ESTIMATIONS OF STREAMFLOW CHARACTERISTICS THROUGHOUT JAPAN USING ARTIFICIAL NEURAL NETWORK

RYOSUKE ARAI
Central Research Institute of Electric Power Industry, 1646Abiko, Abiko-shi, Chiba, Japan, e-mail arai@criepi.denken.or.jp

YASUSHI TOYODA
Central Research Institute of Electric Power Industry, 1646Abiko, Abiko-shi, Chiba, Japan, e-mail toyo@criepi.denken.or.jp

SO KAZAMA
Graduate School of Engineering, Tohoku University, Aoba-yama 6-6-06, Sendai, 980-8579, Japan, so.kazama.d3@tohoku.ac.jp

ABSTRACT

Japan has a diversity of climatic division and one of the heaviest snowfalls in the world. To estimate streamflow characteristics throughout Japan is one of the greatest challenges for hydrologists. Recently, artificial neural networks (ANNs) gained an attention as an approach to estimate streamflow (Q) characteristics in absent of discharge data. We developed ANNs to estimate the Q characteristics inputting enormous basin characteristics throughout Japan. The Q characteristics were obtained from observed discharge data in 448 target basins. We employed the 14 Q characteristics including mean annual runoff height and flow percentiles in flow duration curves. The 175 basin characteristics including climate, land use, geology, soil, and topography were used as input data of the network. The network performed the best in mean annual runoff height (R^2 = 0.72) and the worst in 99 percentile of flow duration curves (R^2 = 0.18). We also evaluated the relationship between Q characteristics and basin characteristics. The results showed that the Q characteristics were explained dominantly by precipitation and aridity index. We found a limited geological effect on low flow regime, whose effect may have been weakened by extreme snowmelt contributions. We validated a relationship between the number of training data and the performance of ANNs. The results showed that increasing training data had a possibility for improvement of the accuracy, especially in low flow regimes. This will be a hope to improve the performance in low flow regimes.

Keywords: absent of discharge data, observed streamflow, cross validation

1. INTRODUCTION

Japan has a climatic diversity including cold, subarctic, temperate, and tropical zone. The climatic conditions influences on not only the overall water balance (Budyko 1974; Zhang et al., 2004; Porporato et al., 2004; Milly, 2001) but also baseflow (Price, 2011). In addition, Japan has one of the heaviest snowfalls in the world (Kazama and Sawamoto, 1997). Some areas experience usually snow depths of more than 5 m even in law latitudes (Kazama et al., 2008). These conditions indicate a difficulty to estimate flow regimes throughout Japan.

In recent years, Beck et al. (2015) and Barbarossa et al. (2018) estimated streamflow regimes globally using artificial neural networks (ANNs). These studies set streamflow (Q) characteristics as the outputs and set basin characteristics (e.g., climate, topography, land use) as the inputs, and then train the ANNs by using global and enormous observed discharge data. In other words, the ANNs can estimate Q characteristics based on a relationship between Q characteristics and basin characteristics in other basins. According to these studies (Beck et al., 2015; Barbarossa et al., 2018), the ANNs showed highly performances. Beck et al. (2015) compared performances of the ANNs to that of some macroscale hydrological models. The results showed that the ANNs outperformed the macroscale hydrological models in all Q characteristics. These facts indicate that the ANNs are one of the most effective approaches to estimate streamflow regimes in absent of discharge data.

In this study, we developed the ANNs to estimate streamflow regimes throughout Japan. The objective of this study is to validate the performance of ANNs quantitatively. Note that there are dams or hydropower weirs in major Japanese basins. Thus, we selected sites gauging discharge with no dams and hydropower weirs upstream as target basins. Next, we prepared the Q characteristics and basin characteristics for the selected basins, and then developed the ANNs. Finally, we validated performances of the ANNs quantitatively.

Table 1. Overview of target basins.

