THE PREDICTION OF HEAVY RAINFALL USING A NEURAL NETWORK AND CONSENSUS FORECAST METHOD FOR THE SYSTEM OF ENVIRONMENTAL CALAMITY PREVENTION

Pak K.S.[1], Jong S.I.[2] Email: Pak1133@scientifictext.ru

[1]Pak Ki Song – Candidate of geographic science, lecturer;
[2]Jong Sang Il – doctor of geographic science, professor, chairman of department,
Department of meteorology, Faculty of the Earth and Environmental Science,
***Kim Il Sung** University, Pyongyang, Democratic People's Republic of Korea*

***Abstract**: in our country during the main rainy seasons (mei – yu and summer) localized extremely heavy rainfall events over lots of regions can frequently result in flooding and landslides. The improvement of the prediction for heavy rainfall is very important to reduce its potential for damage.*
Here, we apply to process heavy rainfall forecasting using an artificial neural network OCF(Objective Consensus Forecasting) strategy, and estimate prediction skill by multi – model in contradistinction to each of models.
***Keywords**: heavy rainfall, neural network, consensus forecast.*

ПРОГНОЗИРОВАНИЕ СИЛЬНЫХ ДОЖДЕЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННОЙ СЕТИ И МЕТОДА КОНСЕНСУС–ПРОГНОЗА ПРИ ПРЕДОТВРАЩЕНИИ ОКРУЖАЮЩЕГО БЕДСТВИЯ

Пак. К.С.[1], Чен С.И.[2]

[1]Пак Ки Сон – кандидат географических наук, преподаватель;
[2]Чен Сан Ир – доктор географических наук, профессор, заведующий кафедрой,
кафедра метеорологии, факультет окружающей среды Земли,
*Университета им. **Ким Ир Сен**а, г. Пхеньян, Корейская Народно-Демократическая Республика*

***Аннотация:** в нашей стране во время основного сезона дождей (мэй - ю и летние), локализованные чрезвычайно ливневые осадки над большим количеством областей часто приводят к наводнениям и оползням. Улучшение прогноза для сильных дождей очень важно, чтобы уменьшить их потенциальный ущерб.*
Здесь мы применяем для обработки прогнозирования осадков нейронную сеть OCF (объективный консенсус прогнозирования), стратегии и оценку навыков прогнозирования с помощью мульти-модели в отличие от каждой из моделей.
***Ключевые слова**: сильные дожди, нейронная сеть, консенсус-прогноз.*

УДК 504.37

The heavy rainfall is decided as a precipitation phenomenon to rain 30 mm per 1hour (50 mm per 3 hours, 100 mm per 12 hours, 200 mm per 24 hours).

Generally, there are four types of forecasts: (1) qualitative forecast (whether we will have a heavy rainfall or not), (2) the probability of heavy rainfall forecast (what is the probability that we will have heavy rain?), (3) the probabilities of classified precipitation (what are the probabilities that the amount of rainfall will belong to the given categories?), and (4) the quantitative precipitation forecast.

This study focuses on the qualitative forecast (the dichotomous forecast) of the occurrence of heavy rainfall in our country. That is, heavy rainfall forecasting has binary values: whether we will have heavy rainfall or not.

The method of heavy rainfall forecasting, firstly was one on the basis of a synoptic chart, and then has been statistically post processed numerical model output.

Traditionally, the primary source of guidance for forecasting surface weather conditions has been statistically post processed numerical model output. In particular, model output statistics (MOS) derived from the Global Forecast System (GFS) provide forecasts of weather parameters at 6 h intervals out to 48 or 60 h, depending on the model.

