공학석사 학위논문

Introduction of Probabilistic Drought Prediction to Korea

가뭄 확률 전망의 국내 도입을 위한 연구

2020년 8월

서울대학교 대학원

건설환경공학부

김 대 호

Introduction of Probabilistic Drought Prediction to Korea

가뭄 확률 전망의 국내 도입을 위한 연구

지도 교수 김 영 오

이 논문을 공학석사 학위논문으로 제출함
2020년 7월

서울대학교 대학원
건설환경공학부
김 대 호

김대호의 공학석사 학위논문을 인준함
2020년 7월

위 원 장 ______________________ (인)

부위원장 ______________________ (인)

위 원 ______________________ (인)

Abstract

The advantage of probabilistic prediction has been verified and acknowledged for several decades so people are making use of the probabilistic prediction in lots of fields, including hydrometeorology. One of the biggest advantages is that it can take into account various events through uncertainty in the predicted value, especially for long-term predictions which have large uncertainties. In Korea, however, the drought prediction is still performed in a deterministic approach. Therefore, the purpose of this study is to apply the probabilistic drought prediction to Korea and then further propose a method to improve the prediction technique.

Accordingly, this study developed an ensemble drought prediction (EDP) system focusing on the hydrological drought measured by natural streamflow in eight basins in Korea. Because of the natural characteristic of drought, it only can be measured indirectly through the hydroclimatic variables. In order to measure the hydrological drought, the streamflow was converted to standardized runoff index (SRI) which is a kind of drought index considering regional characteristics and various time scales for the hydrological drought. Then to generate EDP distribution for 1-month ahead monthly drought prediction, the streamflow simulations of an ESP (Ensemble Streamflow Prediction) were converted to SRI. The deterministic prediction was done by the expected value of EDP distribution, and the probabilistic one was derived by the probability driven from the distribution. Moreover, to improve EDP, soil moisture index (SMI) satellite data provided by APEC climate center (APCC) were used to update EDP via the Bayes' theorem. The regression between SRI and SMI was used as a likelihood function that updates the EDP distribution. Additionally, the APCC precipitation probability forecast was used to update EDP using the PDF ratio method. As a result, three main conclusions were drawn as follows.

(1) The probabilistic drought prediction was 52% better than the deterministic on average in terms of prediction skills. When predicting the short-term drought, the probabilistic approach outperformed even more.

(2) Updating EDP using soil moisture information the via Bayes' theorem makes skill to be improved by 20% on average. It can be said that the soil moisture information corrects EDP if the likelihood function is valid and accurate.

(3) Reflecting the precipitation forecast to EDP via the PDF ratio yielded 6% better performance only for the non-irrigation period. From this, it was found again that reflecting informative data can make better the drought prediction.

Keywords: Drought prediction, Probabilistic approach, Ensemble prediction, Bayes' theorem, PDF ratio

Student number: 2018-29571

TABLE OF CONTENTS

Abstract i

Chapter 1. Introduction 1
1.1 Problem Statement 1
1.2 Research Objectives 2
1.3 Thesis Organization 2

Chapter 2. Literature Review 3
2.1 Drought Measures 3
2.2 Drought Prediction Methods 4
2.2.1 Deterministic Approach 4
2.2.2 Probabilistic Approach 5
2.3 Practical Use of Probabilistic Prediction 8
2.4 Drought Prediction in Korea 13

Chapter 3. Methodology 16
3.1 Ensemble Prediction 16
3.1.1 Concept of Ensemble 16
3.1.2 Ensemble Streamflow Prediction (ESP) 16
3.1.3 Ensemble Drought Prediction (EDP) 19
3.2 Bayes' Theorem 21
3.2.1 Prior Distribution 21
3.2.2 Likelihood function 21
3.2.3 Posterior Distribution 22
3.3 Performance Measures 23
3.3.1 Deterministic Approach 23
3.3.2 Probabilistic Approach 23

Chapter 4. Application 26
4.1 Study Area 26
4.2 Data Sets 28
4.2.1 Observed Data 28
4.2.2 ESP Dataset 28
4.2.3 Soil Moisture Index (SMI) 30
4.3 EDP with SMI 31
4.3.1 Modeling Framework 31
4.3.2 Results and Discussion 36
4.4 EDP with Probabilistic Precipitation Forecast 42
4.4.1 Probabilistic Precipitation Forecast by APCC 42
4.4.2 Modeling Framework 42
4.4.3 Results and Discussion 45

Chapter 5. Conclusion **47**
5.1 Summary and Conclusions 47
5.2 Future Study 49

References **50**

Appendix **58**
A-1 Ensemble Streamflow Prediction Results 58
A-2 Ensemble Drought Prediction Results 61

List of Tables

Table 2.1 Institutions performing probabilistic prediction 9
Table 3.1 Drought Classification criteria 20
Table 4.1 Information about basins 27
Table 4.2 Lag-1 correlation coefficients between SRI and SMI 30
Table 4.3 4-fold cross validation data cases 32
Table 4.4 RMSE of monthly time series regression of each fold 32
Table 4.5 Parameters of likelihood function 33
Table 4.6 N-RMSE of likelihood function SMI~SRI3 40
Table 4.7 N-RMSE of likelihood function SMI~SRI12 40
Table A1.1 Accuracy of ensemble streamflow prediction 58
Table A2.1 RMSE of two EDPs for SRI3 61
Table A2.2 RPSS of two EDPs for SRI3 62
Table A2.3 Averaged BS of two EDPs across eight basins for SRI3 63
Table A2.4 RMSE of two EDPs for SRI12 68
Table A2.5 RPSS of two EDPs for SRI12 69
Table A2.6 Averaged BS of two EDPs across eight basins for SRI12 70
Table A2.7 RMSE of four EDPs for SRI3 75
Table A2.8 RPSS of four EDPs for SRI3 76
Table A2.9 Averaged BS of four EDPs across eight basins for SRI3 77
Table A2.10 RMSE of four EDPs for SRI12 79
Table A2.11 RPSS of four EDPs for SRI12 80
Table A2.12 Averaged BS of four EDPs across eight basins for SRI12 81

List of Figures

Figure 2.1 Frameworks of prediction methods and their interactions 7
Figure 2.2 Flow diagram of combining prediction methods 7
Figure 2.3 Probabilistic SPI prediction by the World Bank 10
Figure 2.4 Example of monthly drought outlook in the U.S. 10
Figure 2.5 Probabilistic drought forecast of each drought phase 11
Figure 2.6 Example of probabilistic streamflow forecasts in the U.S. 12
Figure 2.7 Drought prediction procedure in Korea 14
Figure 2.8 Meteorological drought prediction by KMA 14
Figure 2.9 Hydrological drought prediction by ME 15
Figure 2.10 Hydrometeorological prediction by APCC 15
Figure 3.1 Flow chart of probabilistic drought prediction 17
Figure 3.2 Schematic diagram of ESP procedure 18
Figure 3.3 Schematic diagram of modified TANK model 18
Figure 4.1 Watersheds in Korea 27
Figure 4.2 SRIs on Soyang 29
Figure 4.3 SMI data provided by APCC 30
Figure 4.4 Example of Bayesian update with SMI at Soyang 35
Figure 4.5 Example of EDPs at Soyang 35
Figure 4.28 Differences of BS components between EDP and EDP+S 41
Figure 4.29 Probabilistic precipitation forecast by APCC 43
Figure 4.30 Example of EDP distributions at Soyang 43
Figure 4.31 Example of four EDPs at Soyang 44
Figure A1.1 Ensemble streamflow prediction result at Soyang 58
Figure A1.2 Ensemble streamflow prediction result at Daecheong 58
Figure A1.3 Ensemble streamflow prediction result at Andong 59
Figure A1.4 Ensemble streamflow prediction result at Seomjin 59
Figure A1.5 Ensemble streamflow prediction result at Chungju 59
Figure A1.6 Ensemble streamflow prediction result at Hapcheon 60
Figure A1.7 Ensemble streamflow prediction result at Namgang 60
Figure A1.8 Ensemble streamflow prediction result at Imha 60
Figure A2.1 Averaged RMSE of two EDPs across eight basins for SRI3 ..61
Figure A2.2 RPSS of two EDPs for SRI3 62
Figure A2.3 Averaged BSS of two EDPs across eight basins for SRI3 63
Figure A2.4 BSS at Soyang for SRI3 64
Figure A2.5 BSS at Daecheong for SRI3 64
Figure A2.6 BSS at Andong for SRI3 65
Figure A2.7 BSS at Seomjin for SRI3 65

Figure A2.8 BSS at Chungju for SRI3 66
Figure A2.9 BSS at Hapcheon for SRI3 66
Figure A2.10 BSS at Namgang for SRI3 67
Figure A2.11 BSS at Imha for SRI3 67
Figure A2.12 Averaged RMSE of two EDPs across eight basins for SRI12 68
Figure A2.13 RPSS of two EDPs for SRI3 69
Figure A2.14 Averaged BSS of two EDPs across eight basins for SRI12 ..70
Figure A2.15 BSS at Soyang for SRI12 71
Figure A2.16 BSS at Daecheong for SRI12 71
Figure A2.17 BSS at Andong for SRI12 72
Figure A2.18 BSS at Seomjin for SRI12 72
Figure A2.19 BSS at Chungju for SRI12 73
Figure A2.20 BSS at Hapcheon for SRI12 73
Figure A2.21 BSS at Namgang for SRI12 74
Figure A2.22 BSS at Imha for SRI12 74
Figure A2.23 Averaged RMSE of four EDPs across eight basins for SRI3 75
Figure A2.24 RPSS of four EDPs EDPs for SRI3 76
Figure A2.25 Averaged BSS of four EDPs across eight basins for SRI3 77
Figure A2.26 Reliability diagram of four EDPs SRI3 78
Figure A2.27 Averaged RMSE of four EDPs across eight basins for SRI12 79
Figure A2.28 RPSS of four EDPs EDPs for SRI12 80
Figure A2.29 Averaged BSS of four EDPs across eight basins for SRI12 ...81
Figure A2.30 Reliability diagram of four EDPs SRI12 82

Chapter 1. Introduction

1.1 Problem Statement

Drought, one of the major natural disasters, makes a catastrophic impact on water use in various aspects such as water supply, agriculture, hydro-power generation (Ciais et al., 2005; Grayson, 2013; Mosely, 2015; Van Loon, 2015). Mekonnen and Hoekstra (2016) estimated that about two-thirds of the world population has experienced severe water scarcity. In addition, it is expected that the dry regions will get much drier since global warming has accelerated the hydrological cycle and been resulted in more extremes (Seager et al., 2010; Dai, 2011; Trenberth et al., 2014; Hao et al., 2018;), and Korea is no exception. According to the report from Korea National Drought Information analysis Center (KNDIC), droughts have occurred almost every year since 2000, and even there was a record-breaking multi-year drought from 2013 to 2018 due to lack of precipitation (KNDIC, 2018). Under this circumstance, preparing droughts to prevent catastrophic impacts has become one of the most important challenges for the future.

The hydrological drought prediction is one of the important parts of drought mitigation because it provides drought information to early warning and prevention systems to reduce damages. A high-quality drought prediction can contribute to mitigating drought damage by making the effective operation of reservoirs in Korea including twenty multipurpose dams. The major reasons for uncertainties in drought prediction are from lack of knowledge and nature itself, so the probabilistic approach is required to quantify these uncertainties and thus to derive results that can help decision making in reservoir operation (Demargne et al., 2014). Techniques for predicting the probability of hydrologic conditions, which can take into account both natural and predictive model uncertainties, have been developed over the past decades. Among them, ensemble streamflow prediction (ESP) is most widely used in hydrology since it derives probabilistic forecasts by considering the possible range of streamflow (Palmer, 2017). The ESP, however, makes the simulation of the streamflow which does not directly represent the drought information. Therefore, to obtain the drought information directly, it is necessary to derive the ensemble drought prediction (EDP) by converting the streamflow ensemble into a measure such as drought indices. And in reality, the institutes in the U.S. such as National Oceanic and Atmospheric Administration (NOAA) are making the probabilistic prediction of the meteorological drought through the ensemble method (Yoon et al., 2012; Mo et al., 2019).

In addition, it may be insufficient to make predictions by only referring to the streamflow since the hydrological drought is caused by the interaction of several hydrological factor. In particular, soil moisture has been regarded as a significant factor of the hydrologic process, so studies have been conducted to analyze the impacts of soil moisture on the hydrological drought (Wood and Lattenmaier, 2008; Mahanama et al., 2012; Yuan et al., 2016). Nevertheless, there were insufficient efforts to reflect drought information from soil moisture into the prediction directly.

Meanwhile, the Korean government recently has begun to invest in the improvement of the drought prediction technique, and thus the KNDIC was established in 2016 to provide technical support and to integrate drought forecast and warning systems that were operated by each institution. However, the drought prediction is still being made deterministically without taking into account uncertainties and this can give a false confidence problem in drought management system. Therefore, it is required that the probabilistic drought prediction is introduced in Korea and its advantages should be verified.

1.2 Research Objectives

In Korea, the hydrological drought should be predicted probabilistically, as in the case of other hydro-meteorological conditions, in order to prevent drought effectively. In addition, drought-related information such as soil moisture, whose relationship between the hydrological drought has already been verified, should also be directly reflected in prediction to improve the skill of the predictive model.

Therefore, the purpose of this study is to introduce the probabilistic drought prediction to Korea and to verify the advantage compared to the deterministic approach. Furthermore, in order to improve the drought prediction skill, the drought information from soil moisture is reflected in predicting drought. In the last, the effectiveness of reflecting the information from soil moisture is analyzed by evaluating the drought prediction results.

1.3 Thesis Organization

The literature reviews in chapter 2 focus on probabilistic prediction methods and practical application cases. Chapter 3 introduces the theories to be used in this study, and chapter 4 describes their detailed application methods and results. In the last chapter 5, the main points and conclusions of this study are summarized. Appendix which contains the results of streamflow and drought predictions is located after References.

Chapter 2. Literature Review

Drought cannot be measured or evaluated directly because of its characteristics. Therefore, in general, drought has been measured indirectly through hydrometeorological variables such as precipitation, streamflow, etc. Generally, drought is classified into four types according to aspects of interest, and each definition is as follows (Wilhite and Glantz, 1987; Lloyd-Hughes and Saunders, 2002; Mishra and Singh, 2010; Van Loon and Van Lanen, 2012).

(1) Meteorological drought: Insufficient precipitation

(2) Hydrological drought: Insufficient streamflow (related to precipitation)

(3) Agricultural drought: Drought damage on crops (related to soil moisture)

(4) Socioeconomic drought: Water demand exceeding supply

Among these, hydrological drought is being considered an important issue because it is closely associated with the actual impact on both nature and society (Mishra and Singh, 2011; Cloke and Hannah, 2011; Van Loon, 2015). Referring to the above definition, this study defines the hydrological drought as a situation of low natural streamflow. Based on this background, research cases on hydrological drought, prediction methods, and practical application cases are investigated.

2.1 Drought Measures

The most common method used to measure drought is to derive a drought index representing anomaly levels of dryness through drought-related variables such as precipitation and streamflow. The standardized precipitation index (SPI) is a meteorological drought index indicating precipitation anomaly (Mckee, 1993), and other indices such as standardized runoff index (SRI) and standardized precipitation evaporation index (SPEI) have been developed based on SPI (Shukla and Wood, 2008; Vicente-Serrano et al., 2010). Such drought indices may not contain actual drought information because it just represents anomaly levels of dryness compared to the climatology. Nevertheless, it is widely used due to its spatial and temporal flexibilities and ease of comparison. SRI is often used for measuring the hydrological drought because it considers streamflow. The criteria of SRI for evaluating the depth of drought are usually anomaly levels but sometimes they may be determined by considering water demand.

On the other hand, the hydrological drought is sometimes evaluated through streamflow itself. In this approach, drought properties (duration, deficit, and intensity) corresponding a certain threshold are calculated to represent and analyze severities (Tallaksen et al., 1997; Van Loon, 2015), but this method has less spatial flexibilities than the drought indices because the threshold level should be determined according to streamflow characteristic of the target region, and thus this is commonly used when evaluating drought for a specific region.

2.2 Drought Prediction Methods

The drought prediction is an estimate of how dry in the future. As mentioned earlier, it is common to make predictions through hydrometeorological variables because the drought is measured through them. Therefore, studies and methods in the hydrometeorological prediction, especially associated with drought, are also introduced in this section. The hydrometeorological prediction is generally done in two approaches: deterministic and probabilistic approaches.

2.2.1 Deterministic Approach

Determinism primarily uses dynamical models, which expresses physical mechanisms in the atmosphere, ocean, and continent as mathematical equations, and makes a single-valued prediction. Using dynamical models, however, has a limitation in that the reliability and accuracy decrease exponentially with increasing lead time because the variability of weather conditions is very large (Shukla et al., 2013; Yuan et al., 2015). In order to solve this problem, lots of studies for pre-processing, post-processing, and accurate estimation of initial conditions have been conducted under the lead of research institutes in Europe (Mahanama et al., 2012; Shukla et al., 2014; Wood et al. 2015; Emerton et al., 2016; Yuan et al., 2016; Mendoza et al., 2017).

In the meantime, statistical models have been also used for deterministic prediction. These statistical models derive results in the form of a probability distribution, but in the deterministic approach, they result in a single value through statistics. Streamflow and hydrological drought in Korea were predicted deterministically using a statistical model, resulting in the effective for one- and two-month ahead prediction (Bae et al., 2013; Son and Bae, 2015).

Researchers have demonstrated the disadvantages and limitations of the deterministic approach, advocating the advantages of the probabilistic approach using ensemble prediction (Murphy and Palmer, 1986; Brankovic et al., 1990; Palmer et al., 1993; Molteni et al., 2011). Meanwhile, Krzysztofowicz (2001) pointed out that the catastrophic damage from the great flood in Mississippi was because of a false confidence from the deterministic forecast (NOAA, 1994), and then the necessity of the probabilistic approach in prediction has begun to emerge. Buizza (2008) analyzed the potential economic values from the deterministic and probabilistic streamflow forecasts and as a result, concluding that the deterministic approach could make more loss than the probabilistic one. Besides, determinism has become obsolete due to the development of computation capability (Hao et al., 2018).

2.2.2 Probabilistic Approach

The probabilistic approach produces information about the predicted value and its uncertainty. In 1906, there was the first attempt to quantify uncertainty to predict weather probabilistically (Cooke, 1906), and then it was introduced to practice in the United States in 1969 for the first time.

Probabilistic prediction can be performed in various ways, such as deriving the probability distribution, ensemble prediction method, and deriving the probability of occurrence (Stockdale et al., 2010), among which the ensemble prediction method is the one most widely used. To put it simply, the ensemble prediction is making use of a bunch of deterministic results to make the probability of events. Brankovic et al. (1990) demonstrated that ensemble prediction is more reliable than a single prediction which is deterministic. Traditionally, the ensemble prediction is the entirely statistical method because of the assumption that weather conditions would repeat exactly as they did in the past (Day, 1985), and this approach is still basically adopted. However, there is a problem that the traditional ensemble prediction cannot reflect the actual hydrometeorological conditions at the time of interest, especially with the short lead time. In order to overcome this problem, some researcher began to apply dynamical models to the ensemble prediction. It is a way to regard the results created using many dynamical models as ensemble members. The ensemble prediction is also adopted for the short-term range hydrologic forecast with GCMs (Global Circulation Models) which are the dynamical models for the global scale (Molteni et al., 2011, Saha et al., 2014). However, the limitations of dynamical models mentioned in section 2.2.1 also appear in ensemble prediction. Harrigan et al. (2018) performed streamflow simulation in rivers in the UK using multi-model ensembles from several GCMs, but the accuracy decreased exponentially with increasing lead times.

On the other hand, statistical models such as regression, autoregressive model, Markov chain, machine learning, meta-Gaussian, copula, and their combinations for the probabilistic prediction also have been continuously developed (Mishra et al., 2007; Barros and Bowden, 2008; Durdu, 2010; Hao et al., 2016; Zink et al., 2016).

The statistical models produce probabilistic predictions through the distribution of error terms of results, which are made through statistical assumptions, so thus have limitations in that they cannot predict a possible range of events (Palmer, 2017).

To complement the shortcomings of these dynamical and statistical models, Luo and Wood (2007) argued that it is necessary to combine information from a variety of predictive models, and thus techniques to combine several methodologies are being developed as shown in Figures 2.1 and 2.2 (Hao et al., 2018). Luo et al. (2007) used the Bayesian update proposed by Coelho et al. (2003) to improve the ensemble prediction system for hydrometeorological conditions. It was a method that combines several GCMs and a statistical empirical model using Bayes' theorem. Seo et al. (2019) applied this Bayes' theorem method for the ESP simulation on thirty-five dam watersheds in South Korea, and as a result, the accuracy got improved. As such, lots of studies using statistical techniques are being actively carried out to solve the problem of the ensemble prediction system (Kang et al., 2010; Zhao et al., 2011; Yang et al., 2016; Li et al., 2017). In addition, Qu et al. (2017) performed hydrological prediction in the Fu river of China using BMA (Bayesian Model Averaging) to combine multi-model ensembles. As a result, it resulted in better accuracy than single ensembles and improved accuracy for long lead time.

Ma et al. (2015) found that increasing the prediction accuracy of hydrometeorological variables is directly related to improving the accuracy of drought prediction. However, there are still few studies to apply the methods developed for the hydrometeorology to drought prediction.

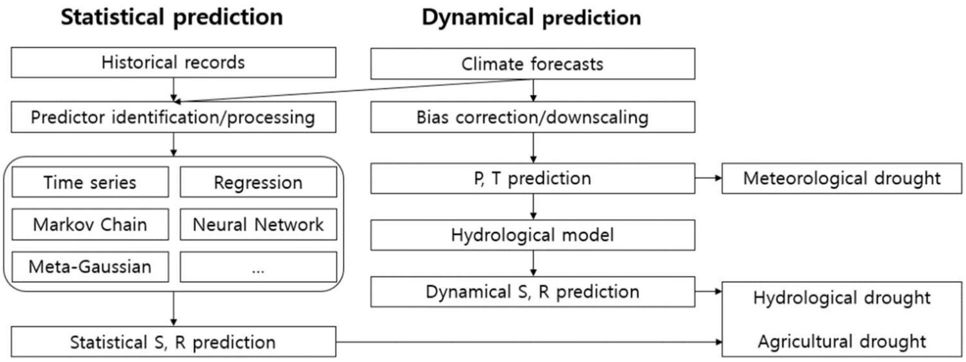


Figure 2.1 Frameworks of prediction methods and their interactions (Hao et al., 2018)

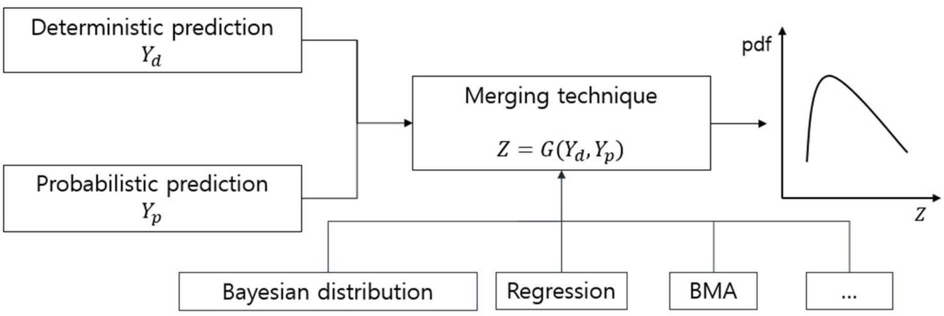


Figure 2.2 Flow diagram of combining prediction methods (Hao et al., 2018)

2.3 Practical Use of Probabilistic Predictions

The examples of the practical use of probabilistic forecast or ensemble prediction in hydrometeorology by institutions around the world are summarized in Table 2.1. As mentioned above, the probabilistic precipitation forecast was first proposed in 1906 (Cooke, 1906), but it was introduced to the practice when the National Weather Service (NWS) began precipitation forecast in 1969. Nowadays, institutions around the world, including Korea, use the ensemble prediction system for the probabilistic prediction. Besides, an international research group called HEPEX (Hydrologic Ensemble Prediction Experiment) is being operated to share and develop the ensemble prediction.

The World Bank and NOAA of the United States are the representative institutes that produce and provide the probabilistic drought prediction. The World Bank produces the world's meteorological drought probabilities in the future using the drought index SPI as shown in Figure 2.3. NOAA makes the probability of future droughts of hydrometeorological variables as shown in Figures 2.4~2.6 and is already using it practically for decision making. In particular, NOAA divides the drought indexes into several phases according to the depth of drought and derives the probability of occurrence for each phase from daily to annual time scale (Mo et al., 2019).

In Europe, thirteen countries of the EU, including the United Kingdom, Germany, and the Netherlands, are collaborated to establish and operate the EFFS (European Flood Forecasting System) to forecast streamflow in major European watersheds. EFFS uses ESP to predict streamflow probabilistically across Europe but not in some regions. In addition, two cooperative research institutes, European Center for Medium-Range Weather Forecasts (ECMWF) and COSMO-LEPS, are working globally as well as Europe and are trying to improve the probabilistic prediction method using the ensemble.