Number of basins	Catchment size [km^2]		Period of records [y]	
	Median	Average	Median	Average
448	72.3	163.6	19	18.6

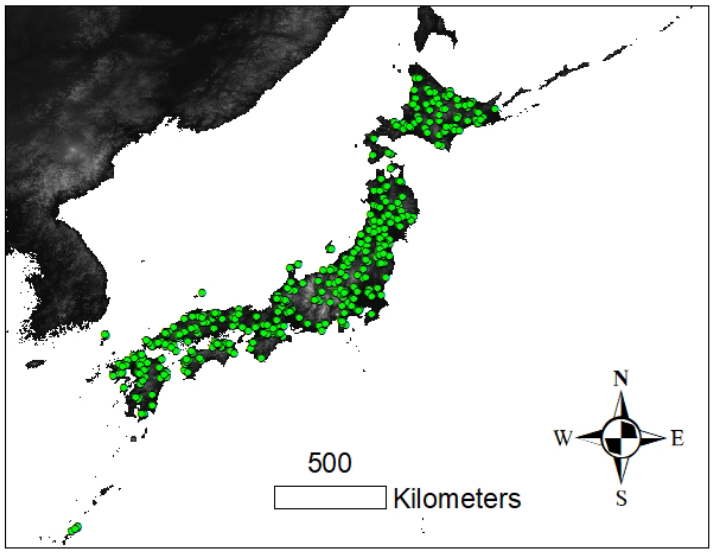

Figure 1. Location of target basins.

2. TARGET BASINS

We selected target basins from gauging points published in Database of Dams (http://mudam.nilim.go.jp/home) and Water Information System (http://www1.river.go.jp/). In order to eliminate anthropogenic impacts on flow regimes, we searched basins with no dams and hydropower weirs upstream using dam almanac (Japan Dam Foundation, 2012) and map of Geospatial Information Authority of Japan (https://maps.gsi.go.jp/). We also set that catchment size was larger than 5 km^2 and the period of discharge records were longer than 5 years as the additional criteria for the selection. Finally, 448 target basins were selected throughout Japan (Fig. 1). The overview of the target basins shown in Table 1.

3. DATASET

The ANNs need *Q* characteristics and basin characteristics. These are prepared as below.

3.1 Q characteristics

In this study, we set the *Q* characteristics annually applying the unit of runoff height. The *Q* characteristics were consisted of mean annual runoff height (QMEAN) and daily flow percentiles in the flow duration curve (Q1, Q5, Q10, Q20, Q30, Q40, Q50, Q60, Q70, Q80, Q90, Q95, Q99). For example, Q20 means a 20 percentile value in the flow duration curve. Since we set these characteristics annually, the total number of data in one characteristic showed 7812.

3.2 Basin characteristics

We set the 175 basin characteristics relating to climate, land use, geology, soil, and topography (Table 2). The data source and the resolution are shown in Table 2. Since these data sources were expressed by GIS, we integrated the basin characteristics using ArcGIS 10.6, and then extracted them from a coverage of each basin. The resolutions of basin characteristics were adjusted by using nearest neighbor interpolation so that we can treat the characteristics even in different resolutions (Table 2). The indices of precipitation and snow were originated from daily data showing more than 1 (mm/d). We applied APHRO_JP (Kamiguchi et al., 2010) which was an observed rainfall database throughout Japan to the data source of precipitation indices. The snow indices were applied to Agro-Meteorological Grid Square Data (Kominami et al., 2015) which was calculated by snow water equivalent model considering heat and radiation balance. While the indices of precipitation and snow considered the time variability corresponding to the period of discharge records, other indices were set as a constant value at the basin. Although aridity index was calculated by dividing mean annual precipitation by annual potential evaporation, mean annual precipitation based on APHRO_JP has a time variability. Thus, we applied the long-term averaged value originated from National Land Numerical

Table 2. Basin characteristics.