This type of forecasting guidance has at least two limitations: 1) the models are run only a few times daily, allowing forecasts to become several hours old before an updated product is made available, and 2) the MOS equations are linear. To alleviate the first limitation, studies have been undertaken to investigate the use of an observations-based forecasting system [7]. In this system, a network of surface observations is used as predictors in a multiple regression technique. It was demonstrated that this approach could improve the accuracy of ceiling and visibility forecasts for the hours between the times that the output from the numerical models is released. Moreover, Leyton showed that the introduction of high – density observation networks and high frequency observations lead to further improvement [4].

The linearity of traditional MOS equations may be a limitation if the underlying relations are nonlinear. To allow for more general relations, nonlinear generalizations of multiple regressions have been utilized to model any nonlinearity and interactions of the underlying processes. Of course, the nonlinearity of the statistical model does not prevent it from capturing linear relations as well.

For example, temperature forecasts from the Advanced Regional Prediction System have been post processed via neural networks, displaying a reduction in bias and error variance of the forecasts [5].

There, it is found that the optimal neural network is indeed nonlinear. As such, the nonlinear statistical post processing yields rainfall forecasts that are more accurate than the model forecasts as well as MOS forecasts. Some recent applications of neural networks include nowcasting of visibility from surface observations [6].

Clemen surveyed over many papers drawn from meteorology [2]. He concluded that 1) accuracy can be substantially improved through the combination of multiple forecasts, 2) simple combinations often work better than more complex methods and 3) combining forecasts should become part of mainstream practice.

The objective consensus forecasting (OCF) system employs bias correction of both multi – model DMO and MOS component forecasts followed by consensus merging. It retains the flexibility of DMOs in generating forecasts for new sites and models and in quickly exploiting observational and numerical system enhancements yet still benefits from available bias corrected MOS forecasts.

Improvement in OCF accuracy as models improve is achieved by weighting bias corrected component DMO forecasts according to their recent accuracy while constraining the sum of weights to one. Gupta and Wilton [3], following a review of the literature on combining forecasts, suggested the following desirable properties of the compositing method.

1) It should not require large quantities of data for estimating weights.

2) It should distinguish between better and poorer available candidate models with the distinction being made on precision (i.e., low MAEs or rmse's) and redundancy (i.e., low cross correlations with other contributing forecasts).

3) Derived weights should be intuitively meaningful.

More precise and less redundant components should be given higher weight. It is not possible to process rainfall forecasts using the OCF strategy because of its discontinuous occurrence. The OCF prediction of rain and its probability of occurrence cannot use either bias correction or MAE weights effectively because rainfall is discontinuous.

Section 1 and section 2 describes basic principle of neural network and consensus method, and data to examine. Respectively, the main results are presented in section 3. Finally, section 4 provides a summary discussion and expected future developments.

1. The basic principle of neural network and the consensus forecast method

1.1 Neural network

When dealing with heavy rainfall modeling, the vector of input variables may consist of IP predictors (x^p) and IR precipitation (x^r) observations, i.e. both autoregressive and exogenous inputs are included, together giving the number of input variables (*g*). The vector of input variables (X_j) at the j–th time instant can be presented as

$$X_j = (x_j^p, \ldots, x_{j-IP+1}^p, x_j^r, \ldots, x_{j-IR+1}^r)^T = (x_1, x_2, \ldots, x_g)^T \qquad (1)$$

and corresponds to output variable y_{j+T} – which represents future precipitation. One can then proceed to the stage of determining the prediction model *F*

$$y_{j+T} = F(X_j) \qquad (2)$$

where *T* is the prediction horizon.

The objective function *J* to be minimized for each model by proper optimization of parameters (*h*) is defined in this paper as:

$$J = \min_h \sum_{j=1}^{n} (d_{j+T} - y_{j+T})^2 \qquad (3)$$

Where *n* is the number of training or validation set outputs and d_{j+T} is the measured value of flow corresponding to the *j* –th input vector.

1.2 Consensus forecast method

The consensus forecast method is one of combining forecast produces of different types of numerical forecast models by weighted average with its forecast skill.