Table 2.1 Institutions performing probabilistic prediction (Cloke and Pappenberger, 2009)

Forecast center	Ensemble NWP input
Advanced Hydrologic Prediction Services (AHPS) from NOAA	US National Weather Service (NOAA)
European Flood Alert System (EFAS) of the European Commission Joint Research Centre	European Centre for Medium Range Weather Forecasts (ECMWF) and Consortium for Small-Scale Modelling-Limited-area Ensemble Prediction System(COSMO-LEPS)
Georgia-Tech/Bangladesh project	ECMWF
Finnish Hydrological Service	ECMWF
Swedish Hydro-Meteorological Service	ECMWF
MAP D-PHASE (Alpine region)/Switzerland	COSMO-LEPS
Vituki (Hungary)	ECMWF
Rijkswaterstaat (The Netherlands)	ECMWF, COSMO-LEPS
Royal Meteorological Institute of Belgium	ECMWF
Vlaamse Milieumaatschappij (Belgium)	ECMWF
Meteo France	ECMWF and Arpege EPS
Land Oberoestereich, Niederoestereich, Salzburg, Tirol (Austria)	Integration of ECMWF into Aladin
Bavarian Flood Forecasting Centre	COSMO-LEPS

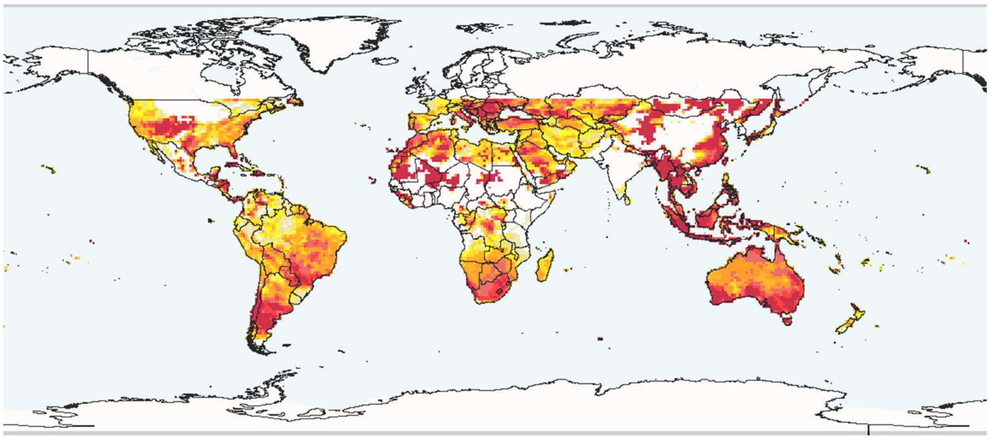


Figure 2.3 Probabilistic SPI prediction by the World Bank
(The World Bank, https://www.worldbank.org/)

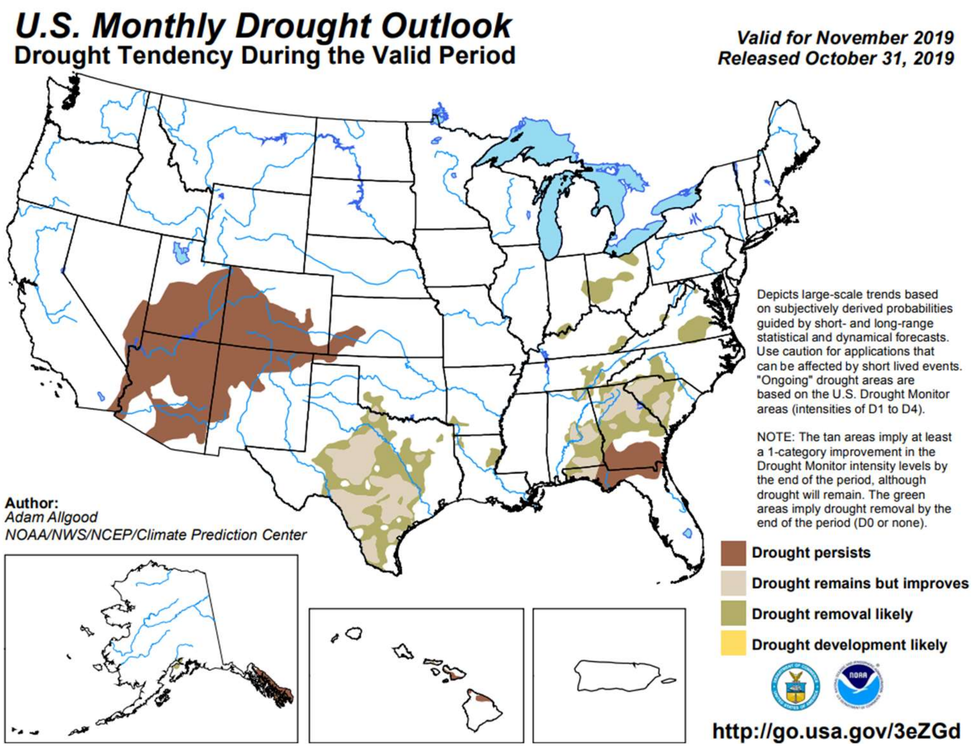


Figure 2.4 Example of monthly drought outlook in the U.S.
(NOAA, https://www.cpc.ncep.noaa.gov/)

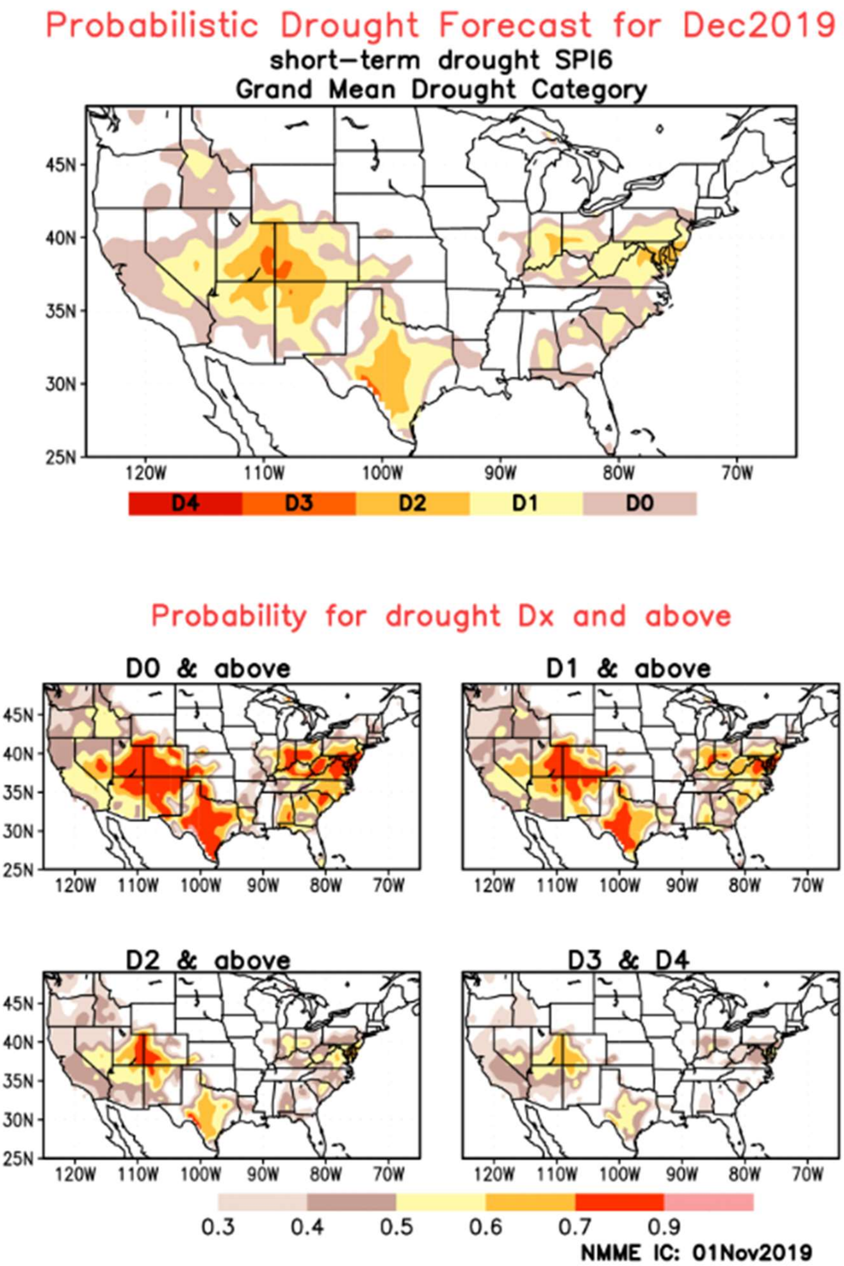
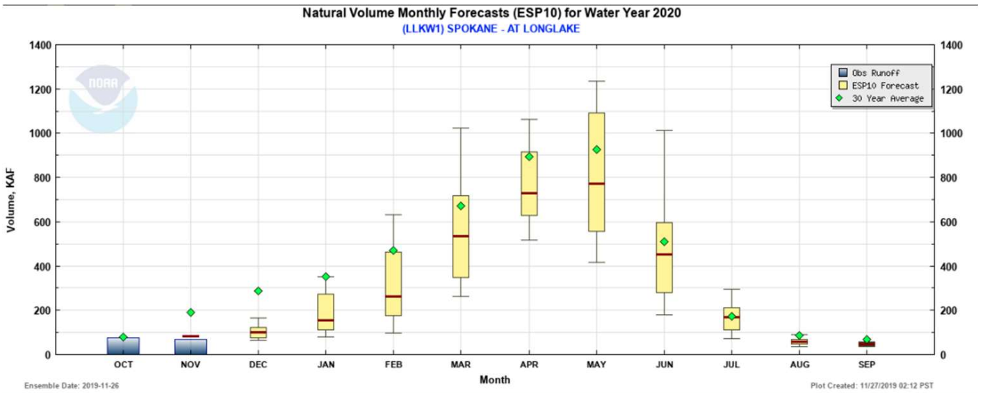


Figure 2.5 Probabilistic drought forecast of each drought phase (Mo et al., 2019)



SPOKANE - AT LONGLAKE
Forecasts For Water Year 2020

Official Water Supply
ESP with 10 days QPF
Ensemble: 2019-11-26 Issued: 2019-11-26

Forecast Period	Forecasts Are in KAF 90 %	50 %	% Average	10 %	Obs Runoff (2019-11-26)	30 Year Average (1981-2010)
OCT					75	78
NOV	83	83	44	83	68*	191
DEC	66	99	35	166		286
JAN	78	153	43	353		353
FEB	98	262	56	633		471
MAR	262	535	80	1022		673
APR	518	728	81	1063		895
MAY	415	773	84	1236		926
JUN	178	453	89	1011		510
JUL	71	169	98	296		172
AUG	37	57	65	91		87
SEP	35	47	67	56		70

Move the mouse over the desired "Forecast Table" to update graph.
* Partial Monthly Total

CSV Download ESP10 Forecast Ensemble

SPOKANE - AT LONGLAKE
Forecasts For Water Year 2020

Experimental Water Supply
HEFS with 15 days EQPF
Ensemble: 2019-11-26 Issued: 2019-11-26

Forecast Period	Forecasts Are in KAF 90 %	50 %	% Average	10 %	Obs Runoff (2019-11-26)	30 Year Average (1981-2010)
OCT					75	78
NOV	83	83	44	83	68*	191
DEC	77	133	46	275		286
JAN	81	185	52	403		353
FEB	104	294	62	766		471
MAR	273	594	88	1085		673
APR	533	755	84	1081		895
MAY	455	818	88	1310		926
JUN	203	464	91	1110		510
JUL	80	179	104	304		172
AUG	39	58	67	94		87
SEP	36	48	69	58		70

Move the mouse over the desired "Forecast Table" to update graph.
* Partial Monthly Total

CSV Download HEFS Forecast Ensemble

SPOKANE - AT LONGLAKE
Forecasts For Water Year 2020

Reference
ESP with 0 days QPF
Ensemble: 2019-11-26 Issued: 2019-11-26

Forecast Period	Forecasts Are in KAF 90 %	50 %	% Average	10 %	Obs Runoff (2019-11-26)	30 Year Average (1981-2010)
OCT					75	78
NOV	83	83	44	89	68*	191
DEC	74	150	52	287		286
JAN	80	189	54	448		353
FEB	101	299	64	754		471
MAR	266	607	90	1147		673
APR	544	777	87	1099		895
MAY	488	819	88	1315		926
JUN	215	465	91	1154		510
JUL	86	181	105	313		172
AUG	41	61	69	96		87
SEP	36	48	69	59		70

Move the mouse over the desired "Forecast Table" to update graph.
* Partial Monthly Total

CSV Download ESP0 Forecast Ensemble

Figure 2.6 Example of probabilistic streamflow forecasts in the U.S. (NOAA, https://water.weather.gov/ahps/)

2.4 Drought Prediction in Korea

The hydrometeorology and drought prediction system of Korea had been independently operated by KMA, the Ministry of Land, Infrastructure, and Transport (MOLIT), Ministry of Environment (ME), and the Ministry of Agriculture, Food and Rural Affairs (MAFRA). Each institution had managed different types of droughts by developing drought index respectively, so it was difficult to prevent the drought beforehand.

The Korean government established KNDIC in 2016 to improve the drought prediction system by recognizing the need for an integrated drought management system after experiencing an unprecedented multi-year drought. The first step, as shown in Figure 2.7, is to produce future drought information for each sector. At this time, the climate forecast produced by KMA is shared with the other institutes. Then, the pieces of drought information are integrated and analyzed and released to the public.

The ensemble method has commonly used for weather and hydrological forecasts in this process. For example, KMA produces weather forecasts using the dynamical model GloSea5 (Global Seasonal Forecasting System 5), and the K-water of ME produces hydrological forecasts using ESP. Although the ensemble method is being used, the prediction for drought is done by the deterministic approach. There is no probability information about the drought condition as shown in Figures 2.8-2.10, a drought forecast conducted by KMA, ME, and APCC Climate Center (APCC). In addition, studies on drought prediction have also been conducted mainly in the deterministic approach (Bae et al., 2013; Son and Bae, 2015).

The studies of the probabilistic prediction for hydrometeorological variables have been ongoing. In addition, the ESP, concept system, was first introduced in Korea in 2001 (Kim et al., 2001), and subsequent studies have continued about the ESP. Seo et al. (2019) upgraded the ESP using the Bayes' theorem and is capable of improving the forecast accuracy. Recently, APCC has produced MME (Multi-Model Ensembles) and used it for the probabilistic prediction (Sohn et al., 2013).

Step	1		2		3
Institution	ME, MAFRA, KMA	➡	Han River Flood Control Office (ME)	➡	National Disaster Management Research Institute
Task	Producing drought information in the future		Integrating pieces of drought information		Examining and publication

Figure 2.7 Drought prediction procedure in Korea

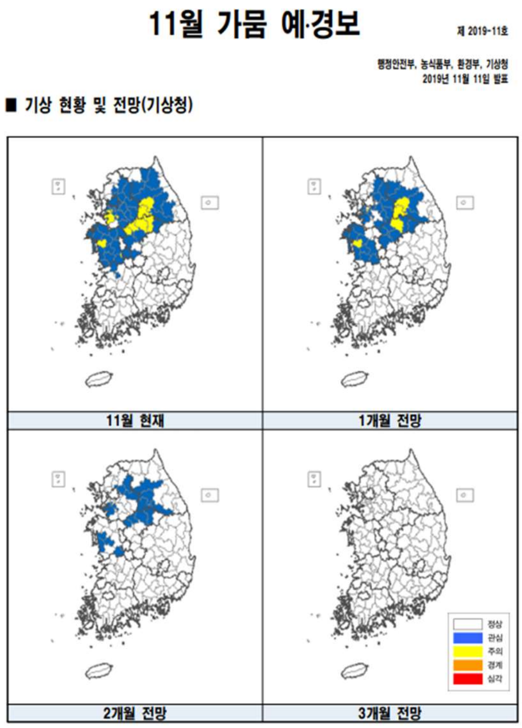


Figure 2.8 Meteorological drought prediction by KMA (https://www.weather.go.kr/)

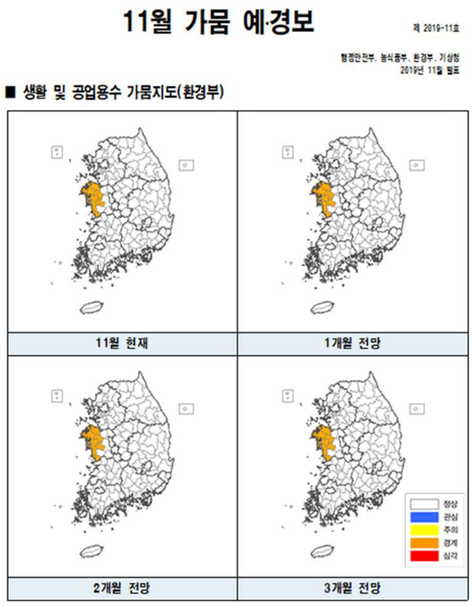


Figure 2.9 Hydrological drought prediction by ME (http://hrfco.go.kr/)

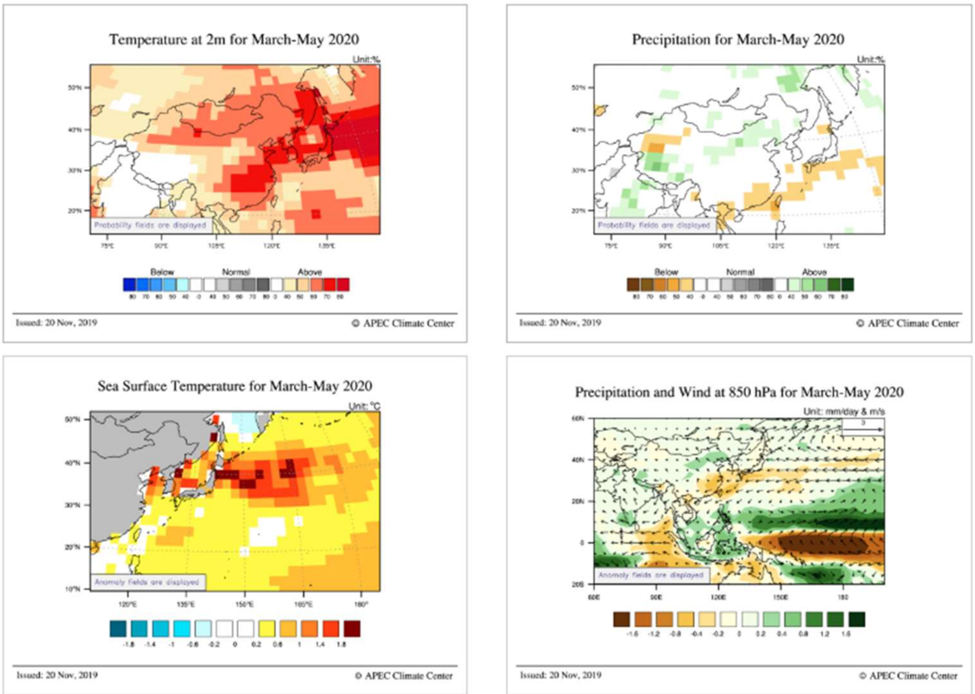


Figure 2.10 Hydrometeorological prediction by APCC (https://www.apcc21.org/)

Chapter 3. Methodology

In this study, the ensemble method is used to predict the probability of drought in the future, and the whole procedure is shown in Figure 3.1. The procedure consists of three main parts: ensemble prediction, empirical model, and Bayesian update. In the ensemble prediction part, EDP is generated by the distribution of SRI which is converted from the ESP simulation results. The empirical model is a regression between SMI (Soil Moisture Index) and SRI. At last, EDP is used as prior information, and the empirical model is used to form the likelihood function. The Bayes' theorem is then applied to produce a posterior distribution which is called EDP+S. The probabilistic and deterministic predictions are derived from the distribution and the expected value of the EDP distribution respectively, and then they are compared through performance metrics.

Additionally, to figure out the availability of making use of climate information on drought prediction, the probabilistic precipitation forecast produced by APCC is reflected in EDP and EDP+S, and they are called EDP+A and EDP+AS respectively. The four EDPs are compared to analyze the effects of SMI and the climate information on drought prediction.

3.1 Ensemble Prediction

3.1.1 Concept of Ensemble

The ensemble consists of a bunch of deterministic prediction series which are called ensemble members. The probabilistic prediction by the ensemble represents the possible range of events rather than just an error bar around a predicted value (Palmer, 2017). Some pre- and post-processing techniques can be used to improve performance (Hamlet and Lattenmaier, 1999; Yao and Georgakakos, 2001; Bradley et al. 2015). In this study, two kinds of post-processing techniques, the Bayesian update and the PDF ratio method are used to improve performance.

3.1.2 Ensemble Streamflow Prediction (ESP)

The ESP is a system that generates streamflow ensembles by inputting climate forcing sampled from observed data to a hydrologic model as shown in Figure 3.2.

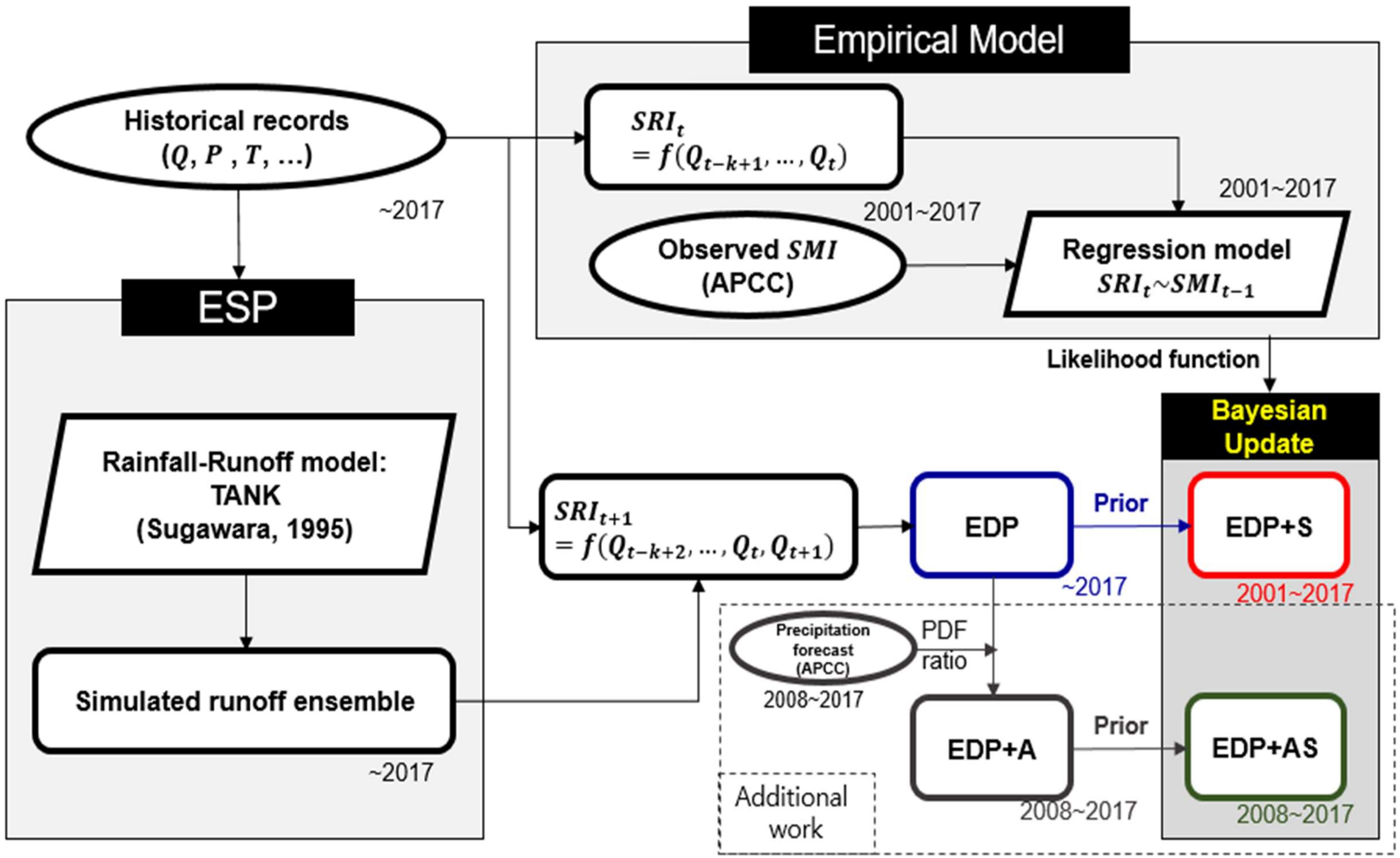


Figure 3.1 Flowchart of probabilistic drought prediction in this study

The ESP consists of three main elements: input ensemble, a hydrologic model with initial conditions, and streamflow ensemble. The initial conditions, such as soil moisture, are estimated using observations just before the time of interest. The number of streamflow ensemble members is equal to the number of input members.

The TANK model which was used in the report Water Vision 2020 (MOLIT, 2016) is used to simulate the natural streamflow at upstream of dam basins. Moreover, to consider the snow accumulation–melting, the modified TANK model by McCabe and Markstrom (2007) which is shown in Figure 3.3 is adopted. The TANK model is a conceptual model to describe the rainfall-runoff process as a structure consisting of four tanks (Sugawara, 1995), and is known to be practical because of small number of required input and parameters. It is suitable for upstream regions since the TANK model produces natural streamflow. The availability of the TANK model is verified for long-term hydrologic simulation in Korea (Kang et al., 2013; Choi et al., 2018; Seo et al., 2018). In case of mid- and long-term forecasts with a lead time longer than 10 days, conceptual hydrologic models such as TANK are appropriate, because the physical models require a long computation time due to complex physical mechanisms.