	Description	Unit	Data source	Time variability	Calculation	Resolution
Climate	Mean annual precipitation	mm/y	APHRO_JP (Kamiguchi et al., 2010)	+		180s
	Mean daily precipitation	mm	APHRO_JP (Kamiguchi et al., 2010)	+		180s
	Mean frequency of daily precipitation	1/d	APHRO_JP (Kamiguchi et al., 2010)	+		180s
	Mean annual maximum daily precipitaion	mm	APHRO_JP (Kamiguchi et al., 2010)	+		180s
	Precipitation seasonality	-	APHRO_JP (Kamiguchi et al., 2010)	+	Calculated following Beck et al. (2015).	180s
	Mean annual snow water equivalent	mm/y	Agro-Meteorological Grid Square Data (Kominami et al., 2015)	+		~1km
	Mean daily snow water equivalent	mm	Agro-Meteorological Grid Square Data (Kominami et al., 2015)	+		~1km
	Mean frequency of daily snowfall	1/d	Agro-Meteorological Grid Square Data (Kominami et al., 2015)	+		~1km
	Mean annual potential evaporation	mm/y	CGIAR-CSI Global-Aridity and Global-PET Database (Zomer et al., 2007)			~1km
	Potential evaporation seasonality	-	CGIAR-CSI Global-Aridity and Global-PET Database (Zomer et al., 2007)		Calculated following Beck et al. (2015).	~1km
	Aridity index	-	National Land Numerical Information (G02)		(Mean annual precipitation)/(Mean annual potential evaporation)	~1km
	Mean temperature	℃	National Land Numerical Information (G02)			~1km
	Maximum temperature	℃	National Land Numerical Information (G02)			~1km
	Minimum temperature	℃	National Land Numerical Information (G02)			~1km
	Amount of global solar radiation	MJ/m^2	National Land Numerical Information (G02)			~1km
	Sunshine duration	h	National Land Numerical Information (G02)			~1km
Land use	Land use classification of National Land Numerical Information (11 groups)	%	National Land Numerical Information (L03-b_r)			~100m
Geology and soils	Large classification of surface geology (7 groups)	%	National Land Numerical Information (G05_003)		Classified following Yokoo and Oki (2009).	~1km
	Geological time (6 groups)	%	National Land Numerical Information (G05_003)		Classified following Yokoo and Oki (2009).	~1km
	Geological classification of Mushiake et al. (1981) (7 groups)	%	National Land Numerical Information (G05_003)		Classified following Mushiake et al. (1981).	~1km
	Soil classification of National Land Numerical Information (79 groups)	%	National Land Numerical Information (G05_004)			~1km
	Soil classification of Yokoo et al. (2001) (3 groups)	%	National Land Numerical Information (G05_004)		Classified following Yokoo et al. (2001).	~1km
Topography	Topographical classification of National Land Numerical Information (40 groups)	%	National Land Numerical Information (G05_002)			3s
	River length	m	HydroSHEDS (Lehner et al., 2008)			3s
	Maximum elevation	m	HydroSHEDS (Lehner et al., 2008)			3s
	Minimum elevation	m	HydroSHEDS (Lehner et al., 2008)			3s
	Catchment size	km^2	HydroSHEDS (Lehner et al., 2008)			3s
	Slope	-	HydroSHEDS (Lehner et al., 2008)			3s
	Runoff recession parameter of Biswal and Marani (2014)	-	HydroSHEDS (Lehner et al., 2008)		Calculated following Biswal and Marani (2010)	3s

Information (http://nlftp.mlit.go.jp/ksj-e/index.html) to the mean annual precipitation in the calculation. The database of National Land Numerical Information was heavily used to calculate the indices of climate, land use, geology, soil, and topography (Table 2). The references to classify soils and geology were shown in Table 2. Biswal and Marani (2010) developed a framework to estimate a runoff recession parameter based on the river network. The runoff recession parameter was estimated by analyzing the elevation data in the target basin (Biswal and Marani, 2010), which was added as a basin characteristic

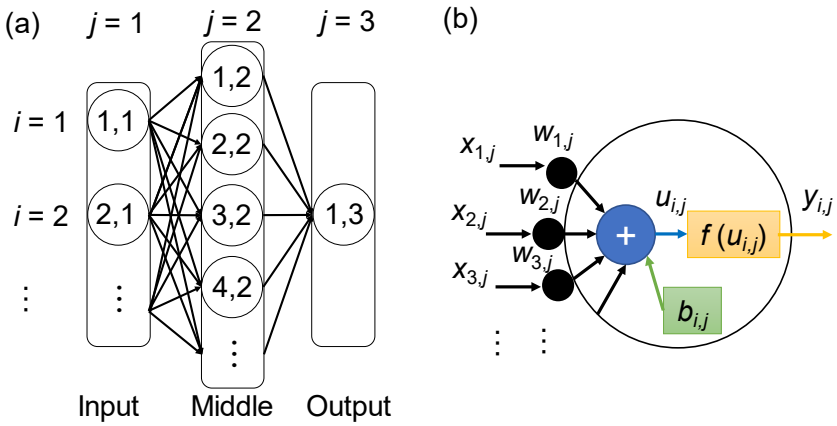

Figure 2. Schematic diagram of (a) structure and (b) function in ANN.

