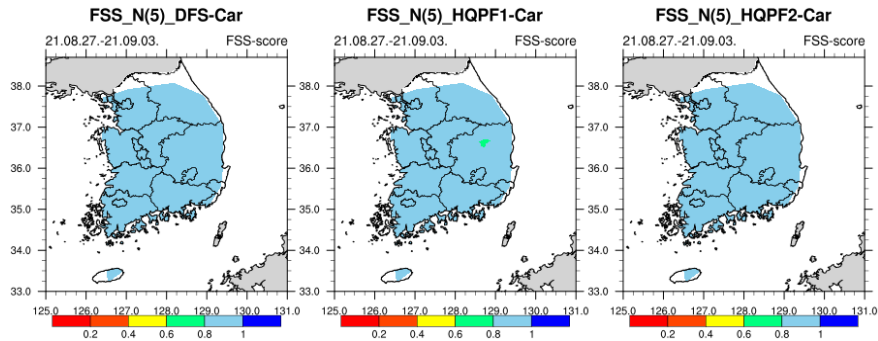


(f) FSS of Man (N=51)

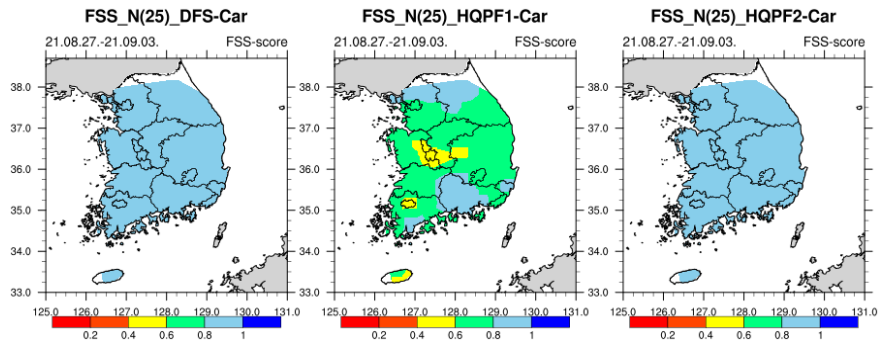
Fig. 5. Spatial of Man Fraction Skill Score (FSS) and FSS (N = Grid distance)



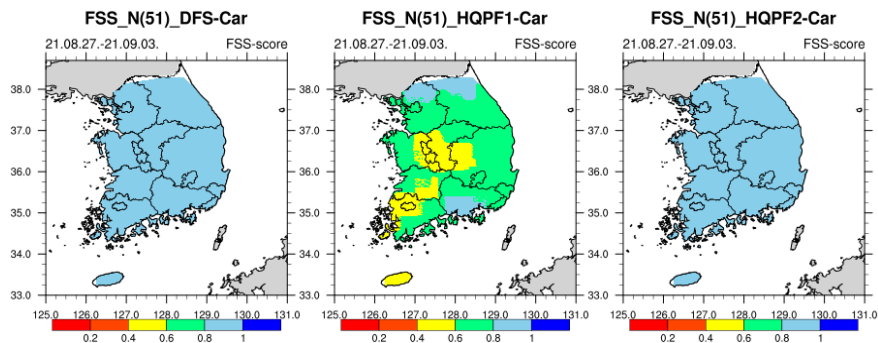
교통(Car)의 FSS값에 대한 공간장과 영역별 FSS값을 비교한 결과(Fig. 6), 보행과 동일하게 관측지점에서 영역이 멀어질수록 예측정확도가 낮아지고 있으나, 5 km 영역에서 HQPF2가 관측과 가장 유사한 강우 영역이 나타났고, 모든 영역에 대하여 HQPF2가 HQPF1보다 관측과 유사한 강우 영역을 보여, HQPF2의 관측 강우 영역에 대한 예측성능이 향상되었음을 확인할 수 있다.