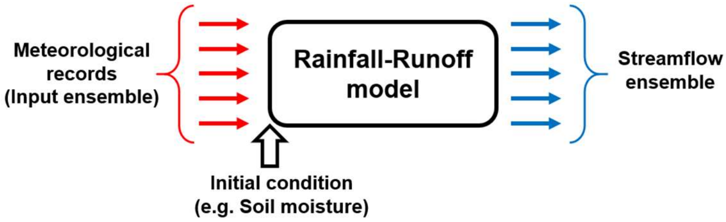


Figure 3.2 Schematic diagram of ESP procedure

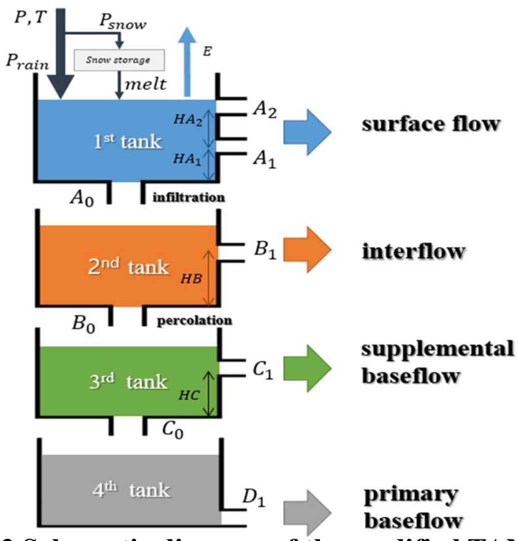


Figure 3.3 Schematic diagram of the modified TANK model

3.1.3 Ensemble Drought Prediction (EDP)

This study determines EDP as an SRI ensemble which is converted from the ESP simulation. As explained in section 2.1, SRI is the drought index used for measuring the hydrological drought. The concept of the drought index was proposed by Mckee et al. (1993) using precipitation, and Shukla and Wood (2008) developed SRI based on that concept. Some probability density functions such as gamma and lognormal can be used to derive SRI easily (Edwards and Mckee, 1997; Shukla and Wood, 2008)

The calculation process of SRI using streamflow (q) is as follows. The first step is calculating the cumulative streamflow Q_j over a given period of k months at month j as Eqn (3.1) where the subscript indicates the month, so when it becomes 0, going down from December of last year. Next, the cumulative probability $F_j(Q_j)$ is estimated with the lognormal function F_j that is already known as appropriate to Q_j. Finally, $F_j(Q_j)$ is converted into the standard normal distribution through Eqns (3.2)~(3.4) to derive SRI_j. For instance, SRI3 is a drought index that represents the anomaly level of cumulative streamflow over a three-month time scale.

$$Q_j = \sum_{i=1}^{k} q_{j-i+1} \tag{3.1}$$

$$t = \begin{cases} \sqrt{\ln\left(\dfrac{1}{F_j(Q_j)^2}\right)}, & 0 < F(Q) \le 0.5 \\ \sqrt{\ln\left(\dfrac{1}{\left(1 - F_j(Q_j)\right)^2}\right)}, & 0.5 \le F(Q) < 1 \end{cases} \tag{3.2}$$

$$SRI_j = \begin{cases} -\left(t - \dfrac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right), & 0 < F(Q) \le 0.5 \\ t - \dfrac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}, & 0.5 \le F(Q) < 1 \end{cases} \tag{3.3}$$

$$\begin{gathered} c_0 = 2.515517,\ c_1 = 0.802583,\ c_2 = 0.010328 \\ d_1 = 1.432788,\ d_2 = 0.189269,\ d_3 = 0.001308 \end{gathered} \tag{3.4}$$

If the above process is expressed as a function $g(\cdot)$, the 1-month lead EDP which is obtained by converting the 1-month lead ESP simulation into SRI can be expressed as Eqn. (3.5) where q_{t+1} is the streamflow ensemble from the ESP, k is a given time scale, and μ_0 and σ_0^2 are the mean and standard deviation of the EDP distribution.

$$\mathrm{EDP_{t+1}} = g(q_{t-k+2},\ \dots,\ q_t,\ q_{t+1}) \sim N(\mu_0, \sigma_0^2) \tag{3.5}$$

The probability of drought occurrence is calculated from the distribution of EDP. This study divides drought into four phases and carries out prediction in terms of the multi-categorical and dichotomous events. The multi-category indicates what level of drought among four phases would occur, and the dichotomous event indicates whether drought above a certain phase occurs or not. The drought phases are determined as shown in Table 3.1, which is based on the general ongoing studies about drought. Let the distribution of EDP is $f(x)$, then the probability between lower bound x_l and upper bound x_u can be calculated as Eqn. (3.6).

$$Prob(x_l < x < x_u) = \int_{x_l}^{x_u} f(x)\,dx \tag{3.6}$$

Table 3.1 Drought classification criteria

Range	Phase	Probability	Cumulative probability
$\mathrm{SRI} > 0$	No drought	0.500	1.000
$-1 < \mathrm{SRI} \le 0$	D0	0.341	0.500
$-1.5 < \mathrm{SRI} \le -1$	D1	0.092	0.159
$-2 < \mathrm{SRI} \le -1.5$	D2	0.044	0.067
$\mathrm{SRI} \le -2$	D3	0.023	0.023

3.2 Bayes' Theorem

Simply speaking, Bayesian inference is that prior knowledge can be updated with new information provided as a form of the likelihood. The frequentist inference is based on deductive inference, but the Bayesian inference makes inductive inference. It mathematically consists of three elements: prior distribution, likelihood function, and posterior distribution, as written in Eqn (3.7).

$$p(D|X) = \frac{p(X|D)p(D)}{p(X)} \tag{3.7}$$

where D is a random variable of interest (i.e., drought index in this study), X is new information for the random variable of interest(i.e., soil moisture in this study), $p(D)$ is the prior distribution, $p(X|D)$ is the likelihood function, $p(X)$ is the marginal distribution of X, and $p(D|X)$ is the posterior distribution. In general, the ESP model is known to have a problem of not being able to estimate the initial conditions well, so this study tries to improve EDP by reflecting soil moisture information via the Bayes' theorem.

3.2.1 Prior Distribution

The prior distribution $p(D)$ is derived from the distribution of EDP. Since EDP is a bunch of drought indexes that follow the standard normal distribution, it can be expressed as Eqn (3.8) where μ_0 and σ_0^2 are the mean and standard deviation of EDP, respectively.

$$p(D) \sim N(\mu_0, \sigma_0^2) \tag{3.8}$$

3.2.2 Likelihood function

The likelihood function is the conditional probability of X given D, where the random variable X is the soil moisture index (SMI) which is the satellite observation data being provided by APCC since 2001. In other study cases, the likelihood function was usually estimated from the past performance of the ensemble model (Luo et al., 2007; Seo et al., 2019). In this study, the likelihood function is estimated by the time series regression between two random variables X and D as shown in Eqn (3.9) to reflect the information to EDP.

$$X_t = b_0 + b_1 D_{t+1} + \epsilon \tag{3.9}$$

where the subscript t means the unit time (month), and b and ϵ are regression parameters and residuals, respectively. The residual ϵ follows a normal distribution that has a zero-mean and standard deviation σ_ϵ, so the regression model can be expressed as Eqn (3.10). The parameters of the regression are estimated monthly, but the notation is omitted for convenience. The k-fold cross-validation method is often used to solve problems such as overfitting that may occur due to the small amount of data.

$$p(X_t|D_{t+1}) \sim N(b_0 + b_1 D_{t+1}, \sigma_\epsilon^2) \tag{3.10}$$

3.2.3 Posterior Distribution

Based on the Bayes' theorem, the posterior distribution follows a normal distribution as written in Eqn (3.11) when both the prior distribution and the likelihood function follow normal (Lee, 1997; Coelho et al., 2004). The parameters of the posterior distribution can be derived by Eqns (3.12)~(3.13) which can be interpreted as a kind of variance weighted average of the prior and likelihood.

$$p(D_{t+1}|X_t) \sim N(\mu_p, \sigma_p^2) \tag{3.11}$$

$$\frac{1}{\sigma_p^2} = \frac{1}{\sigma_o^2} + \frac{b_1^2}{\sigma_\epsilon^2} \tag{3.12}$$

$$\frac{\mu_p}{\sigma_p^2} = \frac{\mu_o}{\sigma_o^2} + \frac{b_1^2}{\sigma_\epsilon^2}\left(\frac{X_t - b_0}{b_1}\right) \tag{3.13}$$

3.3 Performance Measures

The skill of the drought prediction is evaluated in two ways as well as deterministic and probabilistic approaches. RMSE (Root Mean Squared Error) is used to measure the accuracy in the deterministic perspective, and the score metrics RPSS (Rank Probability Skill Score) and BS (Brier Score) are in the probabilistic perspective.

3.3.1 Deterministic Approach

RMSE combining bias and variability is used to evaluate the skill of EDP in terms of determinism. If $\bar{P}$ is a single-valued prediction by the expected value of EDP, RMSE can be calculated as written in Eqn (3.14) where O is the observed value and N is the total number of predictions.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\bar{P}_i - O_i)^2} \tag{3.14}$$

3.3.2 Probabilistic Approach

In this study, the skill scores, RPSS and BS are used for evaluation of probabilistic prediction. RPSS is for multi-categorical outcomes, and BS is one for binary outcomes. They are calculated by differences between the predicted probability and occurrences (i.e., 0 or 1). The single-valued predictions also can be evaluated in the probabilistic approach if they are treated as categorical.

(1) Rank Probability Skill Score

RPSS is a skill score that evaluates a benefit compared to the climatologic prediction and is derived from RPS (Rank Probability Score), a score for multi-categorical outcomes. RPS is the most commonly used measure that is capable of penalizing predictions increasingly, as more probability is assigned to event categories further removed from the actual outcome (Wilks, 2011). RPS is derived from the squared errors computed with respect to the cumulative probabilities in the predictions and observations. Let L be the number of categories (i.e., the number of drought phases in this study), then each category has the predicted probability p_j but the observation o_j takes the value 1 in only one category and 0 otherwise. The cumulative probability of p_j and one of o_j are then defined as Eqn (3.15) and Eqn (3.16) respectively. RPS is the mean of the sum of the squared difference between P_l and O_l.

$$P_l = \sum_{j=1}^{l} p_j \,, \; l = 1, \dots, L \tag{3.15}$$

$$O_l = \sum_{j=1}^{l} o_j \,, \; l = 1, \dots, L \tag{3.16}$$

$$\mathrm{RPS} = \frac{1}{N}\sum_{i=1}^{N}\sum_{l=1}^{L}(P_l - O_l)^2 \tag{3.17}$$

Let RPS_0 be the reference taken by the climatology, RPSS is then calculated as shown in Eqn (3.18). RPSS becomes 1 if it is a perfect prediction, and if the prediction model is worse than climatological prediction it becomes a negative value.

$$\mathrm{RPSS} = 1 - \frac{\mathrm{RPS}}{\mathrm{RPS}_0} \tag{3.18}$$

(2) Brier Score

BS is a score for dichotomous events and a kind of reduced version of RPS. It is very similar to RPS, as shown in Eqn (3.19). BS is the mean of the sum of squared differences between the predicted probability value p_i and the observed occurrence o_i taking the value 1 if the drought occurs and 0 otherwise.

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2 \tag{3.19}$$

BS can be further decomposed into three terms: reliability (REL), resolution(RES), and uncertainty(UNC) which can be expressed in Eqn (3.20). This is called calibration-refinement decomposition to evaluate how well the probabilistic prediction is calibrated (Murphy and Winkler, 1987).

$$\begin{aligned} \mathrm{BS} &= E_P\left[\left(\mu_{O|P} - P\right)^2\right] - E_P\left[\left(\mu_{O|P} - \mu_O\right)^2\right] + \mu_O(1 - \mu_O) \\ &= \mathrm{REL} - \mathrm{RES} + \mathrm{UNC} \end{aligned} \tag{3.20}$$

where $\mu_{O|P}$ is the relative frequency corresponding predicted probability P, and μ_O is the observed frequency. The REL term $E_P\left[\left(\mu_{O|P} - P\right)^2\right]$ quantifies how well the probability predictions are consistent with the corresponding observed frequencies. The RES term can be expressed as the second term $E_P\left[\left(\mu_{O|P} - \mu_O\right)^2\right]$ that indicates a kind of confidence in the prediction by quantifying the variability of observed frequencies around the climatological probability. The last term $\mu_O(1 - \mu_O)$ is the UNC that represents the uncertainty of the events and does not relate to predictions. In a perfect prediction, REL becomes 0, and RES becomes equal to UNC. Simply, the REL term is similar to bias, and the UNC-REL term is similar to variability in a deterministic perspective. Drawing a reliability diagram, all three components of BS can be presented at the same time.

BSS, a relative measure of probabilistic skill to the reference BS, can be defined as Eqn (3.21) because the reference BS is equal to the UNC.

$$\mathrm{BSS} = 1 - \frac{\mathrm{BS}}{\mathrm{UNC}} \tag{3.21}$$

Chapter 4. Application

In this chapter, EDP was applied to the eight dam basins in Korea to make drought predictions in both the deterministic and probabilistic approaches, and then those results were compared. In addition, EDP+S was calibrated by incorporating the soil moisture information into EDP, and the effect of soil moisture information was analyzed. Additionally, the probabilistic precipitation forecasts of APCC using multi-model ensembles (MME) were used for updating EDP and EDP+S with the PDF ratio method, and then the skills were evaluated to analyze the effect of the climate information on the drought prediction.

4.1 Study Area

It is generally recommended to use observation data of more than 30 years to make the hydrological drought index. Therefore, this study selected eight basins where their dams have been operated for more than thirty years. The locations and information of each basin are shown in Figure 4.1 and Table 4.1, and numbers from 1 to 8 are assigned to each basin for convenience.

The Soyanggang and the Chungju dams have major roles in water resource management, flood control system across the Han River. Besides, they are capable of hydroelectric power generation so contributing in many ways to Korean society. The Daecheong dam, the second largest in Korea, plays a key role in flood control and water management across the Geum River basin. The Seomjingang dam is small, but it is the first multi-purpose dam in Korea, which was constructed for stable agricultural water supply to the Jeolla-do, Korea's granary. The Andong, Imha, Hapcheon, and Namgang dams located in the Nakdong River area were constructed for water supply and management at the time when various industries were developed actively throughout the Gyeongsang-do, and they also have played an important role in river maintenance. To sum up, all eight basins to which EDP was applied have very major roles in water resource management and many other aspects.

Table 4.1 Information about basins

Number	Dam basin	Area [km^2]	Annual inflow [$10^6 m^3$]	Period	Source
1	Soyang	2,703	2,148	1973 ~ 2017	K-water
2	Daecheong	3,204	2,722	1981 ~ 2017	
3	Andong	1,584	950	1977 ~ 2017	
4	Seomjin	763	502	1975 ~ 2017	
5	Chungju	6,648	4,872	1986 ~ 2017	
6	Hapcheon	925	573	1989 ~ 2017	
7	Namgang	2,285	2,031	1976 ~ 2017	
8	Imha	1,361	545	1992 ~ 2017	

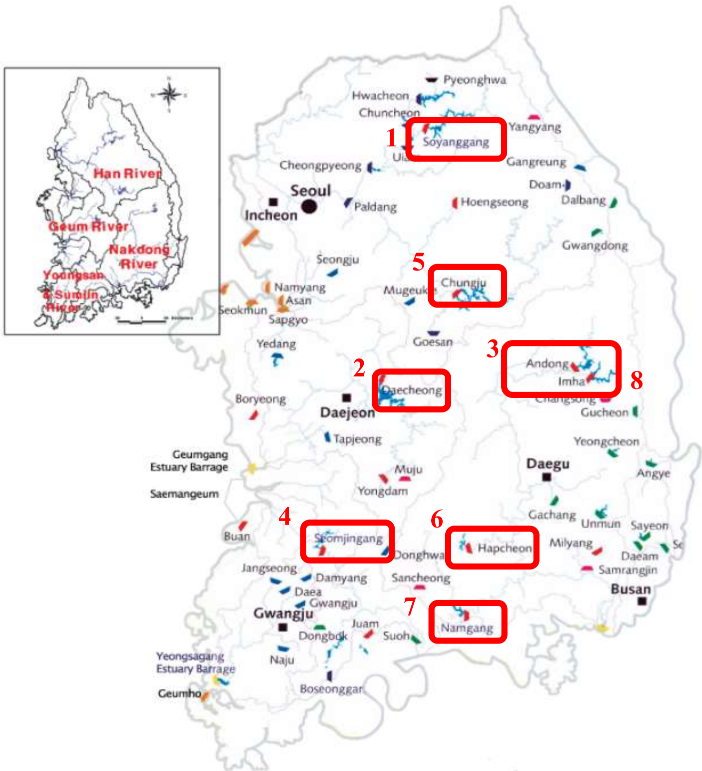


Figure 4.1 Watersheds in Korea (Korea National Committee on Large Dams, https://www.kncold.or.kr/)

4.2 Data Sets

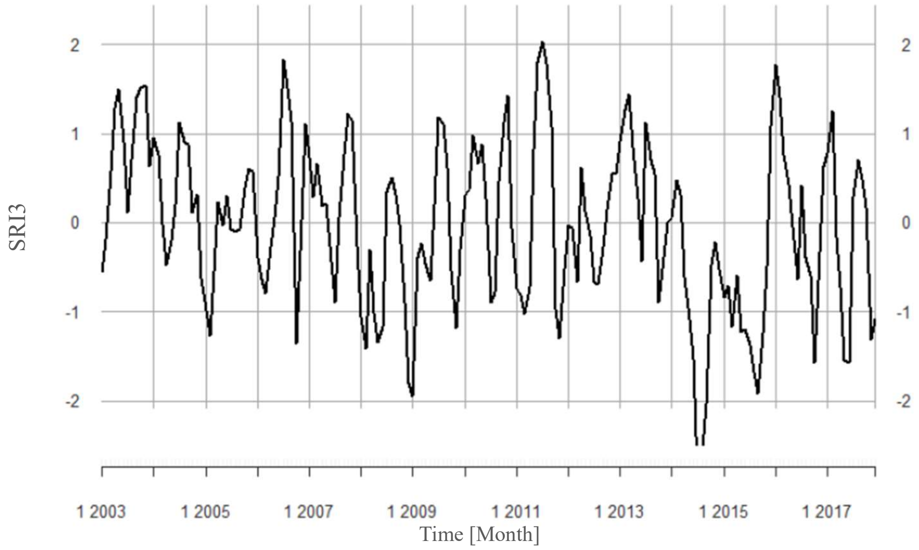
4.2.1 Observed Data Sets

The observed dataset of precipitation, temperature, and streamflow were provided by K-water. Since the observed potential evapotranspiration data are very limited, it should be estimated using the Penman-Monteith equation known to be the best in dry and wet regions (Jensen et al., 1990; Cai et al., 2007). The observations were used to estimate the TANK model parameters, and also for input climate ensembles to the ESP.

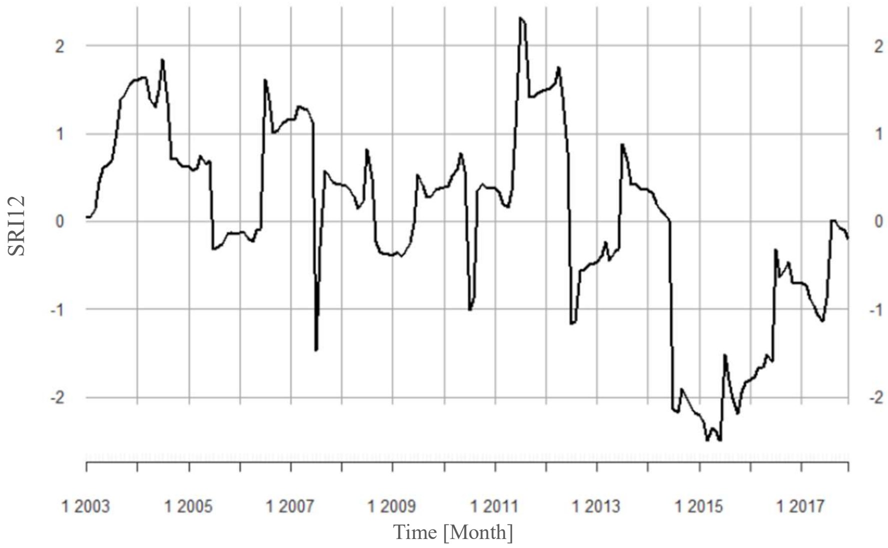
In this study, the observed drought indexes SRI3 and SRI12 were derived using the observed streamflow. These two timescales (3 and 12) are known to best represent Korean short-term and long-term droughts, respectively (Son et al., 2011). For example, SRI3 and SRI12 of Soyang are shown in Figure 4.2. In Korea, from 2014 to 2017, a multi-year drought had occurred due to the lack of precipitation during the summer, and this continued to affect streamflow of winter season and next year. As shown in Figure 4.2, SRI3 hits the lowest value in mid-2014 and continuously presented droughts every summer since 2014, and SRI12 was negative value consistently since 2014, and this well indicated that droughts in summer have impacts on winter. This trend can be found not only at Soyang but also at other basins, so it can be said that the drought indices describe this multi-year drought well.

4.2.2 ESP Dataset

As described in 3.1, EDP is derived by converting the 1-month lead ESP simulations into SRI. In this study, the ESP simulations were produced by using the TANK model of which the parameters were estimated by Seo and Kim (2018) using the SCE-UA (Shuffled Complex Evolution-University of Arizona) algorithm. The parameters of the TANK model were estimated using the observed data sets until 2000, and the model performance was validated by comparing the simulated streamflow with the observed streamflow from 2001 to 2017. NSE values for dam basins in Korea ranged from -0.03 to 0.45, and normalized RMSE (N-RMSE) values ranged from 1.19 to 1.68. If the NSE value is above 0, corresponding prediction is regarded as better than the climatology, and if it is 1, the prediction is perfect. The ESP simulation results are represented in detail in Appendix A-1.



(a) SRI3



(b) SRI12

Figure 4.2 SRIs on Soyang

4.2.3 Soil Moisture Index (SMI)

The APCC has been providing SMI satellite data, a kind of remote sensing data, throughout East Asia since 2001. The SMI is derived from an empirical equation between the relationship between surface temperature and land cover (vegetation). More details about SMI are explained in the paper by Sridhar et al. (2007). Due to the characteristics of the satellite data, there may have severe bias, but in this study, that problem does not matter because SMI is used as the likelihood function via the regression with SRI.

As shown in Figure 4.3, SMI for the study basins should be extracted from the raster data across the whole of East Asia. It cannot represent Korea precisely because the spatial resolution is 1°, so it cannot contain the values of South Korea in detail. Before the time series regression analysis, a cross-correlation analysis between SRI and SMI was done at each basin to check the applicability of the regression. As a result, the lag-1 cross-correlation coefficients at all basins were greater than the critical value 0.136 at a significance level of 5% as shown in Table 4.2, so the regression analysis is possible.

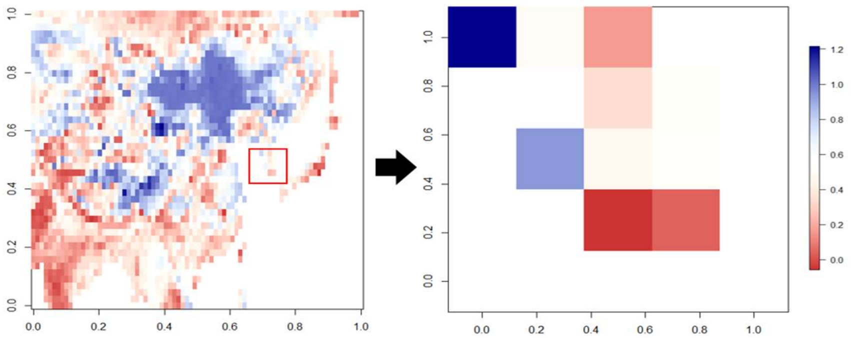


Figure 4.3 SMI data provided by APCC

Table 4.2 Lag-1 correlation coefficients between SRI and SMI

Number	Dam basin	SRI3	SRI12
1	Soyang	0.489	0.260
2	Daecheong	0.516	0.219
3	Andong	0.519	0.262
4	Seomjin	0.473	0.155
5	Chungju	0.506	0.288
6	Hapcheon	0.439	0.142
7	Namgang	0.436	0.155
8	Imha	0.474	0.270

4.3 EDP with SMI

4.3.1 Modelling Framework

EDP for 1-month ahead drought prediction was produced by the procedure described in 3.1.3. Next, in order to generate EDP+S, the likelihood function must be estimated from the time series regression model between SRI and SMI as written in Eqns (3.9) and (3.10) where the parameters are estimated by Eqns (4.1) and (4.2).

$$b_0 = \overline{X_t} - b_1 \overline{D}_{t+1} \tag{4.1}$$

$$b_1 = \frac{\sum (X_{t,i} - \overline{X_t})(D_{t+1,i} - \overline{D_{t+1}})}{\sum \left(X_{t,i} - \overline{X_t}\right)^2} \tag{4.2}$$

where $\overline{X_t}$ and $\overline{D_{t+1}}$ are the average value of SMI and SRI at the month t, respectively. There exists a small number of SMI data because it has been recorded since 2001, so the regression analysis may have a overfitting problem. In order to resolve this problem, the datasets were divided into four to apply the 4-fold cross-validation. The calibration and validation sets of each fold were as shown in Table 4.3. As a result of the regression analysis for each fold, RMSE values of calibration and verification sets were randomly distributed and the differences between folds were small enough assume that there is no overfitting at the likelihood function for all the study basins as shown in Table 4.4.

In this study, to model a more robust likelihood function, the time series regression was fitted with an average of the parameters of 4 folds. All of the parameters (b_0, b_1) and uncertainty (σ_ϵ^2) of the likelihood function for SRI3 and SRI12 were summarized in Table 4.5. Figure 4.4 shows an example of applying EDP and EDP+S to SRI3 at Soyang in July 2014. In the graph, EDP+S is determined by that EDP shifted to the left slightly because the mean value of the likelihood function is located at the left side of EDP and the variance of it is significantly larger than that of EDP. If the variances of the likelihood function is equal to that of EDP, the mean of EDP+S is located at the middle of the prior and likelihood mean values.

As mentioned before, the probabilistic prediction is derived by the calculated probability from the distribution, and the deterministic prediction is derived by the expected value of the distribution. Figure 4.5 is an example of applying EDP and EDP+S at Soyang, presenting only the expected value for convenience.