4. METHODS

4.1 Development of ANNs

4.1.1 Basic design

In this study, we developed ANNs to estimate Q characteristics having a middle layer (Fig. 2(a)) as with previous studies (Beck et al., 2015, Barbarossa et al., 2018). Neurons in the ANNs were distinguished by the location of neuron i and layer j (Fig. 2(a)). The input layer received 175 basin characteristics and the output layer served a Q characteristic. We set 10 neurons in the middle layer. Generally, ANNs can obtain outputs through forward propagation from input layer to output layer. In this process, a following equation is applied:

$$y_{i,j} = f(u_{i,j}) = f(b_{i,j} + \sum_{i=1}^{l} x_{i,j} w_{i,j}) \tag{1}$$

where $y_{i,j}$ is the output of the neuron (i,j), $f(u_{i,j})$ is the activating function, l is the number of neurons at the upper layer, $x_{i,j}$ is the input of the neuron (i,j), and $w_{i,j}$ and $b_{i,j}$ express the weight and bias respectively. The schematic diagram of Eq. (1) is shown in Fig. 2(b). In this study, we employed rectified linear function $f_{mid}(u_{i,j})$ in the middle layer and identity function $f_{out}(u_{i,j})$ in the output layer as the activating function respectively. These equations are expressed as below:

$$f_{mid}(u_{i,j}) = \max(u_{i,j}, 0), \tag{2}$$

$$f_{out}(u_{i,j}) = u_{i,j}. \tag{3}$$

We also applied normalization to the basin characteristics using following equation:

$$\hat{v} = (v - v_{\min})/(v_{\max} - v_{\min}) \tag{4}$$

where v is a target index, v_{min} and v_{max} are the maximum and the minimum of v respectively, and $\hat{v}$ is the normalized v.

4.1.2 Learning method

Learning of ANNs is achieved by optimizing $w_{i,j}$ and $b_{i,j}$ so that the output error is minimized. First, we need to get the error through backward propagation from the output layer to the input layer. In this study, we employed following squared error as the error function:

$$E = \frac{1}{2}(y_{1,3} - T)^2 \tag{5}$$

where E is the error at the output layer, $y_{1,3}$ is the output value, and T is the true value for $y_{1,3}$. Next, we updated $w_{i,j}$ and $b_{i,j}$ by gradient descent method. The equation is expressed as below:

$$w_{i,j}^{(t)} = w_{i,j}^{(t)} - \eta_{i,j} \frac{\partial E}{\partial w_{i,j}^{(t)}} \tag{6}$$

where t is the number of learning steps, and $\eta_{i,j}$ is a learning coefficient. The gradient $\partial E / \partial w_{i,j}^{(t)}$ in Eq. (6) was obtained by backpropagation algorithm (Rumelhart et al., 1986). The update method for $b_{i,j}$ is same as $w_{i,j}$. In fact, we applied stochastic gradient descent (LeCun et al., 1998) to optimize the learning process.

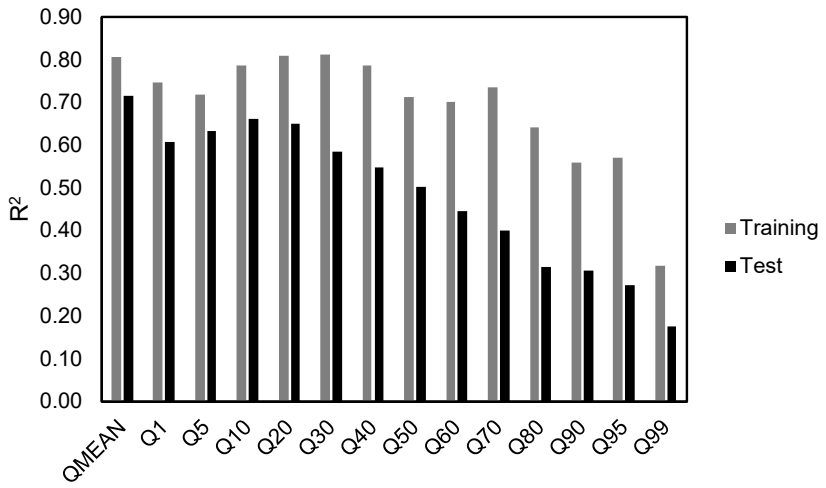

Figure 3. Results of 5-hold cross validation.