The parameterization of numerical forecast model isn't far from perfect and the system errors in the numerical forecast produces remain inevitably because of the others of reason. Thus, the consensus forecast is used to diminish their system errors.

From here we explained about the consensus forecast method concretely.

The consensus method to combine of the different of numerical forecast produces is to weighted mean with their forecast skills on each of numerical forecast model. At this moment, each model's weight is related to its forecast skill and the sum of all model's weights is 1. If F_i are the forecasts and O_i are the observed values on some forecast times at some place with one numerical forecast model, Br (Brier Skill) [1] where

$$Br = \frac{1}{N}\sum_{i=1}^{N}\left(F_i - O_i\right)^2 \qquad (4)$$

Then about n contributing forecast models:

Normalized weighting parameters (w_j) are calculated by using the inverse mean Br from the Br samples of the *n* contributing model forecasts

$$w_j = \frac{\frac{1}{M_j}}{\sum_{j=1}^{K}\frac{1}{M_j}} \qquad (5)$$

Where M_j is the mean *Br* of *j* – th contributing forecast model.

$$M_j = \frac{1}{\mathrm{N}}\sum_{j=1}^{\mathrm{N}}\left(Br\right) \qquad (6)$$

Using these parameters, the consensus forecast P_i based on *n* model forecasts (f_i) is given by

$$P_i = \frac{\sum_{j=1}^{\mathrm{K}} w_j F_{i,\,j}}{\sum_{j=1}^{\mathrm{K}} w_j} \qquad (7)$$

2. Data

The 6-hours weather forecast system of RJTD and EDZW is now running in our country. For our study, 6 hour–interval data during 2011 to 2014 at 43 stations in Korea are used. The observations are area – mean amounts of rainfall. The predicted and obtained from observation, has a binary response (whether the heavy rain occurred or not). In Korea, it is defined that heavy rain occurs when the rainfall is over 30mm h^{-1}, 50 mm per 3 hours, or over 200mm d^{-1}. It is defined by being over 50mm in 6 hours for this study on the 6 weather forecast system.

The 26 synoptic factors at 43 stations are used as potential predictors for heavy rainfall forecasting. These factors include the wind direction and speed, relative humidity, thermal advection, potential precipitation and temperatures. They can be generated by the numerical model, called RJTD and EDZW used in Korea. And some previous observations are added.

Table 1 shows that effective prediction induces for heavy rainfall in the precious researches about one's forecast.

Table 1. Potential Predictors

Acronym	prediction index	Acronym	prediction index
NP	Numerical Precipitation	DH(85)	Difference of gravity potential height between 850hPa and 500hPa
BI	Boyden index	KI	K index
JI	Jafferson index	Rackliff	Rackliff index
MJI(1)	Modified Jafferson index(1)	MJI(1)	Modified Jafferson index(2)
TTI	Total totals index	MTTI(2)	Modified TT index(2)
MTTI(1)	Modified TT index(1)	PWBI	Potential wet-bulb temperature index
LFI	Lifted Force index	CII	Convective instability index
VTG	Vertical temperature gradient	PCI	Potential convection index
TA(1085)	Thermal advection Between 1000hPa and 850hPa	DT1 000	Dew-point temperature at 1 000hPa

CVW	Change of vertical wind	DT925	Dew-point temperature at 925hPa
AI	A index	DT850	Dew-point temperature at 850hPa
SSI	Showalter stability index	DT700	Dew-point temperature at 700hPa
RCH	The rising congealing height	DT500	Dew-point temperature at 500hPa

3. Results

3.1 Prediction of heavy rainfall by neural network

We experiment prediction for heavy rainfall classified by each of predictors using products of two types of numerical forecast model (RJTD and EDZW) with the neural network.

Table 3 and table 4 show PSS of prediction for heavy rainfall classified by each of predictors.