Table 4.3 4-fold cross-validation data cases

Data set	k1	k2	k3	k4
Calibration period	2001 ~ 2013	2001 ~ 2009, 2014 ~ 2017	2001 ~ 2005, 2010 ~ 2017	2005 ~ 2017
Validation period	2014~ 2017	2010 ~ 2013	2006 ~ 2009	2001 ~ 2004

Table 4.4 RMSE of monthly time series regression of each fold
(a) SRI3

Data sets		Dam basin number							
Period	Fold	1	2	3	4	5	6	7	8
Calibration period	k1	0.200	0.138	0.167	0.156	0.203	0.159	0.150	0.174
	k2	0.189	0.159	0.167	0.153	0.191	0.158	0.150	0.169
	k3	0.210	0.170	0.175	0.153	0.206	0.153	0.147	0.180
	k4	0.194	0.158	0.160	0.160	0.188	0.153	0.146	0.169
Validation period	k1	0.200	0.251	0.203	0.192	0.171	0.190	0.181	0.205
	k2	0.245	0.181	0.193	0.188	0.234	0.178	0.165	0.215
	k3	0.152	0.121	0.173	0.192	0.166	0.211	0.188	0.183
	k4	0.212	0.171	0.209	0.158	0.231	0.192	0.175	0.223

(b) SRI12

Data sets		Dam basin number							
Period	Fold	1	2	3	4	5	6	7	8
Calibration period	k1	0.236	0.171	0.196	0.188	0.232	0.184	0.182	0.198
	k2	0.210	0.184	0.197	0.178	0.209	0.178	0.176	0.197
	k3	0.240	0.198	0.188	0.192	0.235	0.185	0.184	0.193
	k4	0.224	0.183	0.189	0.188	0.223	0.180	0.179	0.185
Validation period	k1	0.241	0.268	0.242	0.219	0.230	0.215	0.209	0.228
	k2	0.306	0.221	0.215	0.240	0.294	0.218	0.216	0.222
	k3	0.179	0.164	0.250	0.194	0.189	0.215	0.202	0.248
	k4	0.238	0.218	0.246	0.208	0.234	0.218	0.210	0.299

Table 4.5 Parameters of likelihood function (monthly time series regression)
(a) SRI3

$\mathrm{SMI_t}$~$\mathrm{SRI3_{t+1}}$		Dam basin number							
Parameter	Month	1	2	3	4	5	6	7	8
Intercept (b_0)	1	0.437	0.415	0.406	0.273	0.445	0.298	0.290	0.365
	2	0.406	0.380	0.365	0.197	0.431	0.200	0.171	0.331
	3	0.460	0.449	0.405	0.272	0.485	0.261	0.226	0.361
	4	0.624	0.579	0.531	0.404	0.610	0.378	0.370	0.482
	5	0.599	0.412	0.441	0.349	0.559	0.394	0.378	0.436
	6	0.464	0.378	0.396	0.393	0.409	0.407	0.423	0.384
	7	0.191	0.167	0.193	0.206	0.201	0.253	0.260	0.192
	8	0.552	0.667	0.627	0.721	0.594	0.712	0.716	0.626
	9	0.628	0.791	0.728	0.768	0.665	0.720	0.721	0.722
	10	0.693	0.742	0.745	0.660	0.703	0.679	0.692	0.717
	11	0.484	0.460	0.511	0.398	0.486	0.426	0.428	0.486
	12	0.417	0.386	0.410	0.269	0.421	0.307	0.313	0.383
Slope (b_1)	1	0.105	0.128	0.135	0.118	0.111	0.092	0.093	0.133
	2	0.106	0.138	0.128	0.134	0.139	0.129	0.142	0.113
	3	0.168	0.164	0.144	0.222	0.145	0.177	0.201	0.096
	4	0.244	0.207	0.191	0.126	0.206	0.136	0.131	0.150
	5	0.207	0.100	0.066	0.098	0.155	0.108	0.114	0.070
	6	0.207	0.174	0.133	0.157	0.159	0.093	0.120	0.136
	7	0.096	0.073	0.070	0.141	0.122	0.115	0.113	0.097
	8	0.208	0.129	0.093	0.096	0.180	0.120	0.129	0.104
	9	0.145	0.094	0.117	0.094	0.111	0.097	0.102	0.119
	10	0.149	0.243	0.217	0.151	0.174	0.186	0.166	0.269
	11	0.101	0.092	0.125	0.043	0.109	0.035	0.034	0.146
	12	0.053	0.114	0.136	0.118	0.061	0.084	0.087	0.126
Uncertainty (σ_ϵ^2)	1	0.153	0.109	0.131	0.138	0.158	0.137	0.132	0.143
	2	0.146	0.095	0.141	0.130	0.129	0.137	0.126	0.154
	3	0.141	0.145	0.147	0.187	0.141	0.185	0.169	0.161
	4	0.228	0.131	0.171	0.162	0.219	0.166	0.142	0.196
	5	0.236	0.138	0.179	0.148	0.266	0.145	0.139	0.184
	6	0.177	0.195	0.175	0.244	0.217	0.242	0.230	0.185
	7	0.281	0.193	0.256	0.242	0.277	0.245	0.235	0.249
	8	0.299	0.167	0.236	0.162	0.308	0.187	0.181	0.237
	9	0.268	0.129	0.148	0.122	0.272	0.132	0.128	0.151
	10	0.202	0.112	0.156	0.138	0.171	0.117	0.118	0.160
	11	0.136	0.136	0.123	0.090	0.123	0.079	0.077	0.138
	12	0.119	0.125	0.161	0.105	0.127	0.123	0.118	0.171

Table 4.5(continued) (b) SRI12

SMI_t~$SRI12_{t+1}$		Dam basin number							
Parameter	Month	1	2	3	4	5	6	7	8
Intercept (b_0)	1	0.399	0.343	0.325	0.219	0.401	0.219	0.217	0.325
	2	0.457	0.390	0.362	0.288	0.457	0.266	0.265	0.363
	3	0.559	0.492	0.477	0.363	0.564	0.362	0.356	0.478
	4	0.522	0.365	0.423	0.327	0.531	0.384	0.380	0.430
	5	0.378	0.290	0.354	0.333	0.386	0.399	0.397	0.357
	6	0.190	0.165	0.183	0.197	0.195	0.252	0.250	0.175
	7	0.572	0.682	0.622	0.706	0.578	0.695	0.688	0.612
	8	0.661	0.802	0.740	0.776	0.668	0.722	0.721	0.730
	9	0.692	0.779	0.746	0.648	0.702	0.673	0.674	0.739
	10	0.487	0.473	0.513	0.398	0.494	0.428	0.426	0.512
	11	0.405	0.341	0.385	0.272	0.412	0.312	0.309	0.388
	12	0.064	0.023	0.061	-0.035	0.049	-0.006	0.000	0.088
Slope (b_1)	1	0.030	0.023	0.037	-0.005	0.032	0.015	0.019	0.061
	2	0.015	0.014	0.019	0.001	0.023	-0.009	-0.001	0.044
	3	0.056	0.049	0.022	0.074	0.043	0.004	0.028	0.040
	4	0.075	0.011	0.002	0.005	0.070	-0.052	-0.017	0.002
	5	0.060	-0.057	0.019	0.018	0.053	-0.068	-0.036	0.022
	6	0.051	0.054	0.083	0.101	0.087	0.069	0.081	0.093
	7	0.153	0.166	0.130	0.119	0.138	0.102	0.139	0.128
	8	0.141	0.090	0.129	0.077	0.103	0.092	0.104	0.127
	9	0.134	0.229	0.189	0.183	0.164	0.189	0.175	0.228
	10	0.077	0.086	0.094	0.014	0.104	-0.006	-0.004	0.116
	11	0.044	0.026	0.051	-0.008	0.050	0.004	0.011	0.070
	12	0.399	0.343	0.325	0.219	0.401	0.219	0.217	0.325
Uncertainty (σ_ϵ^2)	1	0.178	0.146	0.175	0.159	0.181	0.139	0.139	0.158
	2	0.136	0.118	0.172	0.155	0.137	0.155	0.154	0.153
	3	0.169	0.175	0.184	0.269	0.169	0.251	0.250	0.171
	4	0.259	0.208	0.232	0.219	0.264	0.210	0.207	0.226
	5	0.250	0.143	0.193	0.183	0.255	0.185	0.188	0.190
	6	0.236	0.190	0.203	0.256	0.242	0.224	0.233	0.204
	7	0.285	0.213	0.171	0.206	0.274	0.198	0.192	0.178
	8	0.325	0.189	0.178	0.201	0.329	0.219	0.201	0.207
	9	0.268	0.243	0.166	0.180	0.280	0.187	0.174	0.189
	10	0.245	0.221	0.214	0.158	0.216	0.143	0.157	0.188
	11	0.222	0.218	0.232	0.154	0.204	0.148	0.148	0.220
	12	0.136	0.167	0.213	0.146	0.125	0.137	0.138	0.200

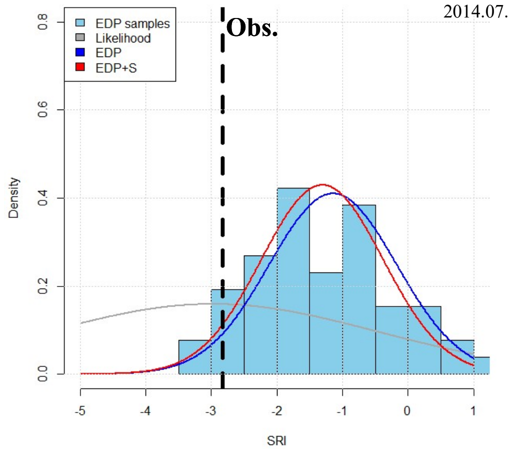
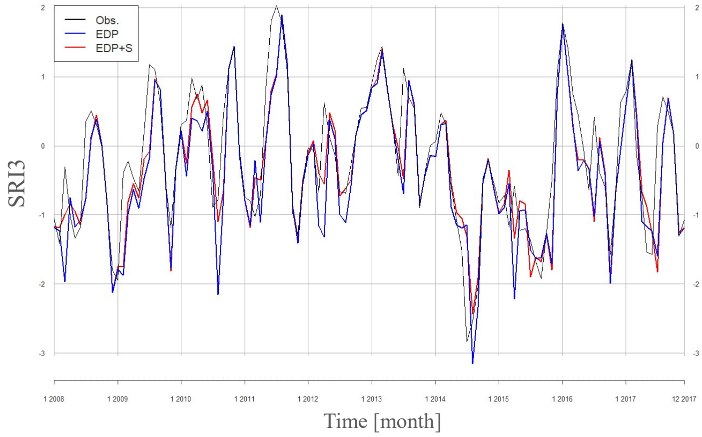
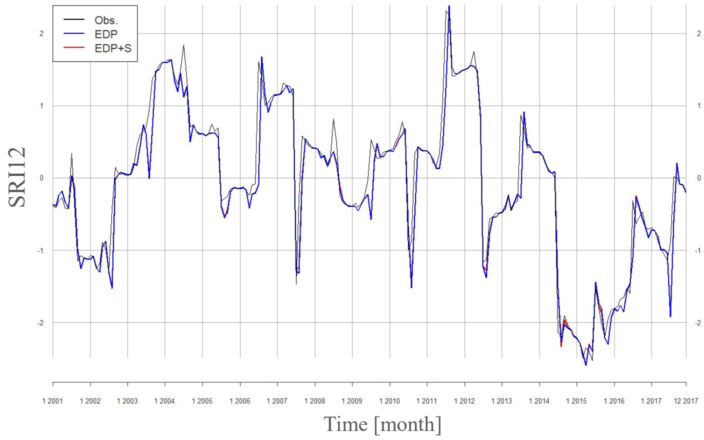


Figure 4.4 Example of Bayesian update with SMI at Soyang



(a) SRI3



(b) SRI12

Figure 4.5 Example of EDPs at Soyang

4.3.2 Results and Discussion

Using EDP and EDP+S, droughts from 2001 to 2017 were predicted at a time one month ahead in both probabilistic and deterministic perspectives, and the skills were evaluated also for irrigation and non-irrigation periods, respectively. As explained before, to derive RPSS and BS of the deterministic prediction, it is necessary to convert the mean value of the EDP distribution into an occurrence or not (i.e., 0 or 1) according to the criteria shown in Table 3.1. For instance, if the mean value of the EDP distribution is -1.3, it is equal to that D1 phase drought will occur 100% in the deterministic approach.

According to the overall results, the larger the basin area, the lower the skill. This may be because of the influence of the ESP which is verified having low skills at large basins in Korea (Seo et al., 2019). In this section, the overall performances of EDP were analyzed in detail, and then the necessity of the probabilistic approach for drought prediction was verified. Lastly, the effect of SMI information was evaluated.

(1) Prediction for short-term drought (SRI3)

The drought prediction for SRI3, which represents short-term drought, was carried out using EDP. The performance measures RMSE, RPSS, BS, and BSS are shown in Tables A2.1~A2.3 and Figures A2.1~A2.11 of Appendix A-2, where DP is the deterministic prediction and PP is the probabilistic prediction. In the heatmaps, the gray indicates that there was no drought case.

RMSE value exceeds 0.5 at all basins, and it becomes even larger in the irrigation period, which is an inevitable problem due to the large variance of streamflow in summer (irrigation period). This means that the determinism can make a false confidence problem because it results in wrong prediction of drought phase more than one phase on average. In this situation, the probabilistic approach is more appropriate. The necessity of the probabilistic approach was further discussed in the section '(3) Necessity of probabilistic approach'.

RPSS, the skill score for the prediction in multi-categorical, is above 0 in all basins, which means that EDP is better than the climatological prediction. On the other hand, according to the values BS, severe droughts (D2 and D3 phase in this study) are difficult to predict using EDP, especially even more for the irrigation period. It is better to have the assumption that severe droughts may occur with the same probability as the observed frequency because BSS are negative values when predicting above D1 phase.

(2) Prediction for long-term drought (SRI12)

The drought prediction for SRI12, which represents long-term drought, was performed using EDP. The performance measures such as RMSE, RPSS, BS, and BSS are shown in Tables A2.4~A2.6 and Figures A2.12~A2.22 of Appendix A-2. The overall trends are similar to the case of the short-term drought prediction, presenting the probabilistic approach is better than the deterministic one.

All the performance metrics of SRI12 are larger than those of SRI3, and it could be because SRI12 has a long persistence. As shown in Figure 4.5, the persistency of SRI12 looks longer than that of SRI3 because SRI12 considers streamflow accumulated for twelve months. It was reported that the persistency and predictability have a positive relationship (Shukla, 1983; Sun and Wang, 2013), and this relationship may lead the drought prediction performance of SRI12 being higher than that of SRI3.

RMSE is below 0.4 even for the irrigation period, so it is expected that the performance of the deterministic prediction will be good as well. RPSS and BS, which represent the performance in the probabilistic approach, have sufficiently good, but when predicting severe drought D2 and D3 at the Daecheong for severe drought they become inaccurate. This may be because of the problem of EDP itself, so it needs to be improved to solve the problem.

(3) Necessity of probabilistic approach

The superiority of the probabilistic approach can be verified through analyzing RPSS, BS, and BSS which can be indicators comparing two approaches. On average, RPSS of the probabilistic one is 100% higher than that of the deterministic for SRI3 (the short-term drought) prediction, and 7% higher for SRI12 (the long-term drought) prediction. In addition, BS of the probabilistic one is also 75% higher for the short-term drought prediction, and 24% for the long-term drought prediction, on average, and it is extremely higher when predicting D1 and D2 phases. When predicting the long-term drought, there is little difference between the probabilistic and deterministic ones because the predictability of SRI12 is sufficiently high even for the deterministic approach. To sum up these results, the probabilistic approach for drought prediction outperforms the deterministic one especially for the short-term drought prediction. Therefore, it can be said that use of the probabilistic approach is especially necessary when predicting the short-term drought.

Due to the anthropogenic activities and climate change, the uncertainty in hydrometeorological variables, including drought, has increased these days (Van

Loon et al., 2016). Under this circumstance, the importance and value of the probabilistic drought prediction for continues to rise, because the probabilistic approach is primarily effective as a tool to help decision-making to prepare for events that have large uncertainties and potentials to cause great losses (Krzysztofowicz, 2001; Palmer, 2017). Subsequently, Buizza (2008) proved that using the probabilistic prediction makes less potential loss than the deterministic prediction in real.

As analyzed above, the deterministic approach may yield errors more than one phase. This can lead to a false confidence, and finally make a catastrophic result like the Great Flood in the U.S. Therefore, it is more appropriate to predict disasters in the probabilistic approach unless the perfect prediction is possible. However, even if the perfect prediction is possible, the uncertainties coming from human activities and nature have to be considered so the probabilistic approach is required.

Of course, it is easy to open and share such drought probability information, but persuading users such as farmers and stakeholders about the importance of the probability information and educating how to recognize it remains a challenging issue. To overcome this, quantitative research is actively conducted in the social science field. Ramos et al. (2013) verified that the probabilistic information for hydrometeorological variables helps to make better decisions by experimental survey research. Furthermore, studies have been conducted to reflect opinions collected through surveys and discussions in hydrological forecasting and dam operation models in order to satisfy various needs of users (Fundel et al., 2018; Kim et al., 2019).

(4) Effectiveness of soil moisture information

Only the probabilistic prediction results of EDP and EDP+S are compared to analyze whether SMI is effective or not. In the above results, it was found that EDP+S makes better predictions than EDP, especially more effective for SRI3. However, there are some basins where the accuracy decreases when the SMI information is used to update EDP. In the four basins such as Seomjin, Hapcheon, Namgang, and Imha, RMSE, RPSS, and BS of the short-term drought prediction with SMI become worse by 2~3%. And, in the Seomjin and Imha basins, RMSE of SRI12 prediction was increased by 2%. The correlations between SMI and SRI of these basins are the smallest four as shown in Table 4.2, and this seems to affect the regression used as the likelihood function.

When focusing on N-RMSE of the residuals of the regression (Tables 4.6 and 4.7), the Seomjin, Hapcheon, and Imha basins have larger than other basins. This means that updating new information in EDP makes a negative effect if the reliability of the likelihood function is not sufficiently good. Maybe this is because the spatial resolution of the satellite data is not high enough so that the satellite cannot capture the value for small basins well. In other words, the quality of the data used for the Bayesian update may be one of the reasons for making the worse prediction. Although the resolution of soil moisture data is too coarse to represent the values of South Korea, reflecting soil moisture into EDP model is found to be effective across all basins. Therefore, in the future, it is expected that various remote sensing data such as satellite and radar that have been proven to be related to drought can be used for a drought study.

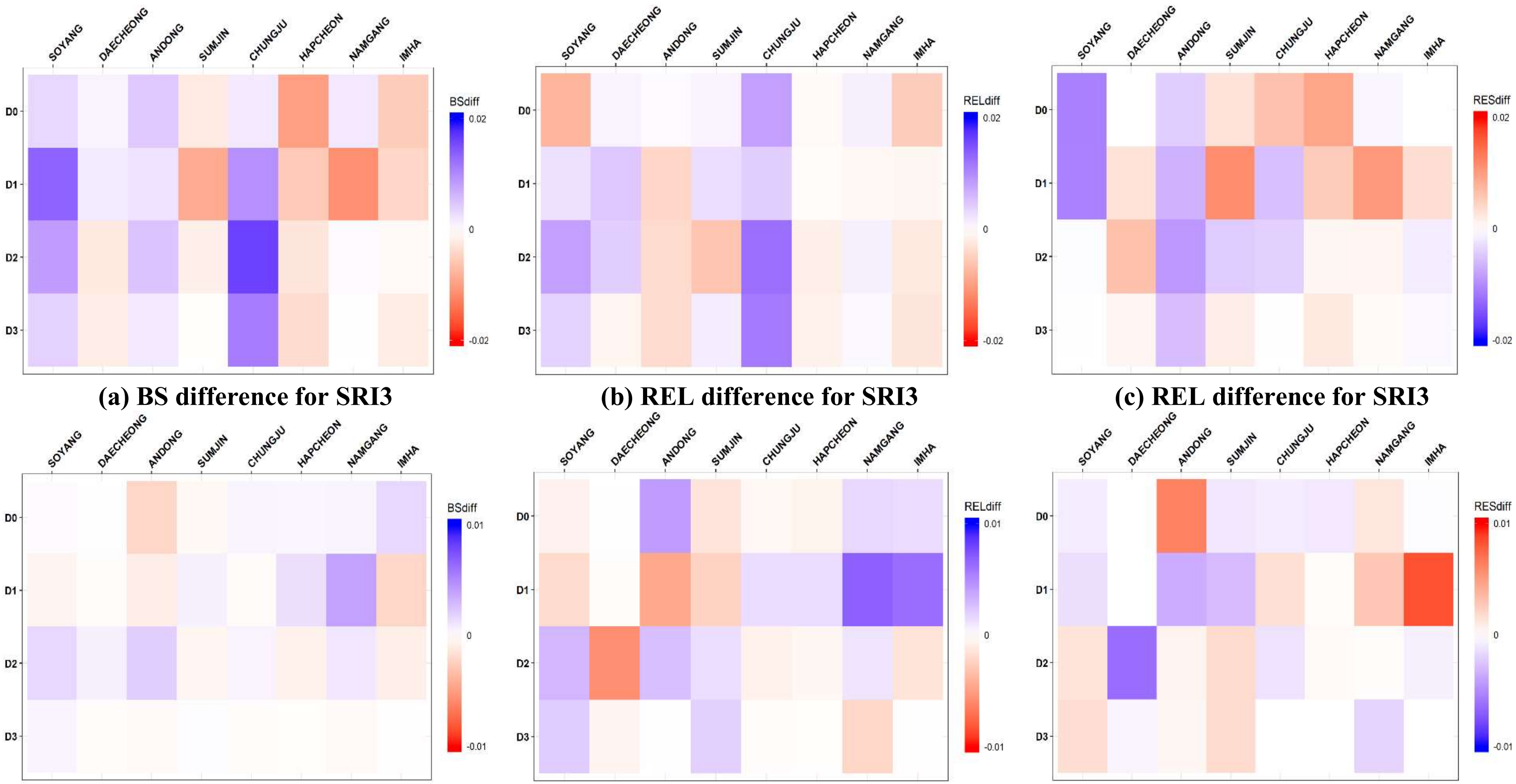
To analyze the effect of SMI on EDP in detail, we focus on the difference of the BS components between EDP and EDP+S. The differences between EDP and EDP+S at each basin are presented as a heatmap in Figure 4.6, where the blue means that EDP+S is better. In the heatmap, it is difficult to discern the variation of REL, but RES is decreased at the all basins except Soyang and Andong after reflecting SMI. This can be said that updating SMI makes EDP consider wider range of possible drought events.

Table 4.6 N-RMSE of likelihood function SMI~SRI3

Basin \ Month	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
Soyang	0.363	0.347	0.404	0.457	0.445	1.421	0.482	0.393	0.336	0.442	0.329	0.338	0.886
Daecheong	0.324	0.429	0.309	0.387	0.626	1.662	0.352	0.333	0.234	0.495	0.457	0.283	0.906
Andong	0.463	0.484	0.351	0.474	0.582	1.385	0.412	0.245	0.224	0.416	0.436	0.336	0.894
Seomjin	0.614	0.697	0.474	0.461	0.666	1.109	0.292	0.207	0.217	0.360	0.394	0.469	0.917
Chungju	0.318	0.319	0.400	0.488	0.535	1.379	0.504	0.389	0.297	0.415	0.313	0.355	0.879
Hapcheon	0.659	0.828	0.513	0.407	0.599	0.924	0.335	0.250	0.184	0.330	0.402	0.430	0.902
Namgang	0.607	0.718	0.474	0.376	0.561	0.879	0.305	0.237	0.186	0.327	0.371	0.408	0.838
Imha	0.519	0.506	0.403	0.468	0.569	1.401	0.434	0.281	0.230	0.425	0.450	0.403	0.937

Table 4.7 N-RMSE of likelihood function SMI~SRI12

Basin \ Month	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
Soyang	0.434	0.436	0.528	0.565	0.646	1.527	0.530	0.384	0.358	0.506	0.362	0.434	1.032
Daecheong	0.427	0.525	0.507	0.502	0.785	1.714	0.317	0.332	0.325	0.534	0.521	0.429	1.064
Andong	0.578	0.584	0.551	0.530	0.733	1.404	0.390	0.255	0.284	0.489	0.561	0.524	1.059
Seomjin	0.859	0.999	0.633	0.618	0.777	1.145	0.292	0.236	0.237	0.385	0.579	0.626	1.136
Chungju	0.441	0.440	0.536	0.557	0.646	1.453	0.548	0.403	0.308	0.454	0.333	0.442	1.009
Hapcheon	0.857	1.001	0.620	0.550	0.659	1.011	0.364	0.259	0.210	0.348	0.488	0.537	1.062
Namgang	0.850	0.995	0.607	0.536	0.659	0.943	0.329	0.249	0.224	0.349	0.484	0.540	1.041
Imha	0.585	0.595	0.561	0.543	0.727	1.475	0.429	0.278	0.279	0.489	0.559	0.508	1.081



(a) BS difference for SRI3 (b) REL difference for SRI3 (c) REL difference for SRI3

(d) BS difference for SRI12 (e) REL difference for SRI12 (f) REL difference for SRI12

Figure 4.6 Differences of BS components between EDP and EDP+S

4.4 EDP with Probabilistic Precipitation Forecast

4.4.1 Probabilistic Precipitation Forecast by APCC

APCC has been using MME to forecast the precipitation probability with the spatial resolution 2.5° across the world and East Asia since 2008. As shown in Figure 4.6, this probabilistic forecast produces the probability of three categories: below, normal, and above, up to six-month ahead. The skill of the APCC precipitation probability forecast in Korea was evaluated in detail by Sohn et al. (2012). More detailed explanation about MME can be found in Min et al. (2009).

4.4.2 Modeling Framework

Using the PDF ratio method, EDP is updated to EDP+A with the probabilistic precipitation forecast. The PDF ratio method was to reflect climate information to an ensemble distribution (Stedinger and Kim, 2010). It is a kind of technique of shifting a distribution by referring to new climate information (i.e., probabilistic forecast). The standard deviation and mean of the normal distribution updated using the PDF ratio are derived from Eqns (4.3)~(4.4).

$$\sigma_1 = \{\mathrm{x_a} - x_b\}/\{\Phi^{-1}(1 - p_a) - \Phi^{-1}(p_b)\} \tag{4.3}$$

$$\mu_1 = x_b - \sigma_1 \Phi^{-1}(p_b) \tag{4.4}$$

where x_a and x_b are terciles corresponding to 0.66 and 0.33 of the normal distribution before update, p_a and p_b are the probability corresponding to above and below respectively, and Φ^{-1} is the inverse function of the normal distribution.

EDP+AS was generated by updating SMI information on EDP+A, using the same likelihood function estimated in section 4.3. Figure 4.8 shows an example of applying four EDPs (EDP, EDP+S, EDP+A, and EDP+AS) to SRI3 at Soyang in July 2014, and the expected values of four EDPs from 2008 to 2017 at Soyang are shown in Figure 4.9.

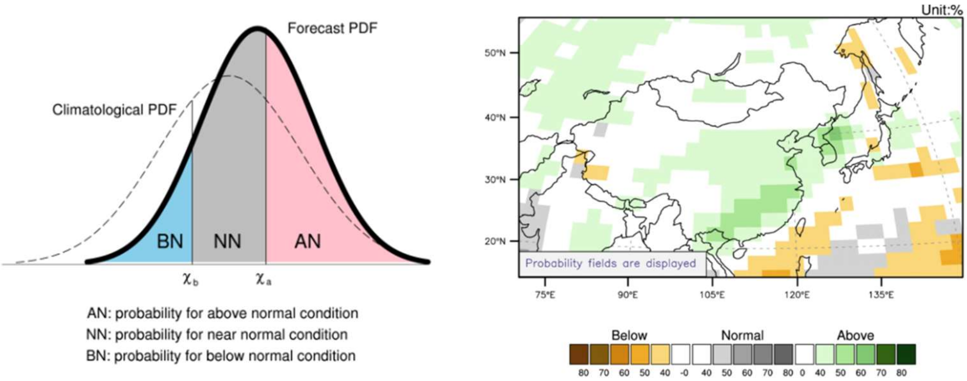


Figure 4.7 Probabilistic precipitation forecast by APCC (https://www.apcc21.org/)

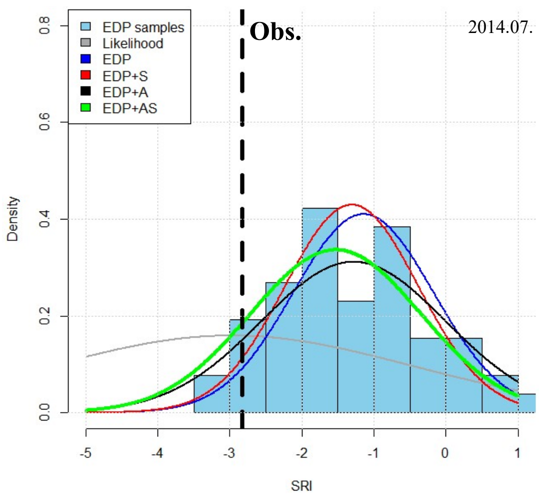
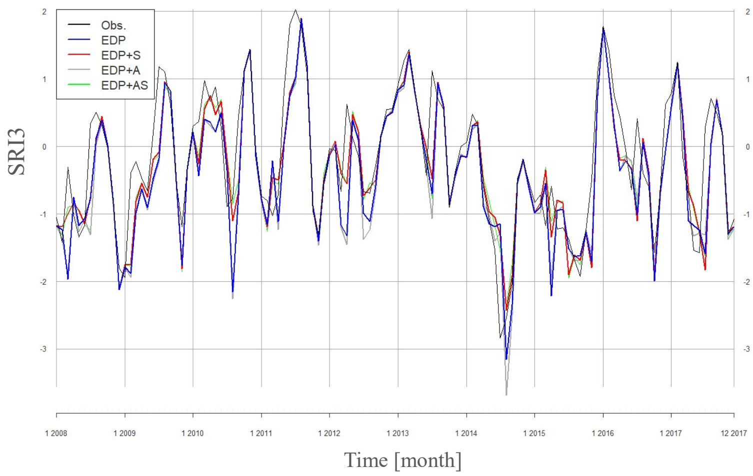
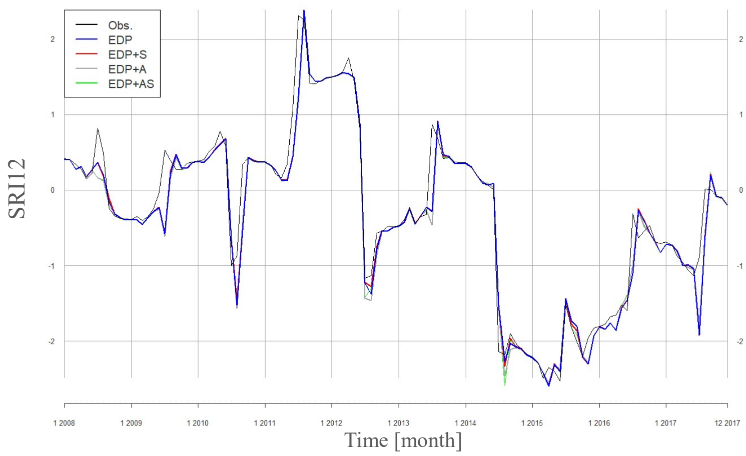


Figure 4.8 Example of EDP distributions at Soyang (EDP, EDP+S, EDP+A, EDP+AS)



(a) SRI3



(b) SRI12

Figure 4.9 Example of four EDPs at Soyang (EDP, EDP+S, EDP+A, EDP+AS)

4.4.3 Results and Discussion

Using four EDPs (EDP, EDP+S, EDP+A, and EDP+AS), the probabilistic drought prediction is performed at eight basins from 2008 to 2017. The performance was evaluated also for irrigation and non-irrigation periods, respectively. In order to easily understand the performance of EDP, reliability diagrams are derived by decomposing BS. All the performance metrics are shown in Figures A2.23~A2.30 and Tables A2.7~A2.12 of Appendix A-2. The overall performances of EDP and EDP+S are similar to those in section 4.3. And in the case of the predictions of SRI12 (the long-term drought), the variation between the four EDPs is not large. Therefore, this section focuses on the effect of the precipitation forecast on the prediction of SRI3 (the short-term drought).