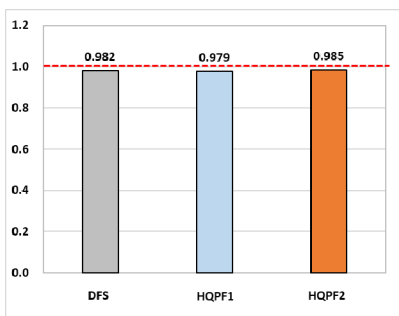
(a) Spatial of Car FSS (N=5)



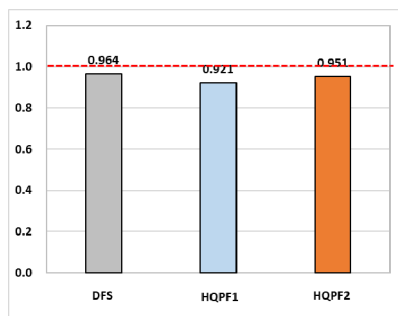
(b) Spatial of Car FSS (N=25)



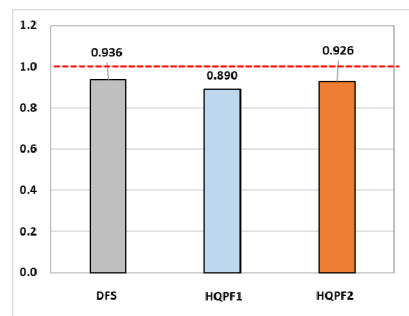
(c) Spatial of Car FSS (N=51)



(d) FSS of Car (N=5)



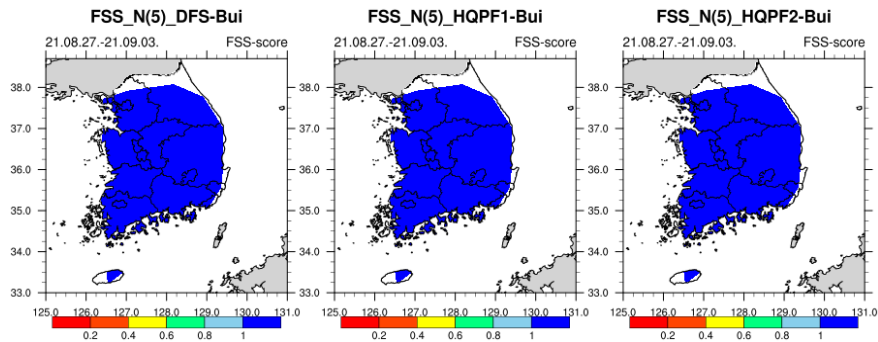
(e) FSS of Car (N=25)



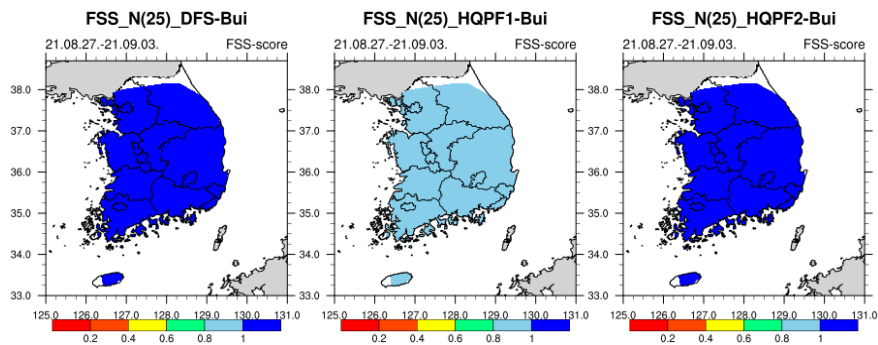
(f) FSS of Car (N=51)

Fig. 6. Spatial of Man Fraction Skill Score (FSS) and FSS (N = Grid distance)

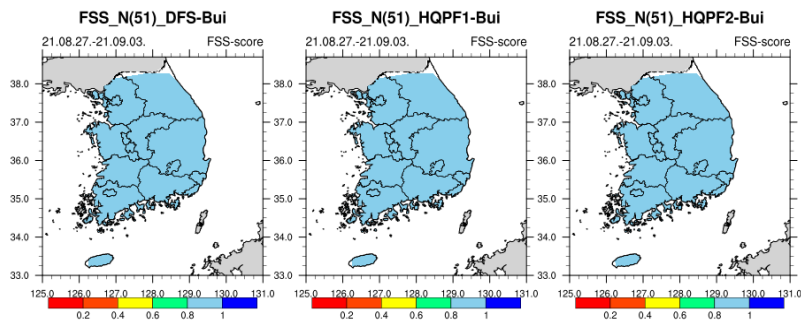
시설(Build)의 FSS값에 대한 공간장과 영역별 FSS값을 비교한 결과(Fig. 7), HQPF2의 강우영역 예측성능이 보행(Fig. 5)과 교통(Fig. 6)보다 관측과 가장 유사한 강우 영역이 나타났으며, 5 km, 25 km, 51 km 모든 영역에서 HQPF2의 강우 영역 예측성능이 가장 높게 나타났다.