Table 2. PSS of prediction for heavy rainfall classified by each of predictors using products of RJTD in prediction period

predictor	**Prediction period(h)**					
	6	**12**	**18**	**24**	**30**	**36**
NP	0.108	0.1054	0.099	0.094	0.088	0.079
BI	0.105	0.102	0.095	0.087	0.071	0.069
JI	0.394	0.367	0.358	0.312	0.304	0.291
MJI(1)	0.502	0.476	0.427	0.398	0.384	0.378
TTI	0.493	0.471	0.456	0.451	0.442	0.425
MTTI(1)	0.341	0.331	0.317	0.296	0.272	0.266
LFI	0.491	0.488	0.466	0.454	0.437	0.429
VTG	0.515	0.498	0.457	0.444	0.426	0.401
TA(1085)	0.061	0.057	0.05	0.041	0.0346	0.022
CVW	0.1553	0.155	0.145	0.142	0.137	0.127
AI	0.489	0.485	0.474	0.461	0.412	0.407
SSI	0.476	0.469	0.451	0.437	0.4012	0.216
RCH	0.095	0.087	0.076	0.071	0.064	0.057
DH(85)	0.129	0.131	0.124	0.119	0.128	0.104
KI	0.437	0.433	0.417	0.408	0.395	0.401
Rackliff	0.099	0.101	0.094	0.088	0.082	0.084
MJI(1)	0.295	0.291	0.288	0.281	0.264	0.257
MTTI(2)	0.448	0.441	0.429	0.411	0.409	0.405
PWBI	0.465	0.451	0.449	0.417	0.406	0.398
CII	0.256	0.236	0.231	0.224	0.213	0.207
PCI	0.188	0.185	0.166	0.161	0.157	0.155
DT1 000	0.109	0.101	0.096	0.094	0.088	0.071
DT925	0.436	0.431	0.42	0.404	0.4	0.389
DT850	0.124	0.117	0.111	0.105	0.11	0.093
DT700	0.184	0.18	0.171	0.154	0.142	0.134
DT500	0.179	0.166	0.16	0.147	0.14	0.133

Table 3. PSS of prediction for heavy rainfall classified by each of predictors using products of EDZW in prediction period

predictor	Prediction period(h)					
	6	**12**	**18**	**24**	**30**	**36**
NP	0.132	0.129	0.127	0.121	0.113	0.109
BI	0.170	0.167	0.164	0.148	0.143	0.139
JI	0.122	0.119	0.113	0.107	0.103	0.101
MJI(1)	0.351	0.343	0.341	0.339	0.337	0.316
TTI	0.413	0.408	0.402	0.396	0.387	0.351
MTTI(1)	0.399	0.391	0.374	0.362	0.354	0.349
LFI	0.477	0.471	0.462	0.454	0.447	0.439
VTG	0.480	0.477	0.473	0.472	0.468	0.452
CVW	0.396	0.387	0.371	0.366	0.359	0.352
AI	0.362	0.363	0.352	0.343	0.332	0.316
SSI	0.449	0.447	0.438	0.432	0.411	0.407
RCH	0.109	0.112	0.103	0.11	0.116	0.097
DH(85)	0.399	0.397	0.382	0.374	0.366	0.358
KI	0.102	0.092	0.099	0.087	0.094	0.076
Rackliff	0.361	0.346	0.327	0.313	0.302	0.299
MJI(1)	0.386	0.383	0.381	0.369	0.374	0.362
MTTI(2)	0.410	0.402	0.396	0.374	0.362	0.358
PWBI	0.224	0.214	0.204	0.206	0.213	0.197
CII	0.256	0.251	0.249	0.248	0.241	0.236
PCI	0.288	0.289	0.257	0.243	0.232	0.245
DT1 000	0.116	0.121	0.117	0.116	0.108	0.100
DT925	0.187	0.18	0.182	0.142	0.165	0.170
DT850	0.193	0.188	0.176	0.160	0.152	0.160
DT700	0.110	0.100	0.101	0.099	0.111	0.098
DT500	0.167	0.166	0.161	0.164	0.154	0.153

As Table 3 and Table 4 shows, in PSS of prediction for heavy rainfall classified by each of predictors using products derived from RJTD and EDZW in prediction period, VTG, LFI, KI, AI are higher than the others. Also, the longer of prediction period, the smaller of forecast skills.