(1) Effect of precipitation forecast

After updating the precipitation forecast, RMSE at eight basins decreases about 2% on average, and RPSS and BS do not change significantly. However, the effect and usefulness can be found when checking the metrics for the irrigation and non-irrigation periods separately. For the irrigation period, the performance metrics of EDP+A and EDP+AS are slightly lower than those of EDP. For the non-irrigation period, however, they become about 6% larger than EDP on average, and up to 19%.

It is assumed that the reason for these results is related to the performance of the precipitation forecast. Sohn et al. (2012) verified that the precipitation forecast by APCC has significant accuracy during winter, the non-irrigation period in Korea. To sum up, if the climate information like the precipitation forecast is informative enough, it is capable of predicting the drought more skillful.

(2) Reliability diagram

A reliability diagram is a graph where the conditional distribution of the observations, given the forecast probability, is plotted against the forecast probability and a perfect prediction is plotted along the 45-degree diagonal. BS and its components can be analyzed through the diagram.

The reliability diagrams shown in Figures A2.26 and A2.30 represent that all four EDPs make overestimation on drought occurrence because the observed frequency is lower than corresponding predicted probability. This overestimation affects RES significantly. As mentioned before, the uncertainty of the droughts gradually increases due to climate change and anthropogenic activities. Under this circumstance, even if a model can make perfect predictions for the past and present, there is no guarantee that it makes the perfect prediction for the future as well. Therefore, a prediction model that can consider a wide range of possible drought is required to prepare the drought in the time of climate change. The cost for the prevention may be wasted because of the overestimation of drought occurrences, so further research should be conducted to evaluate EDP in terms of socio-economy.

Chapter 5. Conclusion

5.1 Summary and Conclusions

This study has proposed a EDP system which predicts hydrological drought probabilistically using an ensemble method in order to demonstrate the necessity of introducing the probabilistic drought prediction to Korea. Among many types of drought, the hydrological drought is especially important because it is a linkage between drought as a natural phenomenon and its impact on human society. The natural hydrological drought can be measured by SRI that can represent both short-term and long-term hydrological. In this study, the hydrological drought has been categorized into four phases depending on the anomaly level. Then, a prediction for each phase is defined as a multi-categorical prediction, and a prediction for the occurrence or not above a certain phase is defined as a dichotomous prediction.

EDP is expressed as an ensemble of SRI which comes out by converting the ESP results. This study has applied EDP to eight dam basins in Korea to predict the short-term and the long-term drought in the deterministic and the probabilistic perspectives and then analyzed their performance metrics. Furthermore, to improve the prediction performance, EDP is updated with soil moisture information using the Bayes' theorem and climate information using the PDF ratio method. For the performance metrics, RMSE (a deterministic measure), and RPSS and BS, (probabilistic measures), are used. RMSE, combining bias and variability, is to evaluate errors of the mean of EDP compared to the observed SRI. RPSS is a skill score for multi-categorical predictions, and BS is the one for dichotomous predictions and can be decomposed into three components (reliability, resolution, and uncertainty) to make a further analysis. Besides, RPSS and BS can also evaluate the accuracy of the deterministic predictions at a probabilistic standpoint, so they are used to compare the probabilistic and the deterministic predictions.

To evaluate the skill of EDP for the short-term drought prediction, the result of SRI3 prediction was analyzed. RMSE exceeded 0.5 on average, and this means that it may yield errors more than one phase if taking the deterministic approach. Consequentially, by analyzing RPSS and BS, the deterministic prediction was inferior to the probabilistic one and even to the climatological prediction. The probabilistic prediction is always better than the climatological prediction in case of the multi-categorical prediction. However, there are some cases that the prediction skill of EDP for the drought phases over D2 is worse than the climatological prediction especially at large basins.

When predicting SRI12 which represents the long-term drought using EDP, the performance metrics are large in general, because it can be cause by the long persistency of SRI12. Thus, it is not easy to discern the differences between the deterministic and probabilistic predictions.

The drought information from SMI is used to update EDP via the Bayes' theorem for improving the prediction performance. The prior distribution is EDP distribution and the likelihood function is estimated as the regression between SMI and SRI. As a result, updating EDP is effective when the residual of the regression model is sufficiently small. In other words, the likelihood function has to be reliable to make EDP improved. Also, the reliability of the regression may depend on the quality of SMI. The SMI satellite data used in this study has a low spatial resolution to capture the soil moisture information of small basins. Consequentially, this makes low reliability of the likelihood function and thus it affects updating EDP negatively. Nevertheless, the Bayesian update with SMI yields 35% lager RPSS and 4% larger BS values than the original EDP. This can be said that the availability of SMI for drought prediction is proved.

Additionally, EDP is updated by reflecting the APCC climate information in EDP distribution via the PDF ratio method. Here, the climate information is the probabilistic precipitation forecast by MME. Updating the precipitation forecast results in the same or slightly lower when compared to EDP, but it improves the prediction performance by 6% for the non-irrigation period. The precipitation forecasts of the APCC have significant skills during winter season across East Asia including Korea and accordingly it could positively affect drought prediction for the non-irrigation period.

Summing up the above results, this study makes three conclusions as follows.

(1) The probabilistic drought prediction was 52% better than the deterministic on average in terms of prediction skills. When predicting the short-term drought, the probabilistic approach outperformed even more.

(2) Updating EDP using soil moisture information the via Bayes' theorem makes skill to be improved by 20% on average. Therefore, it can be said that the soil moisture information corrects EDP if the likelihood function is valid and accurate.

(3) Reflecting the precipitation forecast to EDP via the PDF ratio yielded 6% better performance only for the non-irrigation period. From this, it was found again that reflecting informative data can make better the drought prediction.

5.2 Future Study

This study demonstrated the advantages of the probabilistic drought prediction in terms of accuracy and skill. To state in more practical perspective on using the probabilistic approach, it should be analyzed with economic measures. This can be done if a potential economic value is derived by such as cost-loss analysis.

By comparing four EDPs (EDP, EDP+S, EDP+A, EDP+AS), it was confirmed that additional information about drought does not always make the skill better. This study concluded that the negative effects of the additional information are because of the reliability and quality of the data. In order to underpin this conclusion, it is necessary to update EDP using other additional information and compare the prediction performance. If other information about drought is able to be formed in the likelihood function (or conditional probability), EDP can be updated via the Bayes' theorem consecutively. The biggest advantage of the Bayesian update is that it can consider the new information continuously.

Also, this study demonstrated that the effectiveness of utilizing remote sensing data which is even coarse, so it is expected that various remote sensing data can also be used for drought study in the future.

The other limitation of EDP proposed in this study is that it only considered the natural hydrological drought using SRI. Representing the drought due to the natural hydrologic cycle is the advantage of the drought index but at the same time, it becomes a disadvantage because the drought index does not consider human activities such as water resource management. As described in the introduction, facilities such as reservoirs and dams have been constructed and operated to overcome drought. Accordingly, it can be said that the socio-economic drought introduced in chapter 2 has a closer relation to human society. If EDP is applied to the socio-economic drought prediction, other characteristics of EDP can be analyzed.

References

Bae, D. H., Ah, J.-B., Kim, H.-K., Kim, H.-A., Son, K. H., Jo, S., and Jung, U.-S. (2013). “Development and evaluation of real-time ensemble drought prediction system.” *Atmosphere, Korean Meteorological Society*, 29(3), pp. 113-121.

Barros, A. P., and Bowden, G. J. (2008). “Toward long-lead operational forecasts of drought: An experimental study in the Murray-Darling River basin.” *Journal of Hydrology*, 357, pp. 349-367.

Bradley, A. A., Habib, M., and Schwartz, S.S. (2015). “Climate index weighting of ensemble streamflow forecasts using a simple Bayesian approach.” *Water Resources Research*, 51(9), pp. 7382-7400.

Brankovic, C., Palmer, T. N., Molteni, S., and Cubasch, T. U. (1990). “Extended-range prediction with ECMWF models: Time-lagged ensemble forecasting.” *Quarterly Journal of the Royal Meteorological Society*, 116, pp. 867-912.

Buizza, R. (2008). “The value of probabilistic prediction.” *Atmospheric Science Letters*, 9, pp. 36-42.

Cai, J., Liu, Y., Lei, T., and Pereira, L.S. (2007). “Estimating reference evapotranspiration with the FAO Penman–Monteith equation using daily weather forecast messages.” *Agricultural and Forest Meteorology*, 145(1-2), pp. 22–35.

Choi, S., Kang, S., Lee, D., and Kim, J. H. (2018). “A study on water supply and demand prospects for water resources planning.” *Journal of Korean Society of Hazard Mitigation*, 18(7), pp. 589-596.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., and Valentini, R. (2005). “Europe-wide reduction in primary productivity caused by the heat and drought in 2003.” *Nature,* 437, pp. 529-533.

Cloke, H. L., and Hannah, D. M. (2011). “Large-Scale hydrology: advances in understanding processes, dynamics and models from beyond river basin to global scale.” *Hydrological Processes*, 25, pp. 991-995.

Cloke, H. L., and Pappenberger, F. (2009). “Ensemble flood forecasting: A review.” *Journal of Hydrology*, 375, pp. 613-626.

Coelho, C. A. S., Pezzulli, S., Balmaseda, M., Doblas-Reyes, F. J., and Stephenson, D. B. (2004). “Forecast calibration and combination: A simple Bayesian approach for ENSO.” *Journal of Climate*, 17(7), pp. 1504-1516.

Cooke, W. E. (1906). “Forecasts and verification in Western Australia.” *Monthly Weather Review*, 34(1), pp. 23-24.

Dai, A. (2011). “Drought under global warming: A review.” *Wiley Interdisciplinary Reviews: Climate Change*, 2, pp. 45-65.

Day, G. N. (1985). “Extended streamflow forecasting using NWSRFS.” *Journal of Water Resources Planning and Management*, 111(2), pp. 157-170.

Demargne, J., Wu, L., Regonda, S. K., Brown, J. D., Lee, H., He, M., Seo, D. J., Harman, R., Herr, H. D., Schaake, J., Zhu, Y., and Fresch, M. (2014). “The science of NOAA’s operational hydrologic ensemble forecast service.” *Bulletin of the American Meteorological Society*, 95(1), pp. 79-98.

Durdu, O. F. (2010). “Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey.” *Stochastic Environmental Research and Risk Assessment*, 24(8), pp. 1145-1162.

Edwards, D.C., and McKee, T.B. (1997). “Characteristics of 20th century drought in the United States at multiple scales.” *Atmospheric Science Paper No. 634*, May 1-30.

Emerton, R. E., Stephens, E. M., Pappenberger, F., Pagano, T. C., Weerts, A. H., Wood, A. W., Salamon, P., Brown, J. D., Hjerdt, N., Donnelly, C., Baugh, C. A., and Cloke, H. L. (2016). “Continental and global scale flood forecasting systems.” *Wiley Interdisciplinary Reviews: Water*, 3(3), pp. 391-418.

Fundel, V. J., Fleischhut, N., Herzog, S. M., Gober, M., and Hagedorn, R. (2019). “Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users.” *Quarterly Journal of the Royal Meteorological Society*, 145, pp. 210-231.

Grayson, M. (2013). “Agriculture and drought.” *Nature*, 501, S1.

Hamlet, A. F., and Lettenmaier, D. P. (1999). “Columbia River streamflow forecasting based on ENSO and PDO climate signals.” *Journal of Water Resources Planning and Management*, 125(6), pp.333-341.

Hao, Z., Hao, F., and Singh, V. P. (2016). “A general framework for multivariate multi-index drought prediction based on Multivariate Ensemble Streamflow Prediction(MESP).” *Journal of Hydrology*, 539, pp. 1-10.

Hao, Z., Hao, F., Singh, V. P., Sun, A.Y. and Xia, Y. (2016). “Probabilistic prediction of hydrologic drought using a conditional probability approach based on the meta-Gaussian model.” *Journal of Hydrology*, 542, pp. 772-780.

Hao, Z., Hong, Y., Xia, Y., Singh, V. P., Hao, F., and Cheng, H. (2016). “Probabilistic drought characterization in the categorical form using ordinal regression.” *Journal of Hydrology*, 535, pp. 331-339.

Hao, Z, Singh, V. P., and Xia, Y. (2018). "Seasonal drought prediction: Advances, challenges, and future prospects." *Reviews of Geophysics*, 56(1), pp. 108-141.

Harrigan, S., Prudhomme, C., Parry, S., Smith, K., and Tanguy, M. (2018). "Benchmarking ensemble streamflow prediction skill in the UK." *Hydrology and Earth System Sciences*, 22, pp. 2023-2039.

Jensen, M. E., Burman, R. D., and Allen, R.G. (1990). *Evapotranspiration and irrigation water requirements*. ASCE Manuals and Reports on Engineering Practice No. 70, VA, USA.

Kang, M. G., Lee, J. H., and Park, K. W. (2013). "Parameter regionalization of a Tank model for simulating runoffs from ungauged watersheds." *Journal of Korea Water Resources Association*, 46(5), pp. 519-530.

Kang, T.-H., Hong, I.-P., and Y.-O. Kim. (2010) "Comparison of pre- and post-processor for ensemble streamflow prediction." *Atmospheric Science Letters*, 11, pp. 153-159.

Kim, G. J., Seo, S. B., and Kim, Y.-O. (2019). "Elicitation of drought alternatives based on Water Policy Concil and the role of Shared Vision Model." *Journal of Korea Water Resources Association*, 52(6), pp. 429-440.

Kim, Y.-O., Jeong, D. I., and Kim, H. S. (2001). "Improving water supply outlook in Korea with ensemble streamflow prediction." *Water International*, 26(4), pp. 563-568.

Korea National Drought Information-analysis Center (KNDIC) (2018). *2013~2018 Sustained Drought Analysis and Assessment Report*. Korea.

Krzysztofowicz, R. (2001). "The case for probabilistic forecasting in hydrology." *Journal of Hydrology*, 249(1-4), pp. 2-9.

Lee, P. M. (1997). *Bayesian Statistics: An Introduction,* 2nd edition, Arnold, London, UK, pp. 344.

Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., and Di, Z. (2017). "A review on statistical postprocessing methods for hydrometeorological ensemble forecasting." *Wiley Interdisciplinary Reviews: Water*, 4(6), e1246.

Lloyd-Hughes, B., and Saunders, M. A. (2002). "A drought climatology for Europe." *International Journal of Climatology*, 22, pp. 1571-1592.

Luo, L., and Wood, E. F. (2007). "Monitoring and predicting the 2007 US drought." *Geophysical Research Letters*, 34, L22702.

Luo, L., Wood, E. F., and Pan, M. (2007). "Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions." *Journal of Geophysical Research,* 112, D10102.

Ma, F., Yuan, X., and Ye, A. (2015). "Seasonal drought predictability and forecast skill over China." *Journal of Geophysical Research: Atmospheres*, 120(6), pp. 8264-8275.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., and Reichle, R. (2012). "Soil moisture, snow, and seasonal streamflow forecasts in the United States." *Journal of Hydrometeorology*, 13, pp. 189-203.

McCabe, G. J., and Markstrom, S. L. (2007). *A Monthly Water-Balance Model Driven by a Graphical User Interface (No. 2007-1088)*. US Geological Survey, Reston, VA, USA.

McKee, T. B., Doesken, N. J., Kleist, J. (1993). "The relationship of drought frequency and duration to time scales." *The 8th conference on Applied Climatology*, Anaheim, CA, USA, pp. 179-187.

Mekonnen, M. M., and Hoekstra, A. Y. (2016). "Four billion people facing severe water scarcity." *Science advances*, 2(2), e1500323.

Mendoza, P. A., Wood, A. W., Clark, E., Rothwell, E., Clark, M. P., Nijssen, B., Brekke, L. D., and Arnold, J. R. (2017). "An intercomparison of approaches for improving operational seasonal streamflow forecasts." *Hydrology and Earth System Sciences*, 21(7), pp. 3915-3935.

Min, Y.-M., Kryjov, V. N., and Park, K.-K. (2009). "Probabilistic multimodel ensemble approach to seasonal prediction." *Weather and Forecasting*, 24, pp. 812-828.

Ministry of Land, Infrastructure, and Transport (MOLIT) (2016). *Water Vision 2020 3rd edition*. MOLIT, Korea.

Mishra, A., Desai, V. and Singh, V. (2007). "Drought forecasting using a hybrid stochastic and neural network model." *Journal of Hydrologic Engineering*, 12(6), pp. 626-638.

Mishra, A., and Singh, V. (2011). "Drought modelling-A review." *Journal of Hydrology*, 403, pp. 157-175.

Mo, K. C., Xu, L., and Lattenmaier, D. P. (2019) "Operational probabilistic drought prediction to improve the seasonal drought outlook." *Climate Prediction S&T Digest*, 101.

Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T, and Vitart, F. (2011). *The new ECMWF seasonal forecast system (System 4) (Vol. 49)*. Reading, U. K: European Centre for Medium-Range Weather Forecasts.

Mosely, L. M. (2015). "Drought impacts on the water quality of freshwater systems: A review." *Earth Science Reviews,* 140, pp. 203-214.

Murphy, J., and Palmer, T. (1986). "Experimental monthly long-range forecasts for the United Kingdom. 2. A real-time long-range forecast by an ensemble of numerical integrations." *Meteorological Magazine*, 115(1372), pp. 337-349.

Murphy, A. H., and Winkler, R. L. (1987). "A general framework for forecast verification." *Monthly Weather Review*, 115(7), pp. 1330-1338.

National Oceanic and Atmospheric Administration (NOAA). (1994). *The Natural Disaster Report for the Great Flood of 1993*. U.S. Department of Commerce, Silver Spring, MD, USA.

Palmer, T. N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P., and Tribbia, J. (1993). "Ensemble prediction." *Proceedings of the ECMWF Seminar on validation of models over Europe*, 1, pp. 21-66.

Palmer, T. N. (2017). "The primacy of doubt: Evolution of numerical weather prediction from determinism to probability." *Journal of Advances in Modeling Earth Systems*, 9(2), pp. 730-734.

Qu, B., Zhang, X., Pappenberger, F., Zhang, T., and Fang, Y. (2017). "Multi-model grand ensemble hydrologic forecasting in the Fu river basin using Bayesian model averaging." *Water*, 9(2), pp. 74.

Ramos, M. H., Van Andel, S. J., and Pappenberger, F. (2013). "Do probabilistic forecasts lead to better decisions?" *Hydrology and Earth System Sciences Discussions*, 17, pp. 2219-2232.

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van Den Dool, H. M., Pan, H.-L., Moorthi, S., Behringer, D., Stokes, D., Pena, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., and Xie, P. (2014). "The NCEP climate forecast system version 2." *Journal of Climate*, 27(6), pp. 2185-2208.

Seager, R., Naik, N., and Vecchi, G. A. (2010). "Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming." *Journal of Climate*, 23(17), pp. 4651-4668.

Seo, S. B., and Kim, Y.-O. (2018). "Impact of spatial aggregation level of climate indicators on a national-level selection for representative climate change scenarios." *Sustainability*, 10(7), 2409.

Seo, S. B., Kim, Y.-O., Kang, S.-U., and Chun, G. I. (2019). "Improvement in long-range streamflow forecasting accuracy using the Bayes' theorem." *Hydrology Research*, 50(2), pp. 616-632.

Shukla, J. (1983). "Comments on 'natural variability and predictability.'" *Monthly Weather Review*, 111(3), pp. 581-585.

Shukla, S., McNally, A., Husak, G., and Funk, C. (2014). "A seasonal agricultural drought forecast system for food-insecure regions of East Africa." *Hydrology and Earth System Sciences*, 18(10), pp. 3907-3921.

Shukla, S., Sheffield, J., Wood, E. F., and Lettenmaier, D. P. (2013). “On the sources of global land surface hydrologic predictability.” *Hydrology and Earth System Sciences*, 17(7), pp. 2781-2796.

Shukla, S., and Wood, A. W. (2008). “Use of a standardized runoff index for characterizing hydrologic drought.” *Geophysical Research Letters*, 35(2), L02405.

Sohn, S. J., Min, Y. M., Lee, J. Y., Tam, C. Y., Kang, I. S., Wang, B., Ahn, J. B., and Yamagata, T. (2012). “Assessment of the long-lead probabilistic prediction for the Asian summer monsoon precipitation (1983-2011) based on the APCC multimodel system and a statistical model.” *Journal of Geophysical Research: Atmospheres*, 117(D4).

Sohn, S. J., Tam, C. Y., and Ahn, J. B., (2013). “Development of a multimodel-based seasonal prediction system for extreme droughts and floods: A case study for South Korea.” *International Journal of Climatology*, 33(4), pp. 793-805.

Son, K. H., and Bae, D. H. (2015). “Applicability assessment of hydrological drought outlook using ESP method.” *Journal of Korea Water Resources Association*, 48(7), pp. 581-593.

Son, K. H., Bae, D. H., and Chung, J. S. (2011). “Drought analysis and assessment by using land surface model on South Korea.” *Journal of Korea Water Resources Association*, 44(8), pp. 667-681.

Sridhar, V., Hubbard, K.G., You, J., and Hunt, E.D. (2007). “Development of the soil moisture index to quantify agricultural drought and its ‘User Friendliness’ in severity-area-duration assessment.” *Journal of Hydrometeorology*, 9(4), pp. 660–676.

Stedinger, J. R., and Kim, Y.-O. (2010). “Probabilities for ensemble forecasts reflecting climate information.” *Journal of Hydrology*, 391(1-2), pp. 9-23.

Stockdale, T. N., Alves, O., Boer, G., Deque, M., Ding, Y., Kumar, A., K Yun, W. T. (2010). “Understanding and predicting seasonal-to-interannual climate variability-The producer perspective.” *Procedia Environmental Sciences*, 1, pp. 55-80.

Sugawara, M. (1995). *Tank model, Computer models of watershed hydrology*. Water Resources Publication, Highlands Ranch, CO, USA, pp. 165-214.

Sun, B., and Wang, H. (2013). “Larger variability, better predictability?” *International Journal of Climatology*, 33(10), pp. 2341-2351.

Tallaksen, Lena M., Henrik Madsen, and Bente Clausen. (1997). “On the definition and modelling of streamflow drought duration and deficit volume.” *Hydrological Sciences Journal*, 42(1), pp. 15-33.

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J. (2014). “Global warming and changes in drought.” *Nature Climate Change*, 4(1), pp. 17-22.

Van Loon, A. F. (2015). “Hydrological drought explained.” *Wiley Interdisciplinary Reviews: Water*, 2(4), pp. 359-392.

Van Loon, A. F., Stahl, K., Baldassarre, G. D., Clark, J., Rangecroft, S., Wanders, N., Gleeson, T., Van Dijk, A. I. J. M., Tallaksen, L. M., Hannaford, J., Uijlenhoet, R., Teuling, A. J., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., and Van Lanen, H. A. J. (2016). “Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches.” *Hydrology and Earth System Sciences*, 20(9), pp. 3631-3650.

Van Loon, A. F., and Van Lanen, H. A. J. (2012). “A process-based typology of hydrological drought.” *Hydrology and Earth System Sciences*, 16, pp. 1915-1946.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). “A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index.” *Journal of climate*, 23(7), pp. 1696-1718.

Wilhite, D. A., and Glantz, M. H. (1985). “Understanding the drought phenomenon: the role of definitions.” *Water International*, 10(3), pp. 111-120.

Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences*. Academic Press, Waltham, MA, USA.

Wood, A. W., and Lattenmaier, D. P. (2008). “An ensemble approach for attribution of hydrologic prediction uncertainty.” *Geophysical Research Letters*, 35(14).

Wood, E. F., Schubert, S. D., Wood, A. W., Peters-Lidard, C. D., Mo, K. C., Mariotti, A., and Pulwarty, R. S. (2015). “Prospects for advancing drought understanding, monitoring and prediction.” *Journal of Hydrometeorology*, 16(4), pp. 1636-1657.

Yang, D., Yan, X. Q., Xie, Q., Zhang, Y., Ren, X., and Tang, Y. (2016). “Probabilistic versus deterministic skill in predicting the western North Pacific-East Asian summer monsoon variability with multimodel ensembles.” *Journal of Geophysical Research: Atmospheres*, 121(3), pp. 1079-1103.

Yao, H., and Georgakakos, A. (2001). “Assessment of Folsom Lake response to historical and potential future climate scenarios: 2. Reservoir management.” *Journal of Hydrology*, 249(1), pp. 176-196.

Yoon, J. H., Mo, K., and Wood, E. F. (2012). "Dynamic-model-based seasonal prediction of meteorological drought over the contiguous Unites States." *Journal of Hydrometeorology*, 13(2), pp. 463-482.

Yuan, X., Ma, F., Wang, L., Zheng, Z., Ma, Z., Ye, A., and Peng, S. (2016). "An experimental seasonal hydrological forecasting system over the Yellow River basin—Part I: Understanding the role of initial hydrological conditions." *Hydrology and Earth System Sciences*, 20(6), pp. 2437-2451.