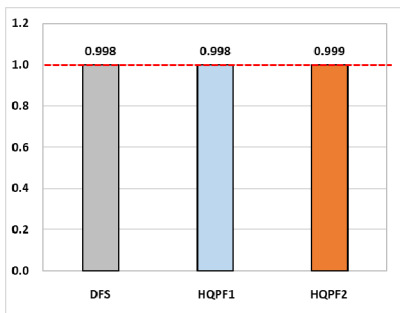
(a) Spatial of Build FSS (N=5)



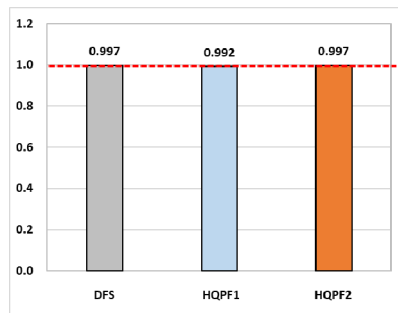
(b) Spatial of Build FSS (N=25)



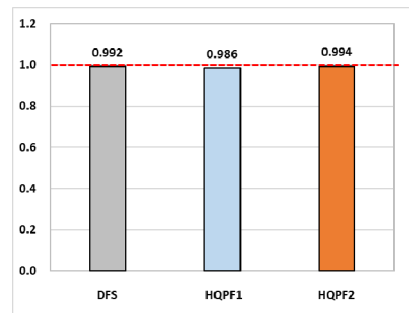
(c) Spatial of Build FSS (N=51)



(d) FSS of Build (N=5)



(e) FSS of Build (N=25)



(f) FSS of Build (N=51)

Fig. 7. Spatial of Build Fraction Skill Score (FSS) and FSS (N = Grid distance)

## 4. 결론

본 연구는 국지성 집중호우에 의한 침수피해 대비를 위해 호우 영향정보를 생산하는 HQPF1의 예측정확도 향상을 위해 머신러닝의 학습기간을 확대하였고, 멤버별 확률매칭(PM) 기법 적용 과정에서 발생하는 낮은 강우강도에 대한 과대예측 성능을 개선하고자 앙상블 기법 분석 및 확률매칭 기법 프로세스 변경을 통해 HQPF2를 생산하였다. 생산된 HQPF2의 예측 성능 검증을 보다 다양한 측면에서 분석하기 위해 기존 연구에서 활용한 RMSE를 통한 정량적 검증에 수치적 공간장 검증과 발생가능성 검증을 추가로 진행하였다. 발생가능성 검증은 신뢰도 다이어그램과 RPS를 이용하였으며, 수치적 공간장 검증은 FSS를 이용하였고 2021년 8월 27일 ~ 2021년 9월 3일 호우 사례를 바탕으로 생산된 동네예보, HQPF1, HQPF2의 예측 성능을 비교·검증하였다.

본 연구의 결론은 다음과 같다. 첫째, 강우강도별, 선행시간별 예측강우량과 관측강우량 비교결과, 전체 강우강도와 10 mm 초과 30 mm 이하의 강우강도의 일부 선행시간에서 HQPF2의 RMSE가 높게 나타났지만, 낮은 강우강도인 10 mm 이하의 강우에서 HQPF2의 RMSE가 모든 선행시간에 대하여 가장 낮게 나타나 머신러닝 과정의 학습자료 기간확대로 인한 예측정확도가 약 34.2% 향상됨을 확인하였고, HQPF1의 과대모의 경향을 개선하는 효과 또한 확인하였다.

둘째, HQPF1과 HQPF2의 발생가능성 검증결과 신뢰도 다이어그램을 통하여 개선 전과 개선 후 모두 관측값 대비 과대예측하는 경향이 보였으나, HQPF2가 모든 침수피해 대상체, 모든 예측 구간에서 과대예측 정도가 줄어들어 관측값과 유사해졌으며, RPS값을 통해 모든 침수피해 대상체에 대하여 예측성능이 향상됨을 확인하였다.