Fig 1 shows PSS of prediction for heavy rainfall classified by each of predictors in contradistinction to PSS of prediction for heavy rainfall by all predictors using products derived from RJTD.

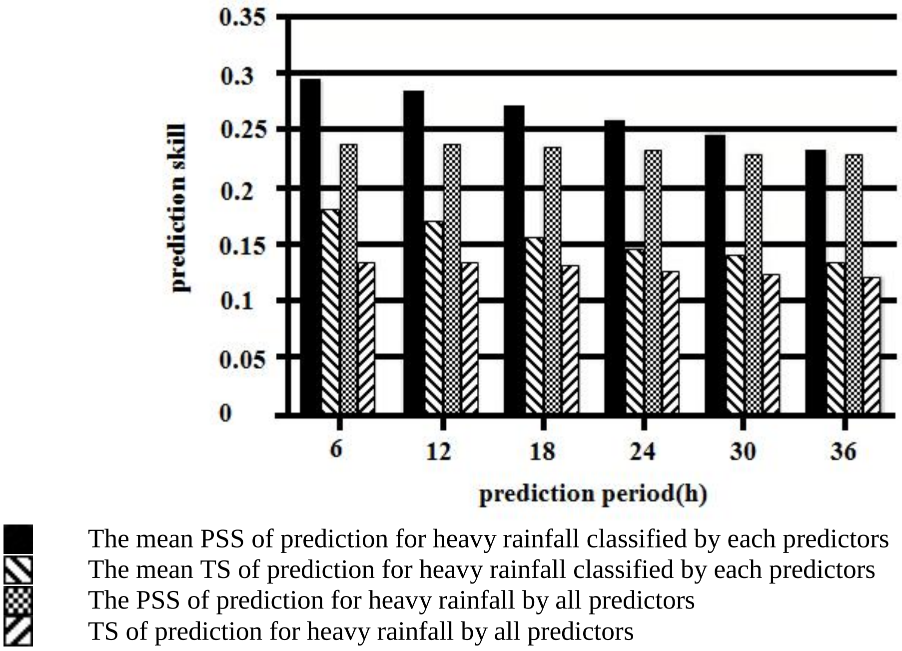


Fig. 1. PSS of prediction for heavy rainfall classified by each of predictors and of prediction for heavy rainfall by all predictors using products derived from RJTD

As Fig 1 shows, the mean PSS and TS of prediction for heavy rainfall classified by each predictors is about 0.06 and 0.05 higher than the PSS and TS of prediction for heavy rainfall classified by all predictors, and. through these courses, prediction skill for heavy rainfall with all predictors as the input dataset of neural network is low, but the change of input datasets of neural network improve prediction skills for heavy rainfall by ones.

Generally, the smaller the coefficient of correlations between variables used input datasets of neural network are, the bigger the exact of neural network's training is, so we estimated the biggest PSS classified by prediction periods so that collated predictors with small coefficient of correlations.

Table 5 shows collated predictors with biggest PSS and table 6 is ones classified by prediction periods.