Yuan, X., Wood, E. F., and Ma, Z. (2015). "A review on climate-model-based seasonal hydrologic forecasting: Physical understanding and system development." *Wiley Interdisciplinary Reviews: Water*, 2(5), pp. 523-536.

Zhao, L., Duan, Q., Schake, J., Ye, A., and Xia, J. (2011). "A hydrologic post-processor for ensemble streamflow predictions." *Advances in Geosciences*, 29, pp. 51-59.

Zink, M., Samaniego, L., Kumar, R., Thober, S., Mai, J., Schäfer, D., and Marx, A. (2016). "The German drought monitor." *Environmental Research Letters*, 11(7), 074002.

Appendix

A-1. Ensemble Streamflow Prediction Results

Table A11 Accuracy of ensemble streamflow prediction from 2001 to 2017

Basin	Basin number	NSE	N-RMSE
Soyang	1	0.394	1.292
Daecheong	2	-0.027	1.682
Andong	3	0.396	1.225
Seomjin	4	0.447	1.189
Chungju	5	0.110	1.579
Hapcheon	6	0.390	1.258
Namgang	7	0.401	1.242
Imha	8	0.121	1.686

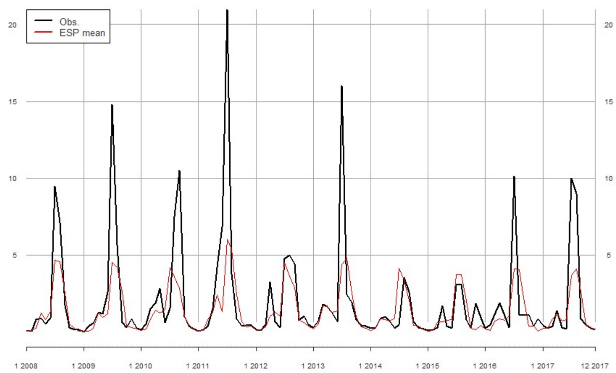


Figure A1.1 Ensemble streamflow prediction result at Soyang

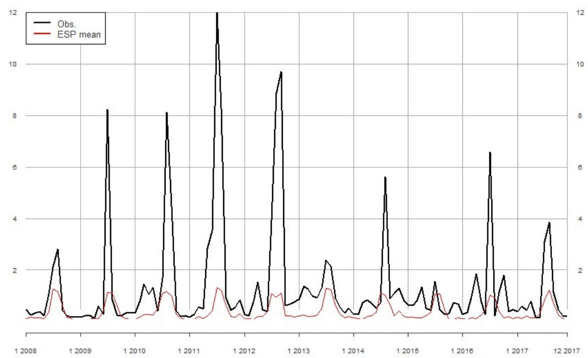


Figure A1.2 Ensemble streamflow prediction result at Daecheong

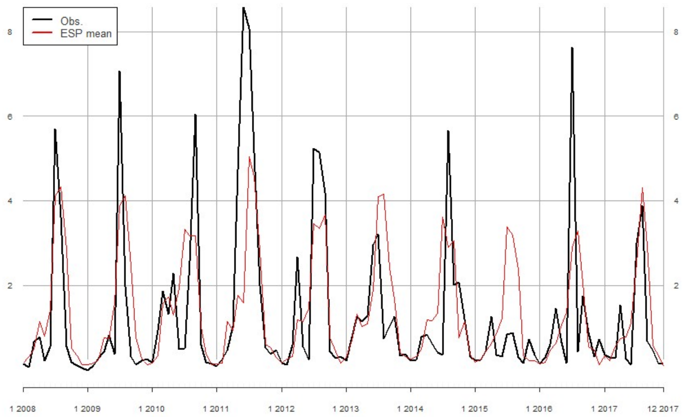


Figure A1.3 Ensemble streamflow prediction result at Andong

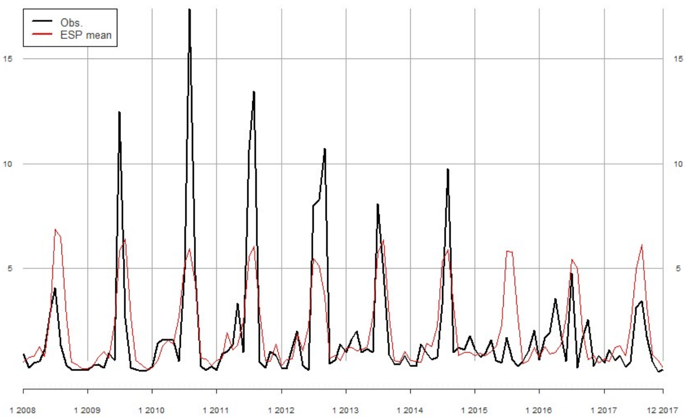


Figure A1.4 Ensemble streamflow prediction result at Seomjin

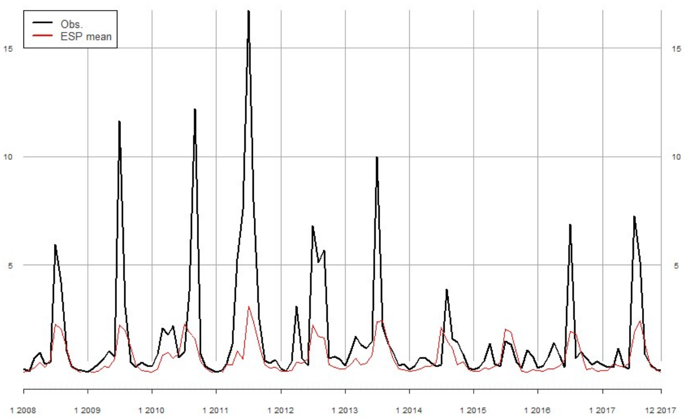


Figure A1.5 Ensemble streamflow prediction result at Chungju

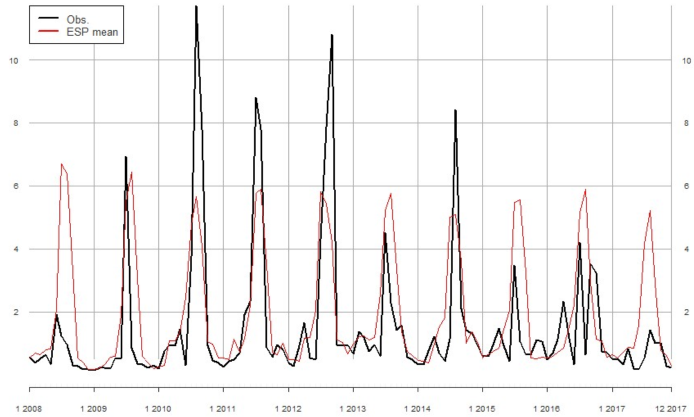


Figure A1.6 Ensemble streamflow prediction result at Hapcheon

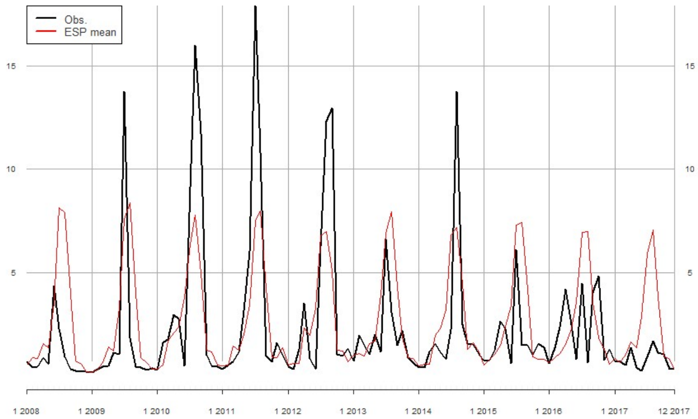


Figure A1.7 Ensemble streamflow prediction result at Namgang

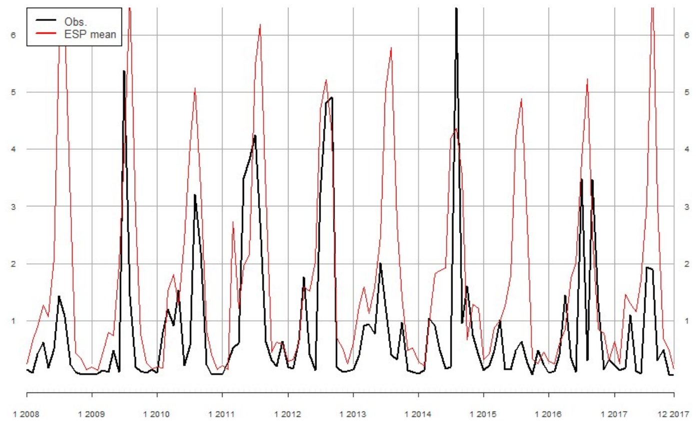


Figure A1.8 Ensemble streamflow prediction result at Imha

A-2. Ensemble Drought Prediction results

Table A2.1 RMSE of two EDPs for SRI3

Case	Soyang	Daecheong	Andong	Seomjin	Chungju	Hapcheon	Namgang	Imha
EDP	0.609	0.676	0.600	0.524	0.791	0.501	0.509	0.655
EDP+S	0.551	0.658	0.582	0.532	0.739	0.519	0.518	0.642

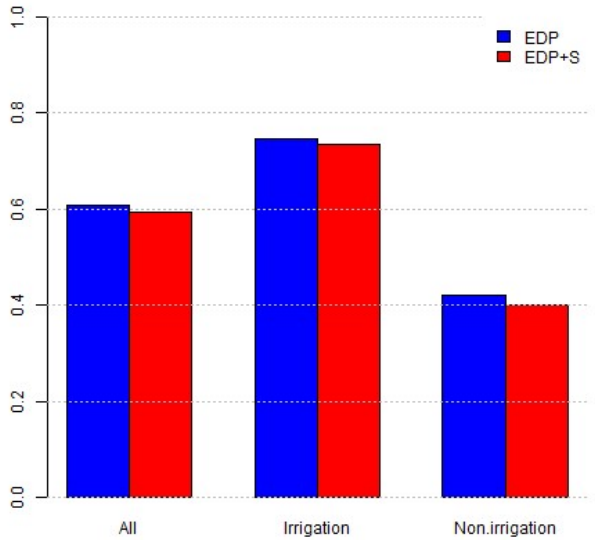
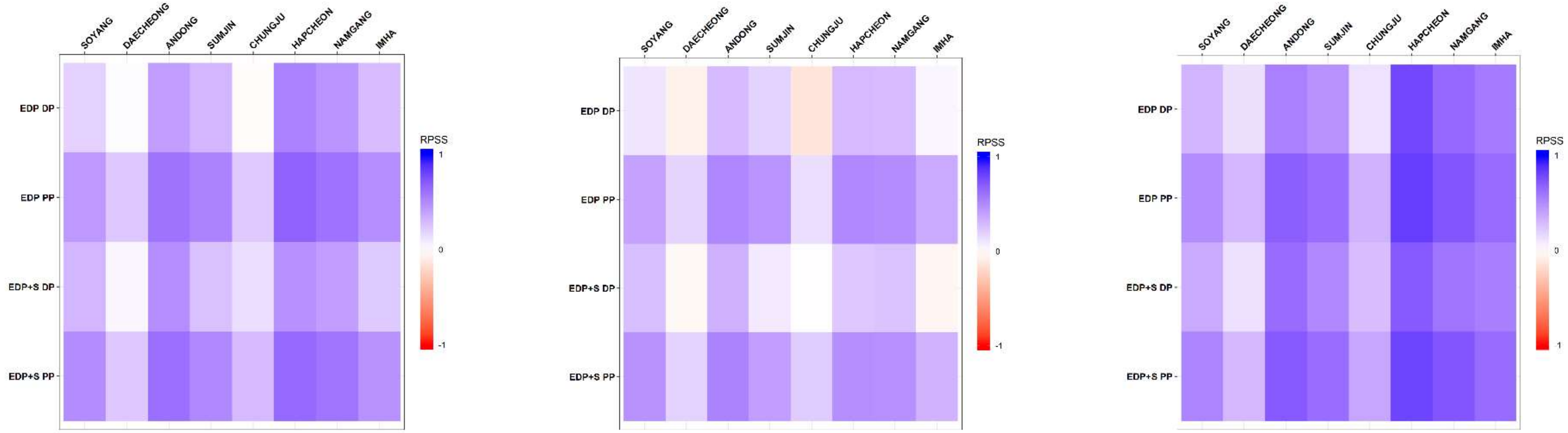


Figure A2.1 Averaged RMSE of two EDPs across eight basins for SRI3

Table A2.2 RPSS of two EDPs for SRI3

Basin \ Case	All				Irrigation				Non-irrigation			
	EDP		EDP+S		EDP		EDP+S		EDP		EDP+S	
	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP
Soyang	0.197	0.438	0.313	0.493	0.106	0.397	0.274	0.467	0.318	0.493	0.366	0.529
Daecheong	0.012	0.233	0.033	0.235	-0.073	0.181	-0.033	0.189	0.132	0.308	0.127	0.300
Andong	0.415	0.603	0.487	0.624	0.285	0.517	0.336	0.536	0.546	0.689	0.638	0.712
Seomjin	0.312	0.530	0.268	0.509	0.193	0.455	0.090	0.417	0.472	0.631	0.508	0.632
Chungju	-0.016	0.229	0.134	0.296	-0.139	0.142	0.007	0.219	0.122	0.326	0.277	0.382
Hapcheon	0.542	0.674	0.474	0.643	0.297	0.515	0.236	0.482	0.787	0.835	0.714	0.804
Namgang	0.459	0.606	0.417	0.595	0.283	0.490	0.254	0.473	0.652	0.733	0.597	0.729
Imha	0.284	0.488	0.227	0.470	0.041	0.359	-0.051	0.328	0.571	0.640	0.554	0.637

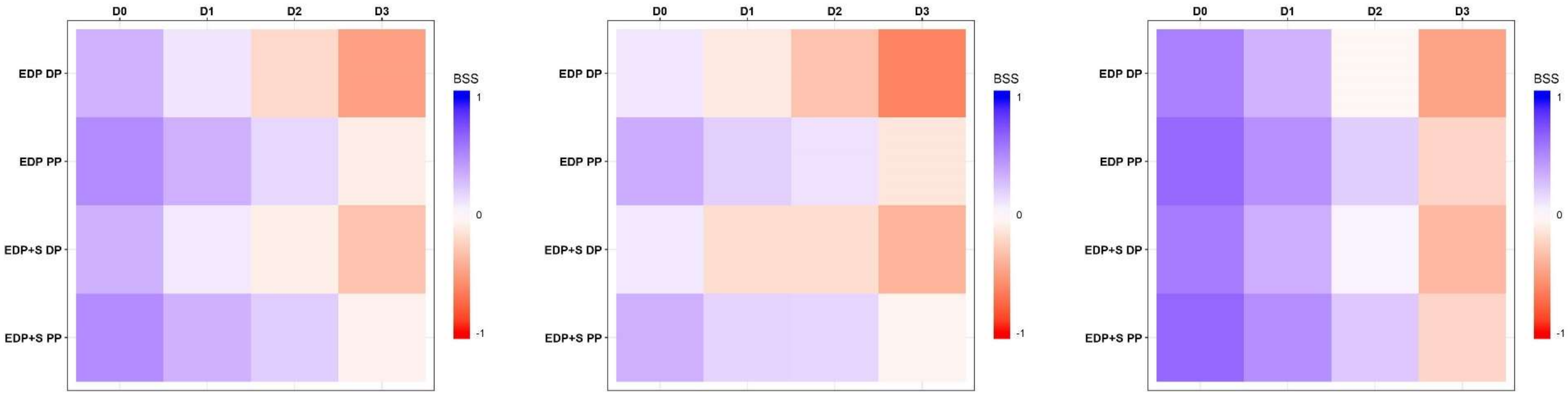


(a) All (b) Irrigation (c) Non irrigation

Figure A2.2 RPSS of two EDPs for SRI3

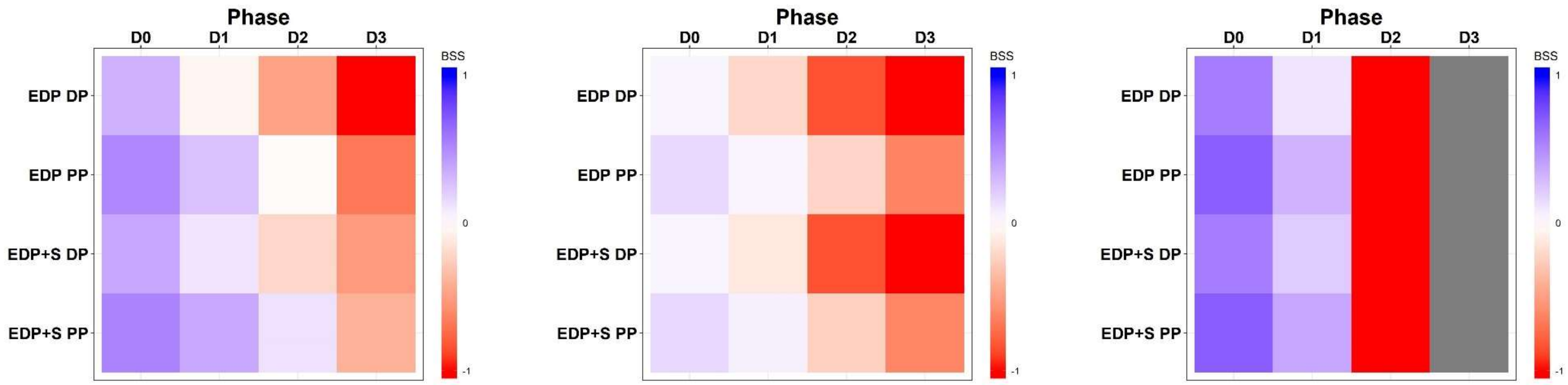
Table A2.3 Averaged BS of two EDPs across eight basins for SRI3

Phase \ Case	All				Irrigation				Non-irrigation			
	EDP		EDP+S		EDP		EDP+S		EDP		EDP+S	
	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP
D0	0.1642	0.0345	0.1631	0.0344	0.2194	0.0617	0.2209	0.0616	0.1091	0.0070	0.1054	0.0071
D1	0.1317	0.0307	0.1319	0.0304	0.1642	0.0534	0.1694	0.0530	0.0993	0.0078	0.0944	0.0078
D2	0.0784	0.0175	0.0711	0.0170	0.1005	0.0335	0.0892	0.0324	0.0564	0.0013	0.0530	0.0014
D3	0.0331	0.0095	0.0286	0.0095	0.0490	0.0152	0.0411	0.0152	0.0172	0.0037	0.0162	0.0038



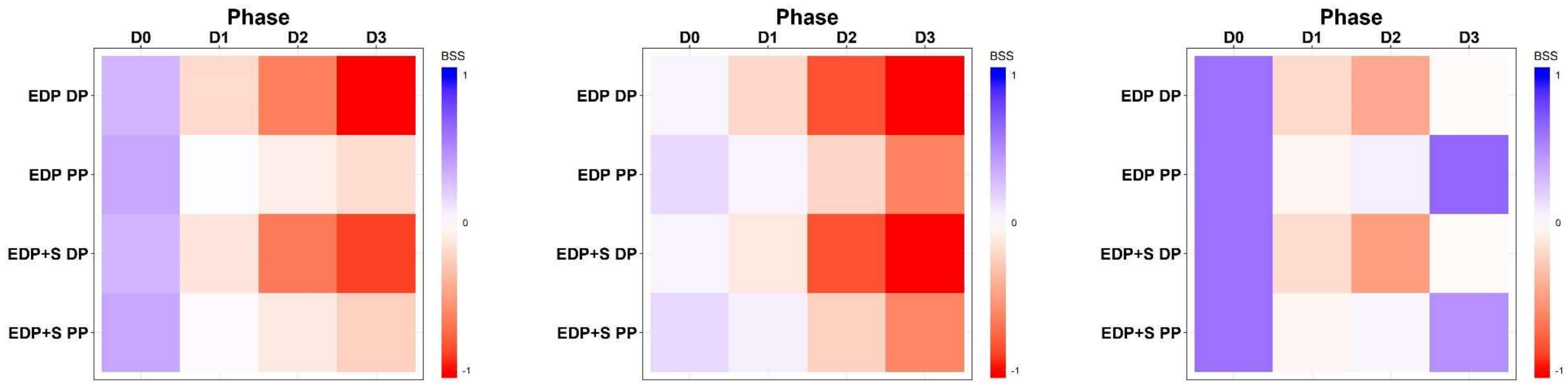
(a) All (b) Irrigation (c) Non irrigation

Figure A2.3 Averaged BSS of two EDPs across eight basins for SRI3



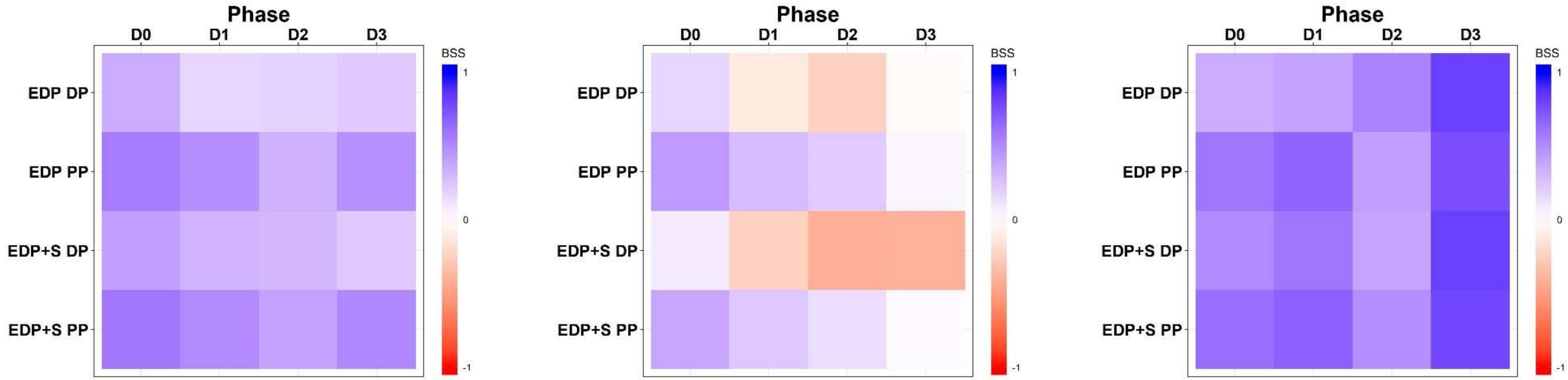
(a) All (b) Irrigation (c) Non irrigation

Figure A2.4 BSS at Soyang for SRI3



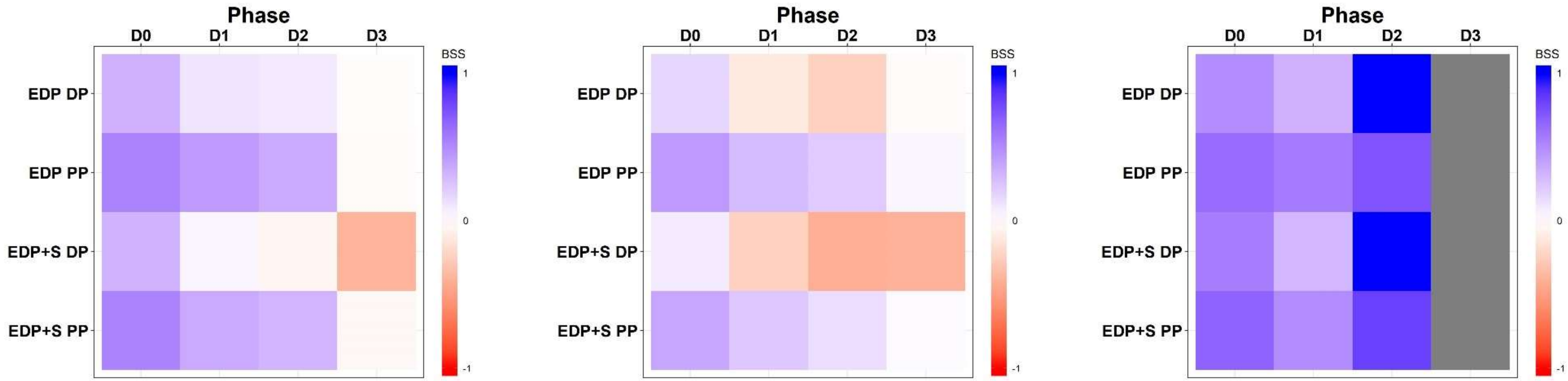
(a) All (b) Irrigation (c) Non irrigation

Figure A2.5 BSS at Daecheong for SRI3



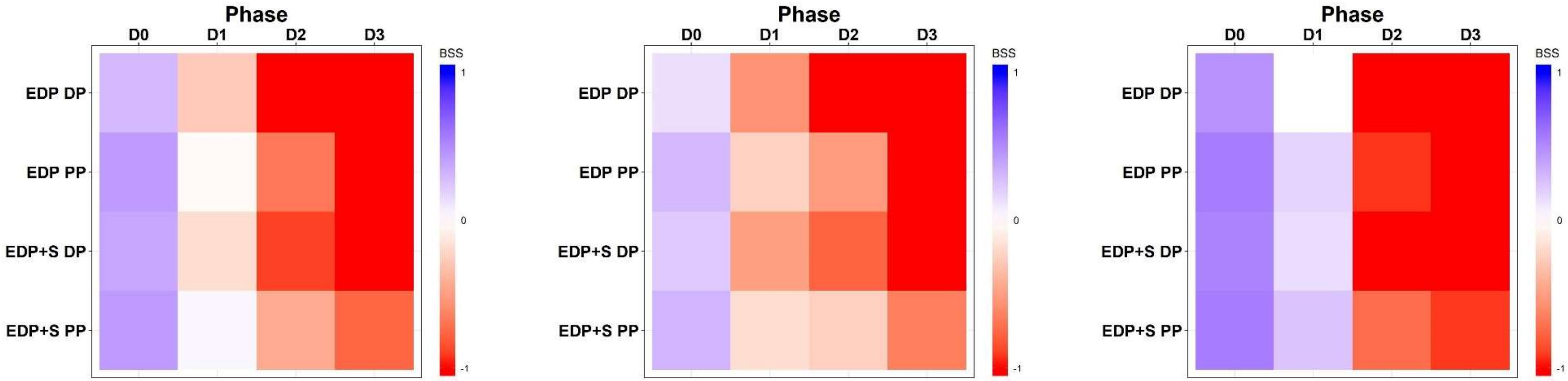
(a) All **(b) Irrigation** **(c) Non irrigation**