셋째, 관측지점 대상 인근 5 km, 25 km, 51 km 격자에 대하여 동네예보와 HQPF1, HQPF2의 강수확률 값을 FSS를 통해 비교하였다. 강우 영역이 관측지점과 멀어질수록 모든 침수피해 대상체에 대한 동네예보, HQPF1, HQPF2의 강수 영역의 예측정확도가 낮아졌으나, HQPF2가 HQPF1보다 관측과 유사한 강우 영역을 관측 강우 영역에 대한 예측성능이 향상됨을 확인하였다.

본 연구를 통해 생산된 HQPF가 일부 예측 선행시간에 대한 예측정확도 향상 효과가 상대적으로 낮게 나타났지만, 대부분 예측 선행시간에서 낮은 강우강도에 대한 예측정확도 향상 및 과대예측 성향이 개선되었음을 확인하였다. 향후, 꾸준한 머신러닝 튜닝, 학습기간 확대 및 앙상블 멤버 특성 반영을 통해 집중호우로 인한 침수피해 예방에 활용이 가능할 것으로 기대된다.

## Acknowledgment

This work was funded by the Korea Meteorological Institute (KMI2021-00311) Research project for the Advancement of Predictive Rainfall Production Technology for Disaster Impact Model in Heavy Rainfall.

## References

- Ghada, W., Eastrella, N., and Meanzel, A. (2019). Machine Learning Approach to Classify Rain Type based on This Disdrometers and Cloud Observations. *Atmosphere*. 10(5): 251-268.
- Hong, W. C. (2008). Rainfall Forecasting by Technological Machine Learning Models. *AMC*, 200(1): 41-57.
- Kim, S. H., Kim, H. M., Kay, J. K., and Lee, S. W. (2015). Development and Evaluation of the High Resolution Limited Area Ensemble Prediction System in the Korea Meteorological Administration. *Korean Meteorological Society*. 25(1): 67-83.
- Ko, C. M., Jeong, Y. Y., Ji, Y. K., Lee, Y. M., and Kim, B. S. (2020). A Study on Hydrological Rainfall Adjustment using Machine Learning and Probability Matching Method during Heavy Rainfall Season. *Journal of Climate Research*.

15(4): 257-267.

- Lee, H. S., Jee, Y. K., Lee, Y. M., and Kim, B. S. (2021). Optimizing Hydrological Quantitative Precipitation Forecast (HQPF) based on Machine Learning for Rainfall Impact Forecasting. *Journal of Environmental Science International*. 30(12): 891-905.
- Lee, S. H., Kang, D. H., and Kim, B. S. (2018). A Study on the Method of Calculating the Threshold Rainfall for Rainfall Impact Forecasting. *Journal of the Korean Society of Hazard Mitigation*. 18(7): 93-102.
- Roberts, N. M. (2008). Assessing the Spatial and Temporal Variation in Skill of Precipitation Forecasts from an NWP Model. *Meteorological Applications*. 15(1): 163-169.
- Sumi, S. M., Zaman, M. F., and Hirose, H. (2012). A Rainfall Forecasting Method using Machine Learning Models and its Application to the Fukuoka City Case. *Int. J. Appl. Math. Comput. Sci.* 22(4): 841-854.
- Valipour, M., Sefidkouhi, G., Ali, M., Raeini-Sarjaz, M., and Guzman, S. M. (2019). A Hybrid Data-driven Machine Learning Technique for Evapotranspiration Modeling Various Climates. *Atmosphere*. 10(6): 311-325.
- Zamami, J. M., Cao, C., Ni, X., Bashir, B., and Talebiesfandarani, S. (2019). PM2.5 Prediction based on Random Forest, XGBoost, and Deep Learning using Multisource Remote Sensing Data. *Atmosphere*. 10(7): 373-391.

#### **Korean References Translated from the English**

- 고철민, 정영윤, 지용근, 이영미, 김병식 (2020). 집중호우 시기 기계학습 및 PM기법을 이용한 수문학적 강우보정에 관한 연구. *건국대학교 기후연구소*. 15(4): 257-267.
- 김세현, 김현미, 계준경, 이승우 (2015). 기상청 고해상도 국지 양상불 예측 시스템 구축 및 성능 검증. *한국기상학회*. 25(1): 67-83.
- 이석호, 강동호, 김병식 (2018). 호우영향예보를 위한 한계강우량 산정 방법 연구. *한국방재학회논문집*. 18(7): 93-102.
- 이한수, 지용근, 이영미, 김병식 (2021). 호우 영향예보를 위한 머신러닝 기반의 수문학적 정량강우예측(HQPF) 최적화 방안. *한국환경과학회지*. 30(12): 891-905.