Table 4. Collated predictors with biggest PSS (prediction period=6h)

Collated predictors	PSS	TS
AI, TTI, NP,CVW, PWBI, CII	0.573	0.292
PCI, DT925,KI, DH(85), TA(1085), BI	0.562	0.287
DH(85), AI, PCI, CII, DT500, DT700, TTI, TA(1085)	0.560	0.276
PWBI, DT1000, MTTI(2), DH(85), RCH, TA(1085)	0.551	0.263
MTTI(1), SSI, TA(1085)	0.551	0.261

Table 5. Collated predictors with biggest PSS

Prediction period	Collated predictors	PSS	TS
6h	AI,TTI,NP,CVW,PWBI,CII	0.573	0.292
12h	TTI,AI,DT1000,TA(1085),RCH,NP	0.496	0.268
18h	MJI(2),DT(925),KI,DH(85),TA(1085),BI	0.481	0.251
24h	AI,TTI,NP,CVW,PWBI,CII	0.472	0.248
30h	DH(85),AI,MJI(2),CII,DT500,DT700,TTI,TA(1085)	0.458	0.236
36h	DT1000,MJI(1),CVW,TA(1085)	0.432	0.218

The coefficient of correlations between AI, TTI, NP, CVW, PWBI, CII used input datasets of neural network are 0.006~0.009.The PSS of prediction for heavy rainfall by these collated predictors are higher about 0.25(PSS) and 0.1(TS) than the mean of prediction skills for heavy rainfall classified by each of predictors, are higher also about 0.31(PSS) and 0.16(TS) than the prediction skill for heavy rainfall by all predictors (Fig 1).

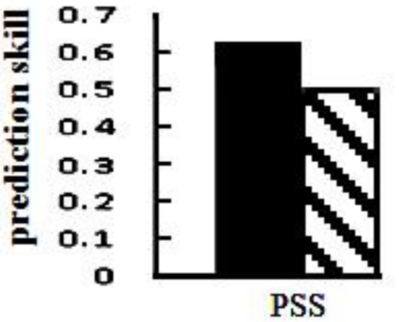
3.2 Prediction for heavy rainfall by consensus method.

a. The consensus forecasting for heavy rainfall by collated predictors.

Let’s see that PSS of prediction for heavy rainfall by effective collated predictors for input datasets of neural network is higher than by all predictors or by each of predictors.

In equation (4), F_i, O_i have binary values: whether we will have heavy rainfall or not.

Fig 2 shows PSS of prediction for heavy rainfall by the consensus method of the series of effective collated predictors classified by each method. As figure 2 show, PSS of prediction for heavy rainfall by first consensus method of collated predictors using products derived from RJTD and EDZW in 6–h period is higher about 0.08 ~ 0.21 than the others, and higher about 0.3than mean ones by each predictors. For example it is higher about 0.11 than by VTG (0.505).



RJ

ED

PSS of prediction for heavy rainfall by consensus method of collated predictors using products derived from RJND in 6–h period

Fig. 2. PSS of prediction for heavy rainfall by the consensus method of the series of effective collated predictors

Table 6 shows the first consensus method of the series of effective collated predictors with the highest PSS by RJTD .On occasion by EDZW are same.

Table 6. Consensus method of the series of effective collated predictors with the highest PSS by RJTD

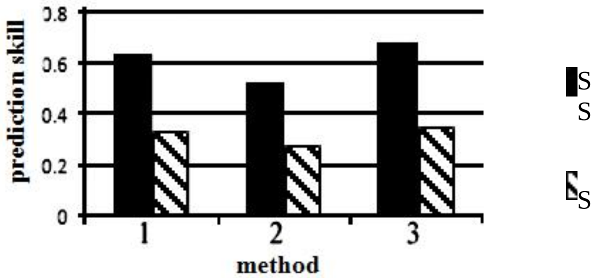
effective collated predictors	PSS
AI, CII, PWBI, BI, PCI, DT1000, DH85, NP, TTI	0.55
SSI, MJI(1), DT500, DT925, TA(1085)	0.51
JI,DT700, MTTI(1), MJI(2), CVW, RCH, Rackliff, DT850	0.468
KI,MTTI(2)	0.35
VTG	0.48
LFI	0.52