Figure A2.6 BSS at Andong for SRI3



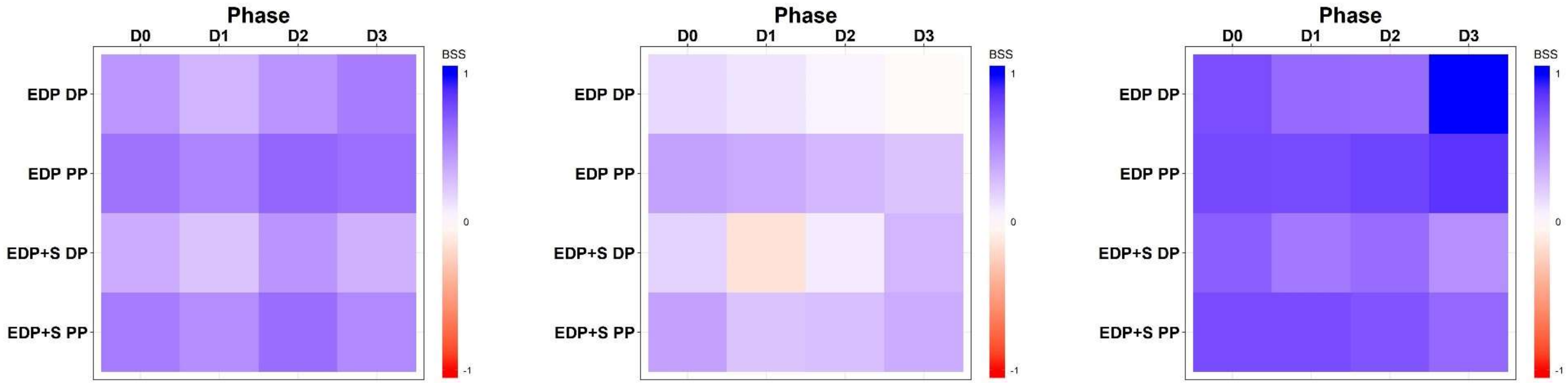
(a) All **(b) Irrigation** **(c) Non irrigation**

Figure A2.7 BSS at Seomjin for SRI3



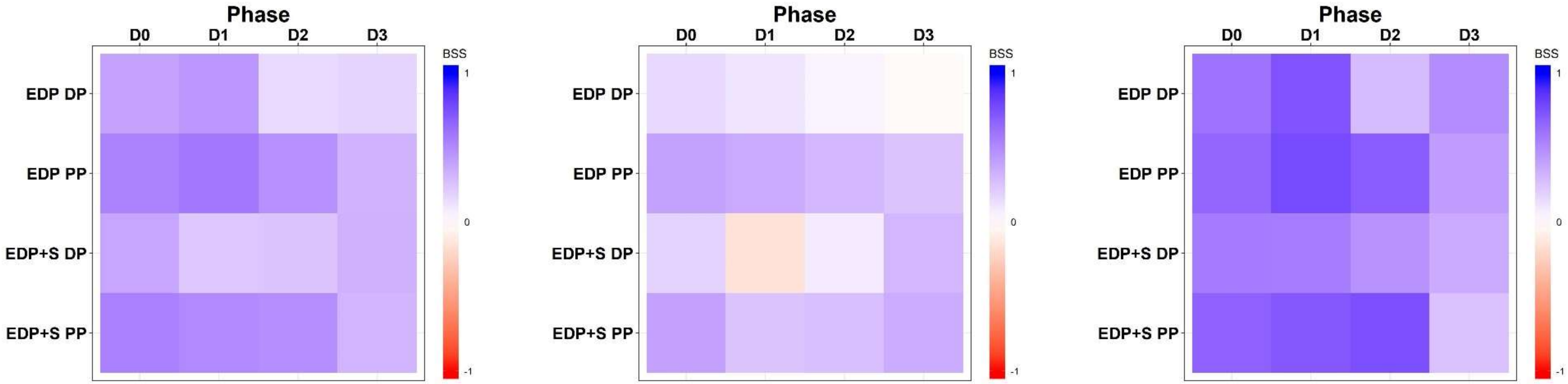
(a) All (b) Irrigation (c) Non irrigation

Figure A2.8 BSS at Chungju for SRI3



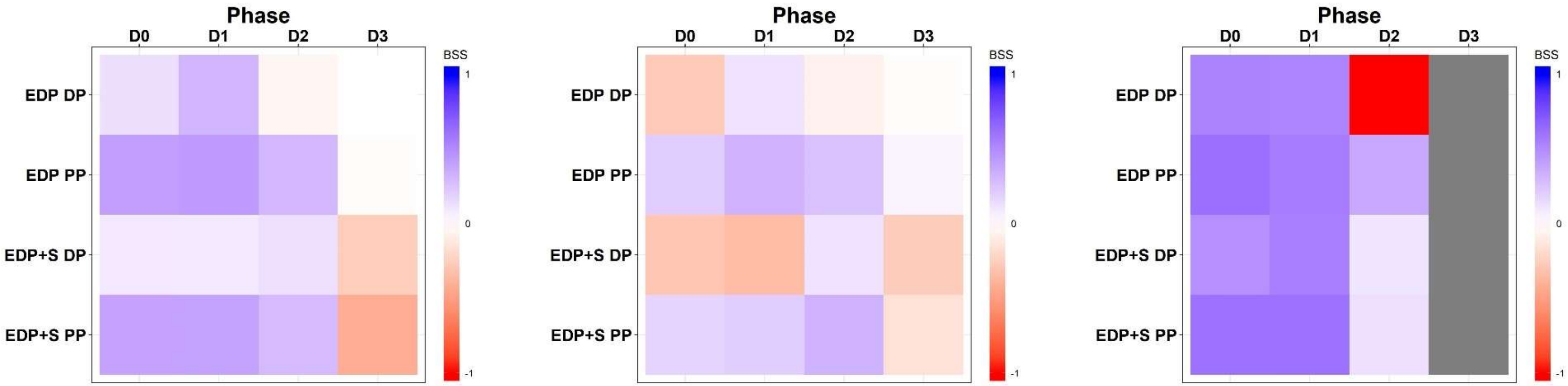
(a) All (b) Irrigation (c) Non irrigation

Figure A2.9 BSS at Hapcheon for SRI3



(a) All (b) Irrigation (c) Non irrigation

Figure A2.10 BSS at Namgang for SRI3



(a) All (b) Irrigation (c) Non irrigation

Figure A2.11 BSS at Imha for SRI3

Table A2.4 RMSE of two EDPs for SRI12

Case	Soyang	Daecheong	Andong	Seomjin	Chungju	Hapcheon	Namgang	Imha
EDP	0.293	0.288	0.279	0.270	0.311	0.245	0.276	0.259
EDP+S	0.286	0.286	0.281	0.267	0.309	0.247	0.268	0.262

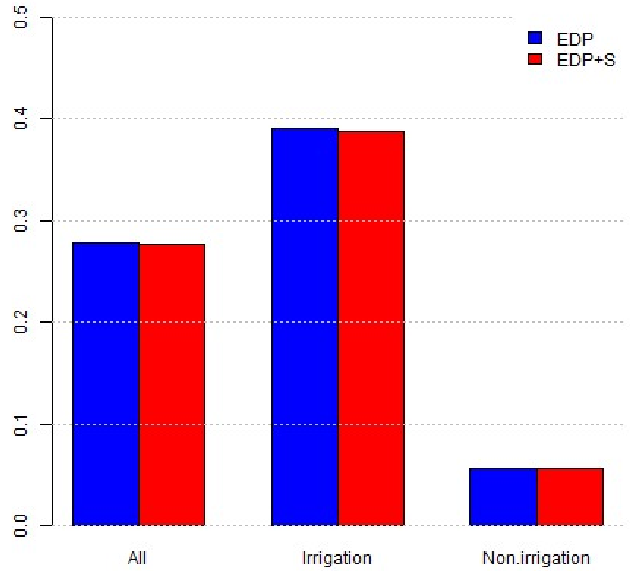
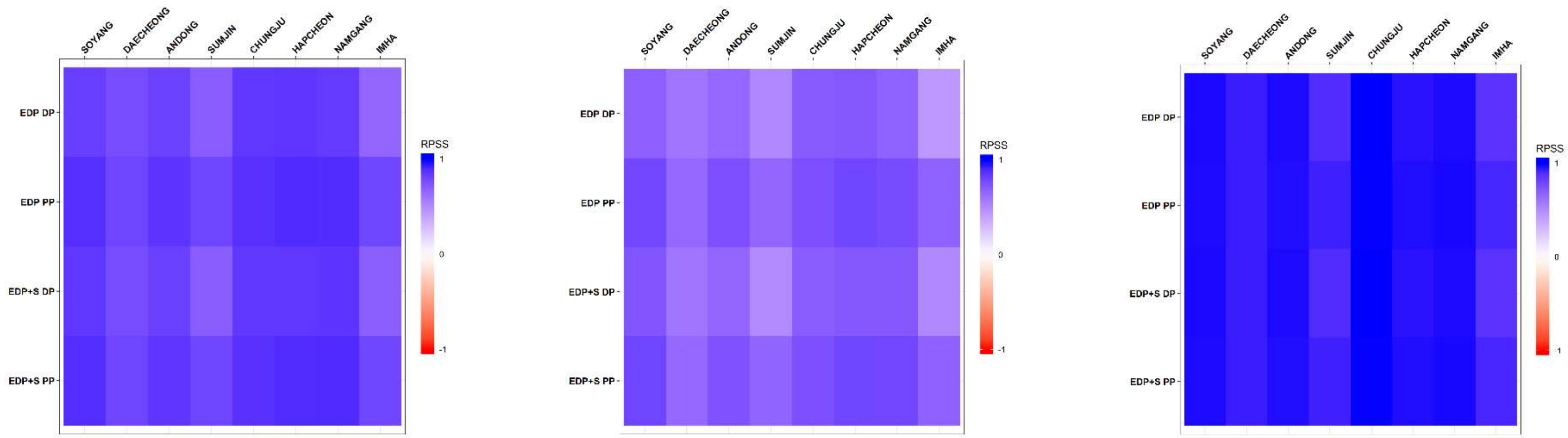


Figure A2.12 Averaged RMSE of two EDPs across eight basins for SRI12

Table A2.5 RPSS of two EDPs for SRI12

Case / Basin	All				Irrigation				Non-irrigation			
	EDP		EDP+S		EDP		EDP+S		EDP		EDP+S	
	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP
Soyang	0.830	0.885	0.857	0.890	0.683	0.791	0.736	0.801	0.986	0.985	0.986	0.985
Daecheong	0.773	0.800	0.773	0.801	0.594	0.649	0.594	0.649	0.946	0.947	0.946	0.947
Andong	0.814	0.864	0.818	0.861	0.652	0.755	0.659	0.747	0.985	0.979	0.985	0.980
Seomjin	0.703	0.801	0.701	0.802	0.512	0.664	0.501	0.667	0.891	0.935	0.897	0.935
Chungju	0.854	0.877	0.853	0.878	0.708	0.755	0.705	0.757	1.000	0.998	1.000	0.998
Hapcheon	0.858	0.898	0.856	0.897	0.726	0.800	0.721	0.798	0.969	0.981	0.969	0.981
Namgang	0.843	0.891	0.865	0.899	0.676	0.777	0.723	0.794	0.984	0.988	0.984	0.987
Imha	0.660	0.798	0.692	0.798	0.444	0.677	0.509	0.680	0.872	0.917	0.872	0.915

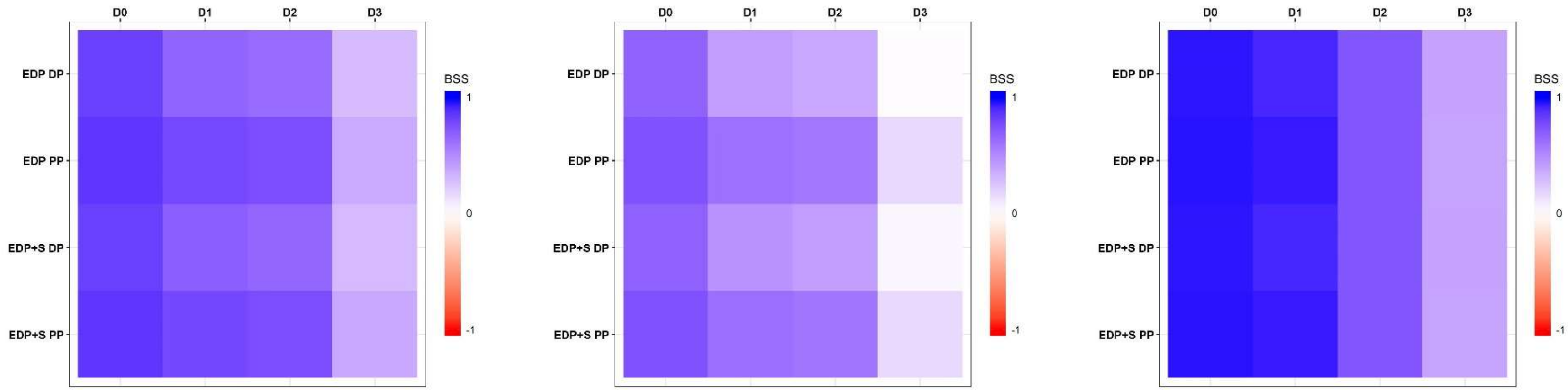


(a) All **(b) Irrigation** **(c) Non irrigation**

Figure A2.13 RPSS of two EDPs for SRI12

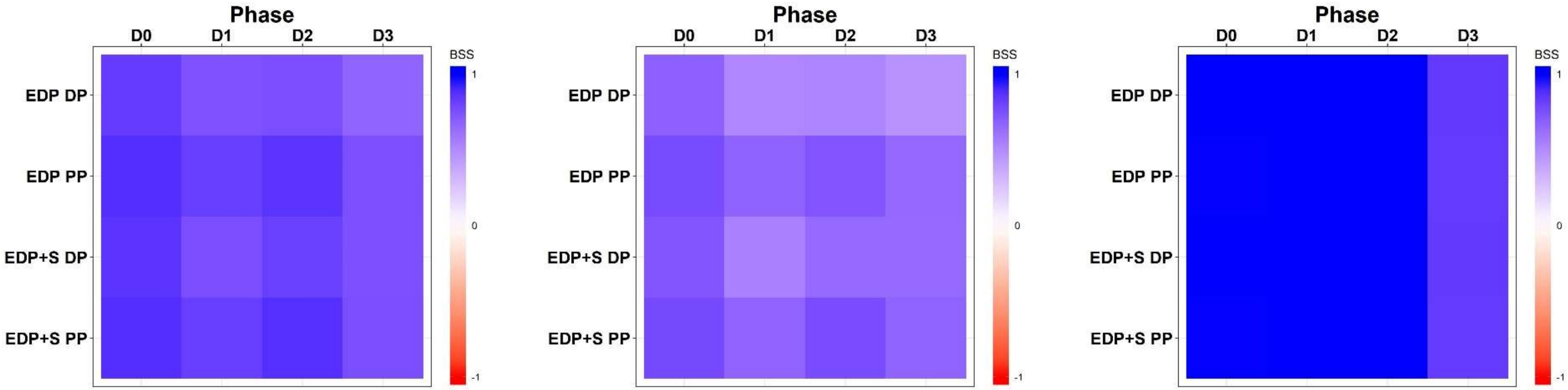
Table A2.6 Averaged BS of two EDPs across eight basins for SRI12

Phase \ Case	All				Irrigation				Non-irrigation			
	EDP		EDP+S		EDP		EDP+S		EDP		EDP+S	
	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP	DP	PP
D0	0.0441	0.0345	0.0435	0.0344	0.0797	0.0617	0.0784	0.0616	0.0086	0.0070	0.0086	0.0071
D1	0.0453	0.0307	0.0424	0.0304	0.0784	0.0534	0.0729	0.0530	0.0123	0.0078	0.0119	0.0078
D2	0.0251	0.0175	0.0233	0.0170	0.0490	0.0335	0.0453	0.0324	0.0012	0.0013	0.0012	0.0014
D3	0.0116	0.0095	0.0110	0.0095	0.0196	0.0152	0.0184	0.0152	0.0037	0.0037	0.0037	0.0038



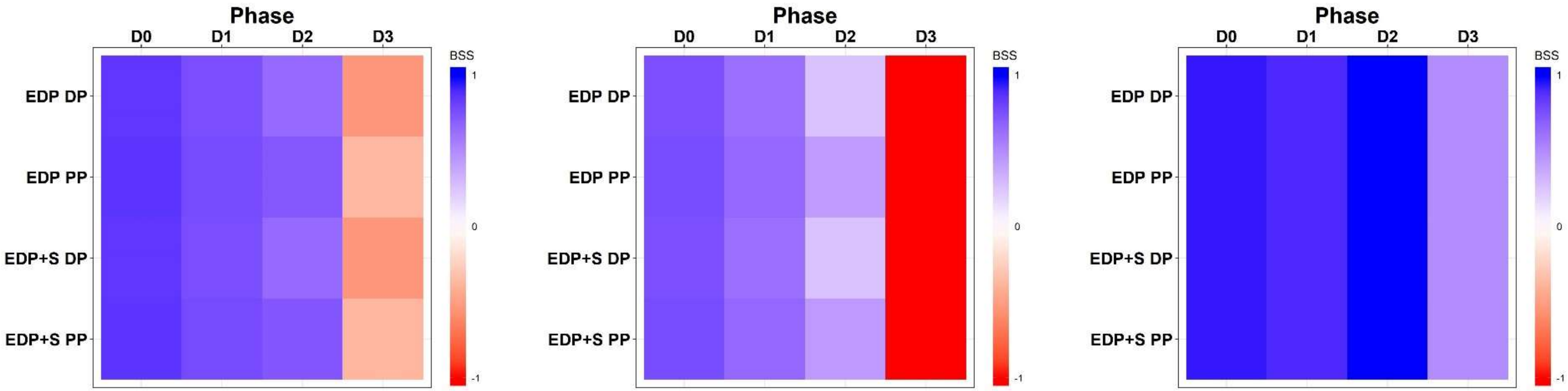
(a) All (b) Irrigation (c) Non irrigation

Figure A2.14 Averaged BSS of two EDPs across eight basins for SRI12



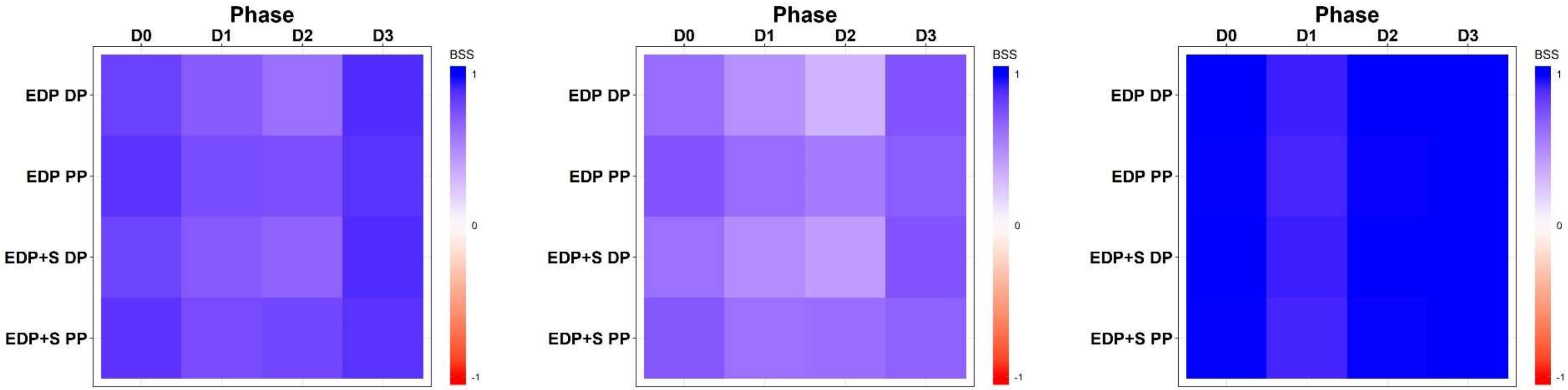
(a) All (b) Irrigation (c) Non irrigation

Figure A2.15 BSS at Soyang for SRI12



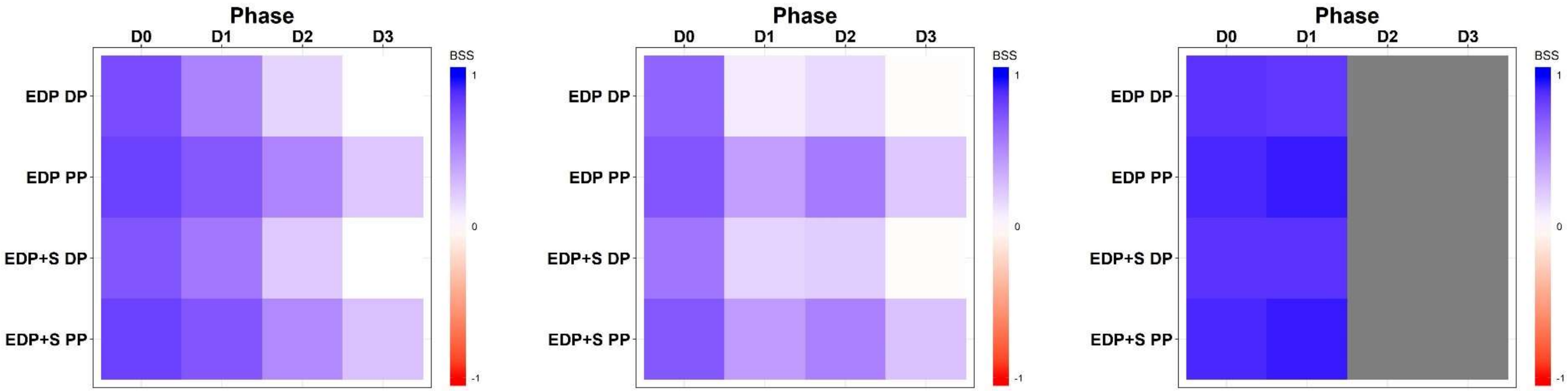
(a) All (b) Irrigation (c) Non irrigation

Figure A2.16 BSS at Daecheong for SRI12



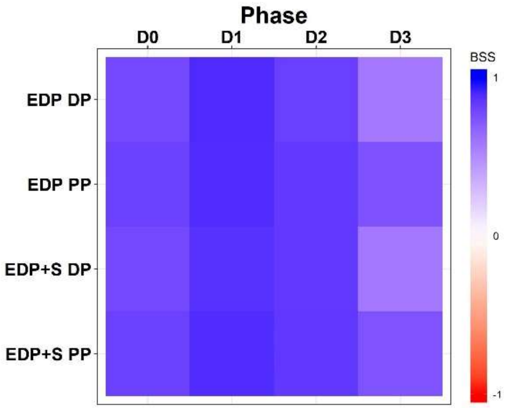
(a) All **(b) Irrigation** **(c) Non irrigation**

Figure A2.17 BSS at Andong for SRI12

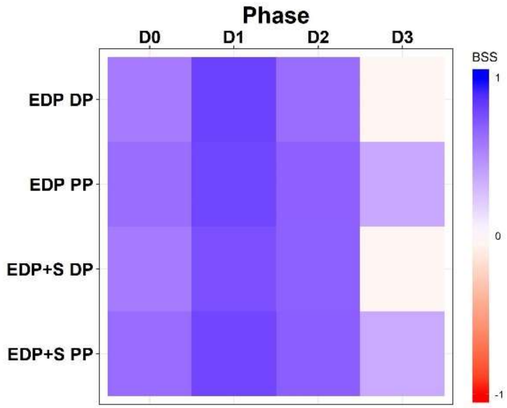


(a) All **(b) Irrigation** **(c) Non irrigation**

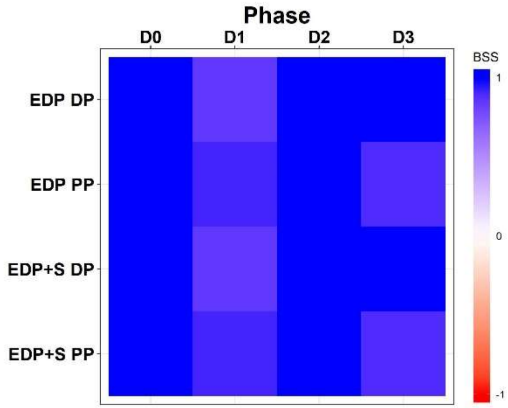
Figure A2.18 BSS at Seomjin for SRI12



(a) All

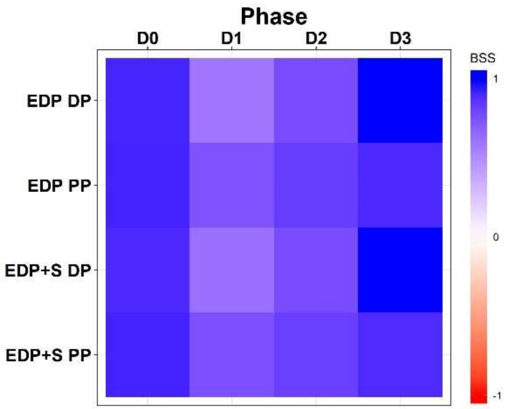


(b) Irrigation

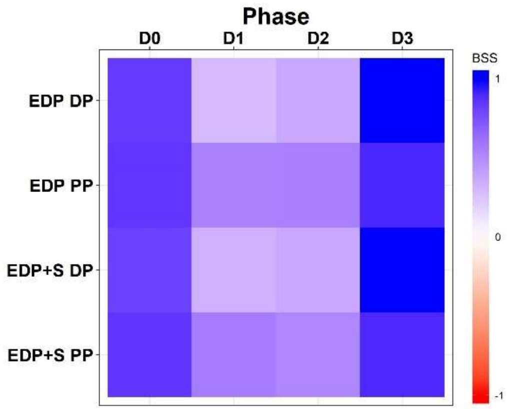


(c) Non irrigation

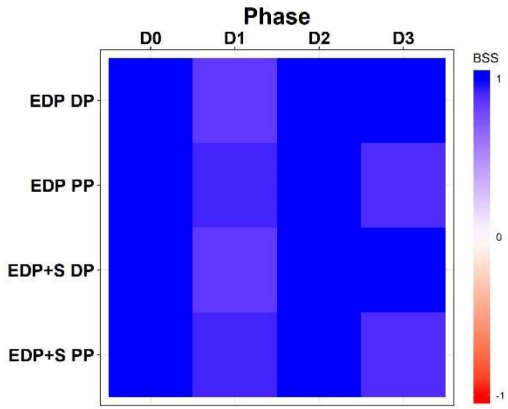
Figure A2.19 BSS at Chungju for SRI12



(a) All

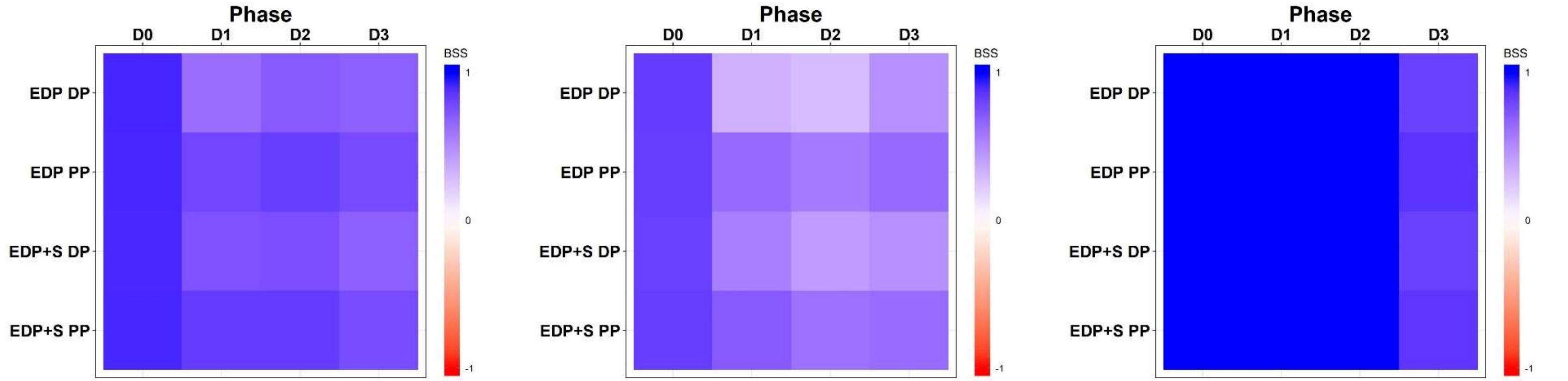


(b) Irrigation



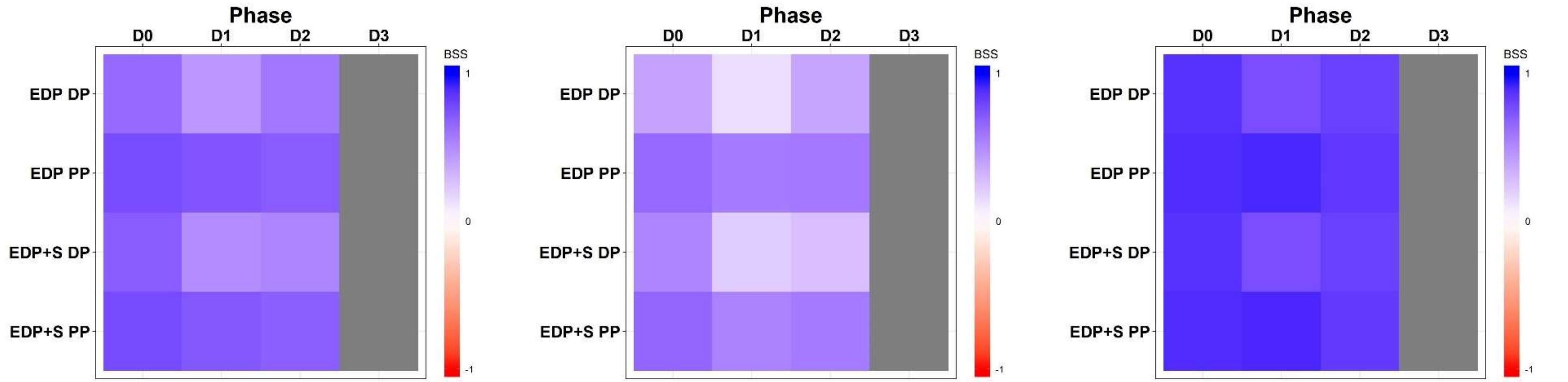
(c) Non irrigation

Figure A2.20 BSS at Hapcheon for SRI12



(a) All (b) Irrigation (c) Non irrigation

Figure A2.21 BSS at Namgang for SRI12



(a) All (b) Irrigation (c) Non irrigation

Figure A2.22 BSS at Imha for SRI12

Table A2.7 RMSE of four EDPs for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)