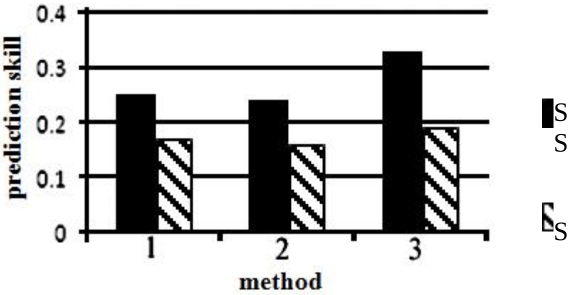
Starting from these facts, we can understand that the forecasting skill for heavy rainfall by the consensus method of the of series of effective collated predictors is higher remarkably than by all predictors and by each of predictors using products derived from RJTD as well as EDZW.

b. Consensus forecasting by combining RJTD and EDZW

Let’s see PSS and TS for heavy rainfall is high when we forecast by the above-mentioned consensus method of combing RJTD and EDZW. Here, from equation (4~7) we estimated PSS and TS for heavy rainfall by multi – model consensus method in contrasts of RJTD and EDZW.

Fig. 3 shows that PSS and TS for heavy rainfall by multi – model consensus method are high about 0.051 and 0.13 than by RJTD and EDZW.





[1] – PSS and TS of prediction for heavy rainfall by consensus method of collated predictors using products derived from RJTD
[2] – PSS and TS of prediction for heavy rainfall by consensus method of collated predictors using products derived from EDZW
[3] – PSS and TS for heavy rainfall is high when we forecast by the consensus method of combing RJTD and EDZW

Fig. 3. PSS and TS for heavy rainfall by multi-model consensus method in contrasts of RJTD and EDZW

[1] – PSS and TS of prediction for heavy rainfall by consensus method of all predictors using products derived from RJTD
[2] – PSS and TS of prediction for heavy rainfall by consensus method of all predictors using products derived from EDZW
[3] – PSS and TS for heavy rainfall is high when we forecast by the consensus method of combing RJTD and EDZW

Fig. 4. PSS and TS for heavy rainfall by multi-model consensus method in contrasts of RJTD and EDZW with all predictors

Next, we estimated PSS and TS for heavy rainfall by multi – model consensus method in contrasts of RJTD and EDZW with all predictors for input datasets of neural network. AS figure 4 shows, PSS and TS for heavy rainfall by multi – model consensus method is high about 0.06 and 0.08 than RJTD and EDZW.

4. Discussion

The conclusions of this study are follows:

– The skill of prediction for heavy rainfall by effective collated predictors for input datasets of neural network is higher than by all predictors or by each of predictors.

– The skill of prediction for heavy rainfall using products derived from RJTD is higher than from EDZW.

– The prediction skill for heavy rainfall by multi – model consensus method is high about than RJTD and EDZW.

References

1. *Applequist S. et al.* Comparison of Methodologies for Probabilistic Quantitative Prediction Forecasting. Weather and Forecasting, 2002. № 17. Pp. 783 - 799.
2. *Clemen R.T.* Combining forecasts, a review and annotated bibliography. Int. J. Forecasting, 1989. 5, 559 – 583.
3. *Gupta S. and Wilton P.C.* Combination of forecasts: An extension. Manage. Sci., 1987. 33, 356 – 372.
4. *Leyton S.M. and Fritsch J.M.* Short-term probabilistic forecasts of ceiling and visibility utilizing high-density surface weather observations. Wea. Forecasting, 2003. 18, 891 - 902.
5. *Marzban C.A* Neural network for post-processing model output: ARPS. Mon. Wea. Rev., 2003. 131, 1103 – 1111.
6. *Pasini A., Pelino V. and Potesta S.* A neural network for visibility nowcasting from surface observations: Results and sensitivity to physical input variables. Geophys. Res., 2001. 106, 14951 – 14959.
7. *Vislocky R.L. and Fritsch J.M.* An automated, observations-based system for short-term prediction of ceiling and visibility. Wea. Forecasting, 1997. 12, 31 – 43.