Case	Soyang	Daecheong	Andong	Seomjin	Chungju	Hapcheon	Namgang	Imha
EDP	0.608	0.593	0.563	0.491	0.741	0.486	0.502	0.633
EDP+S	0.528	0.586	0.538	0.491	0.657	0.500	0.502	0.624
EDP+A	0.624	0.604	0.578	0.504	0.762	0.488	0.514	0.594
EDP+AS	0.523	0.595	0.564	0.504	0.647	0.504	0.513	0.624

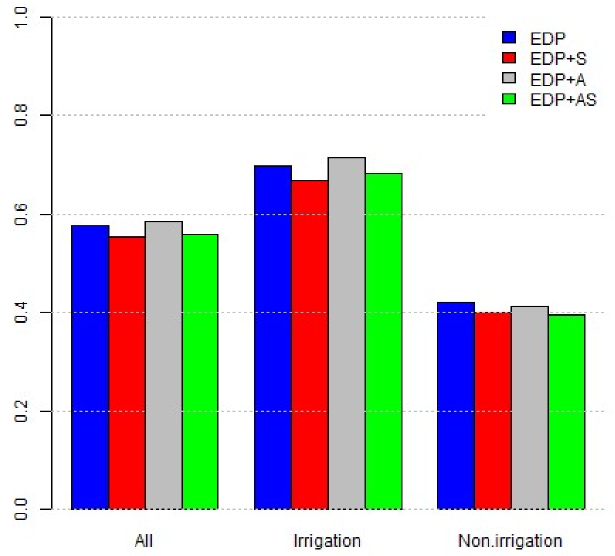
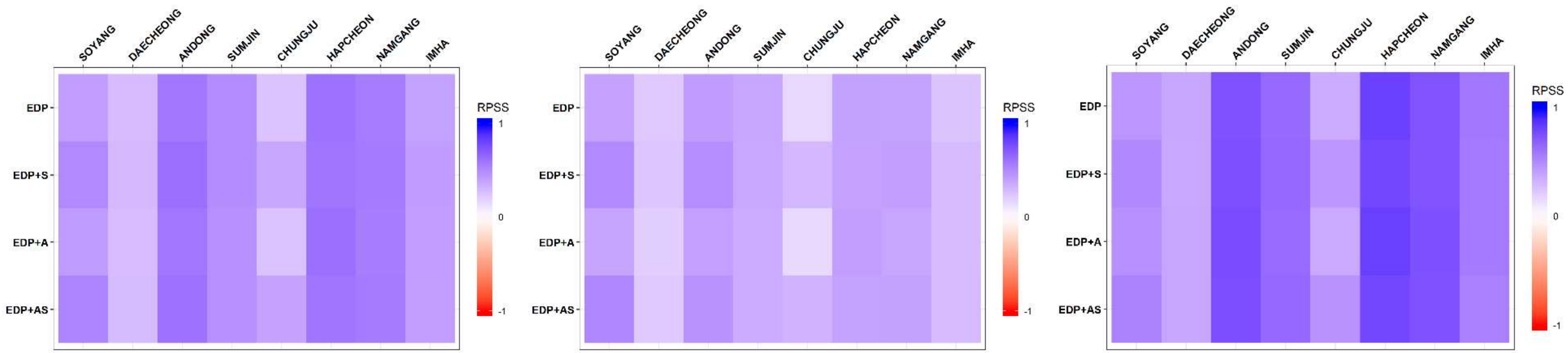


Figure A2.23 Averaged RMSE of four EDPs across all eight basins for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)

Table A2.8 RPSS of four EDPs for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)

Case / Basin	All				Irrigation				Non-irrigation			
	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS
Soyang	0.418	0.505	0.427	0.525	0.393	0.504	0.387	0.516	0.452	0.507	0.480	0.536
Daecheong	0.289	0.297	0.281	0.292	0.229	0.244	0.209	0.229	0.374	0.373	0.383	0.382
Andong	0.585	0.619	0.588	0.614	0.424	0.484	0.408	0.461	0.748	0.754	0.768	0.766
Seomjin	0.494	0.493	0.476	0.478	0.381	0.372	0.352	0.351	0.645	0.654	0.641	0.649
Chungju	0.258	0.374	0.256	0.394	0.169	0.307	0.160	0.328	0.358	0.449	0.363	0.468
Hapcheon	0.611	0.592	0.617	0.594	0.403	0.394	0.411	0.402	0.820	0.791	0.823	0.787
Namgang	0.561	0.570	0.559	0.568	0.397	0.414	0.383	0.404	0.741	0.741	0.752	0.749
Imha	0.402	0.420	0.419	0.412	0.247	0.290	0.286	0.296	0.584	0.571	0.575	0.548

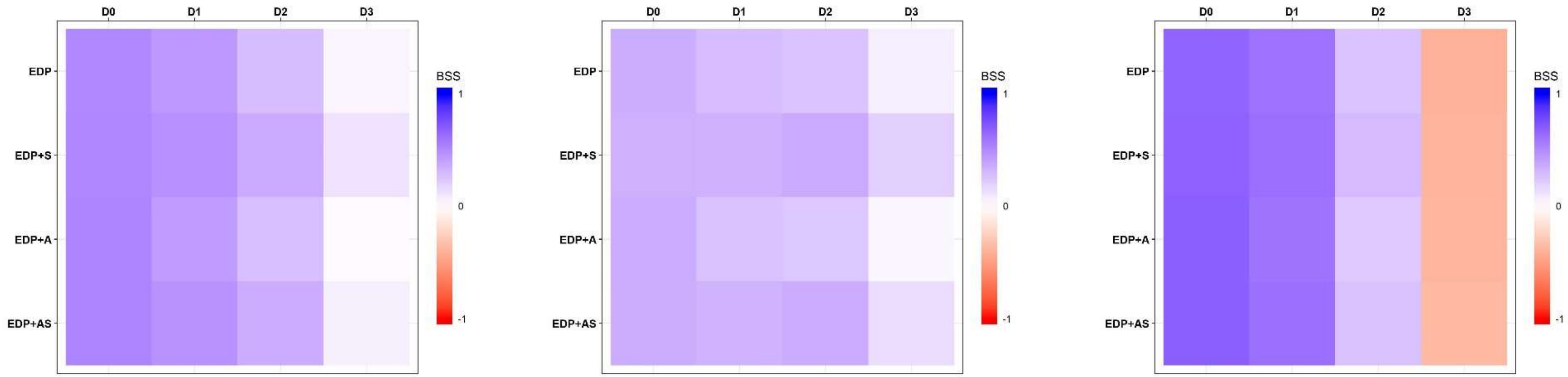


(a) All **(b) Irrigation** **(c) Non irrigation**

Figure A2.24 RPSS of four EDPs for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)

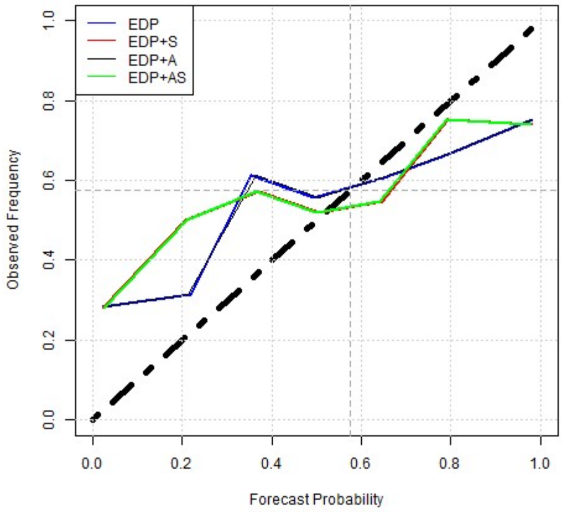
Table A2.9 Averaged BS of four EDPs across eight basins for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)

Case / Phase	All				Irrigation				Non-irrigation			
	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS
D0	0.1127	0.1123	0.1094	0.1098	0.1466	0.1490	0.1456	0.1469	0.0784	0.0756	0.0730	0.0728
D1	0.1003	0.0939	0.1026	0.0948	0.1352	0.1255	0.1392	0.1270	0.0650	0.0618	0.0658	0.0625
D2	0.0656	0.0581	0.0660	0.0585	0.0830	0.0708	0.0845	0.0717	0.0481	0.0456	0.0476	0.0456
D3	0.0304	0.0273	0.0308	0.0276	0.0419	0.0355	0.0431	0.0370	0.0190	0.0190	0.0183	0.0183

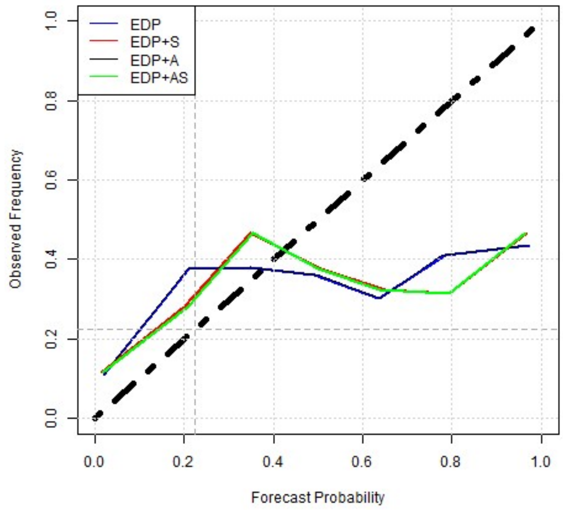


(a) All **(b) Irrigation** **(c) Non irrigation**

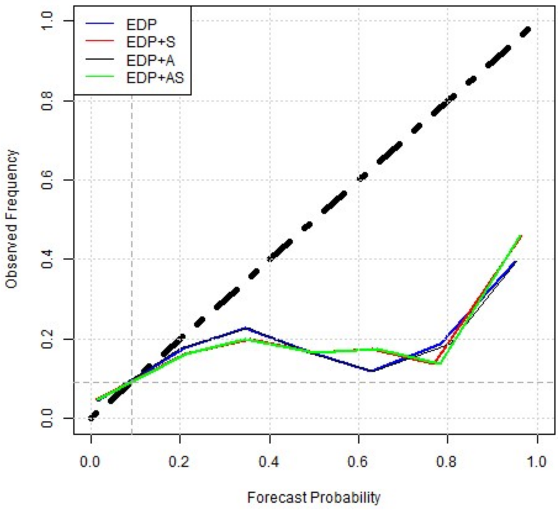
Figure A2.25 Averaged BSS of four EDPS across all eight basins for SRI3 (EDP, EDP+S, EDP+A and EDP+AS)



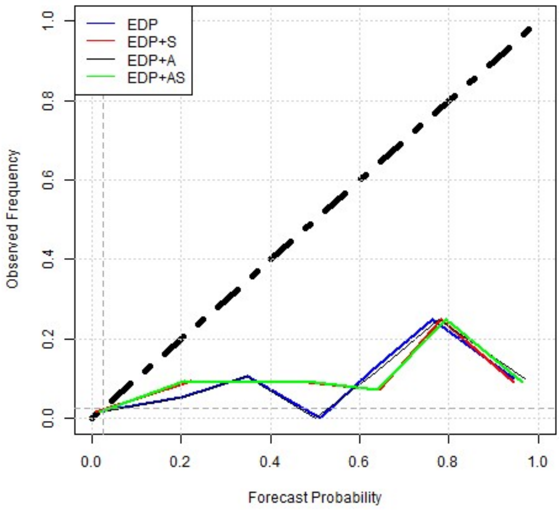
(a) D0



(b) D1



(c) D2



(d) D3

Figure A2.26 Reliability diagram of four EDPs for SR3

Table A2.10 RMSE of four EDPs for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)

Case	Soyang	Daecheong	Andong	Seomjin	Chungju	Hapcheon	Namgang	Imha
EDP	0.278	0.258	0.236	0.255	0.264	0.222	0.242	0.253
EDP+S	0.272	0.257	0.238	0.245	0.262	0.227	0.237	0.255
EDP+A	0.287	0.262	0.245	0.263	0.269	0.224	0.251	0.242
EDP+AS	0.279	0.261	0.250	0.252	0.266	0.227	0.243	0.254

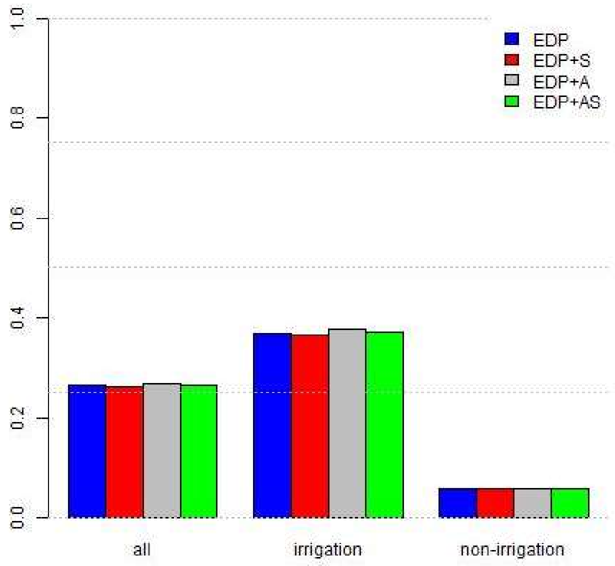
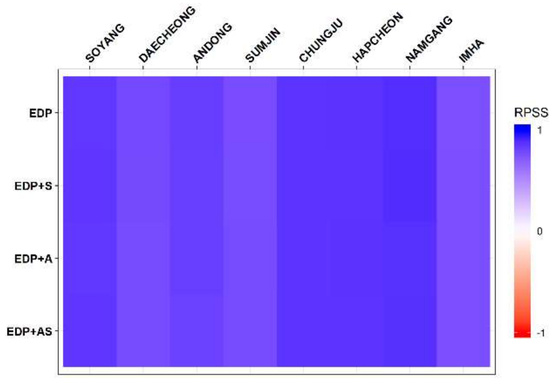


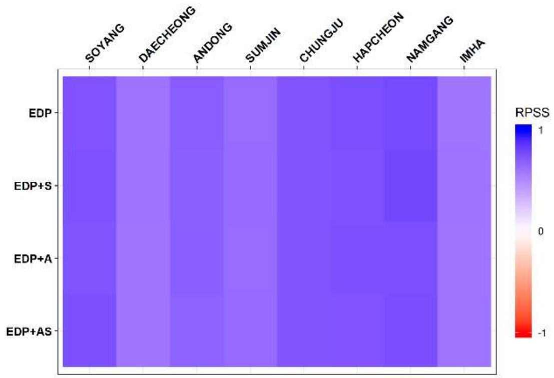
Figure A2.27 Averaged RMSE of four EDPs across eight basins for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)

Table A2.11 RPSS of four EDPs for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)

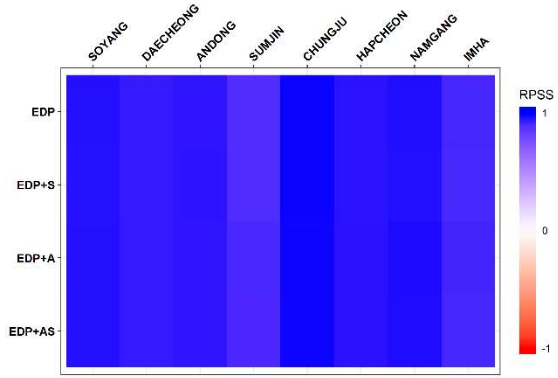
Case / Basin	All				Irrigation				Non-irrigation			
	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS
Soyang	0.854	0.858	0.853	0.860	0.739	0.748	0.737	0.751	0.975	0.975	0.975	0.975
Daecheong	0.783	0.783	0.777	0.777	0.605	0.604	0.594	0.593	0.955	0.955	0.955	0.955
Andong	0.832	0.826	0.826	0.817	0.705	0.692	0.694	0.677	0.965	0.967	0.964	0.965
Seomjin	0.770	0.771	0.777	0.777	0.641	0.646	0.641	0.643	0.896	0.895	0.910	0.909
Chungju	0.866	0.867	0.865	0.866	0.734	0.736	0.735	0.737	0.996	0.997	0.994	0.994
Hapcheon	0.871	0.867	0.871	0.866	0.757	0.747	0.752	0.741	0.967	0.967	0.970	0.970
Namgang	0.888	0.891	0.879	0.882	0.780	0.788	0.756	0.764	0.980	0.978	0.984	0.982
Imha	0.758	0.760	0.764	0.761	0.594	0.604	0.600	0.600	0.919	0.914	0.926	0.920



(a) All



(b) Irrigation

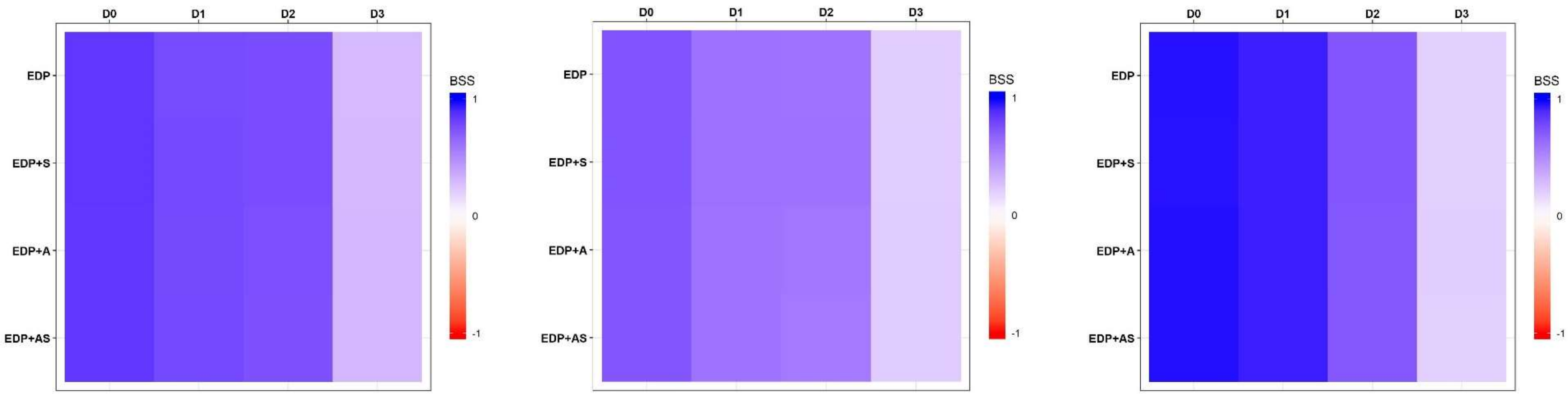


(c) Non irrigation

Figure A2.28 RPSS of four EDPs for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)

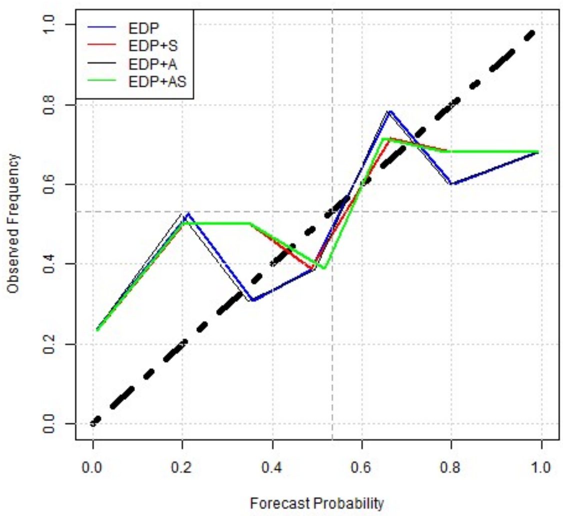
Table A2.12 Averaged BS of four EDPs across eight basins for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)

Case / Phase	All				Irrigation				Non-irrigation			
	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS	EDP	EDP+S	EDP+A	EDP+AS
D0	0.0313	0.0313	0.0322	0.0321	0.0565	0.0563	0.0590	0.0588	0.0060	0.0062	0.0054	0.0055
D1	0.0408	0.0409	0.0407	0.0408	0.0691	0.0691	0.0693	0.0697	0.0125	0.0125	0.0117	0.0118
D2	0.0229	0.0225	0.0238	0.0242	0.0434	0.0423	0.0446	0.0455	0.0022	0.0023	0.0028	0.0029
D3	0.0120	0.0120	0.0117	0.0117	0.0197	0.0197	0.0195	0.0194	0.0043	0.0044	0.0040	0.0041

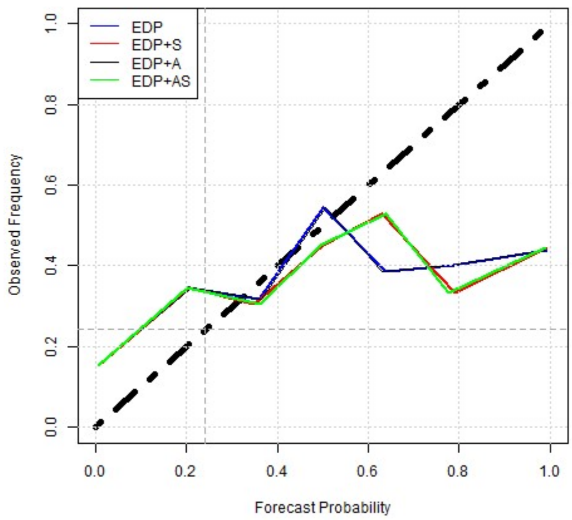


(a) All **(b) Irrigation** **(c) Non irrigation**

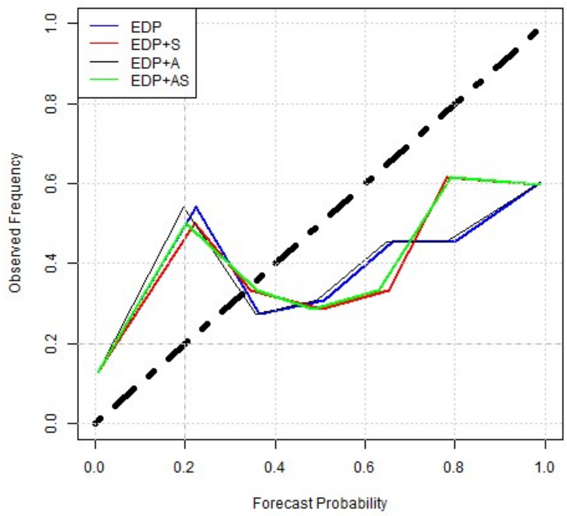
Figure A2.29 Averaged BSS of four EDPs across eight basins for SRI12 (EDP, EDP+S, EDP+A and EDP+AS)



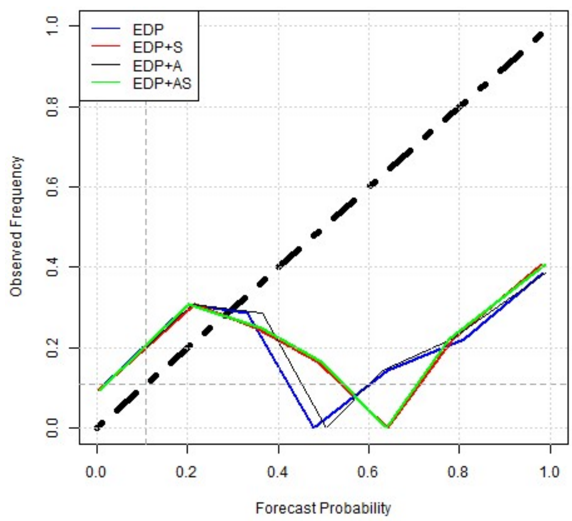
(a) D0



(b) D1



(c) D2



(d) D3

Figure A2.30 Reliability diagram of four EDPs for SRI12