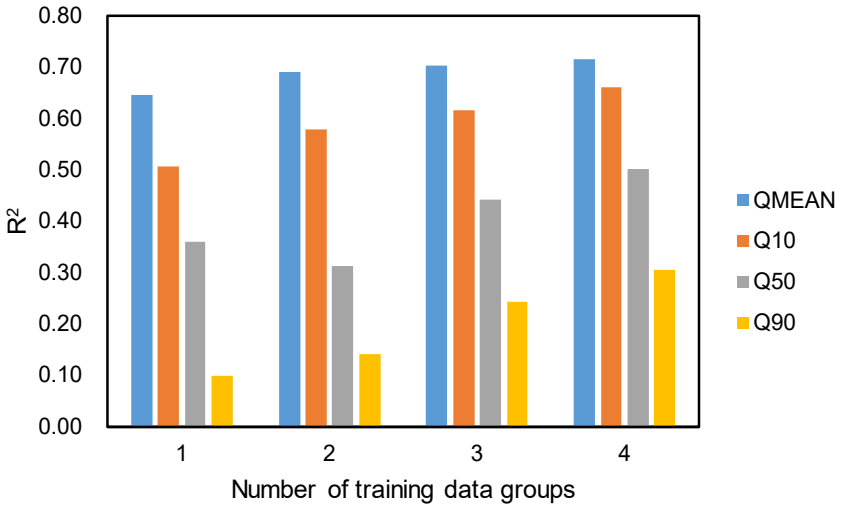

Figure 4. Results of the impact of the number of training data on the performance.

4.2 Validation of ANNs

In this study, we conducted 5-hold cross validation to evaluate the accuracy and the generalization performance of ANNs. First, we separated all target basins (i.e., 448) into 5 groups. Next, we set 1 group as the test data and set the other 4 groups as the training data. After the learning process, we calculated the determination coefficient R^2 between true values and output values for training data groups to evaluate the training accuracy. In order to evaluate the generalization performance of ANNs, we calculated the R^2 between true values and output values for test data group. This validation step was repeated 5 times to include the all target basins as the test data group. Finally, we averaged these R^2. Furthermore, we evaluated the impact of the number of training data on the performance through changing the number of training data groups from 1 to 4.

4.3 Rank correlation analysis

Generally, it is difficult to understand a relationship between the input values and the output values in ANNs. Thus, we conducted Spearman's rank correlation analysis between the Q characteristics and the basin characteristics.

5. RESULTS

The results of 5-hold cross validation are shown in Fig. 3. The all R^2 for training data showed higher than that of test data (Fig. 3). The highest performance of test data was shown in QMEAN (R^2 = 0.72). According to the results from Q1 to Q99, there was a peak of the performance in Q10 (R^2 = 0.66). In addition, the performance decreased remarkably in low flow regimes (e.g., Q99: R^2 = 0.18). We also found the performance increasing with the number of training data groups increasing (Fig. 4). Note that we focused on the results of QMEAN, Q10, Q50, and Q90 for simplicity from here (i.e., Fig. 4, Table 3 - 6).

Table 3. Top 10 basin characteristics for Spearman' rank correlation coefficient ($p < 0.001$) in QMEAN.

Description	ρ
Mean annual precipitation	0.76
Aridity index	-0.74
Mean frequency of daily precipitation	0.62
Agricultural land exept rice paddy	-0.47
Podozolic soil (dry)	0.44
Mean daily snow water equivalent	0.43
Mean daily precipitation	0.41
Forest area	0.39
Rocky area	0.37
Podozolic soil (wet)	0.35

Table 4. Top 10 basin characteristics for Spearman' rank correlation coefficient ($p < 0.001$) in Q10.

Description	ρ
Mean annual precipitation	0.70
Mean daily precipitation	-0.70
Mean annual maximum daily precipitaion	0.67
Aridity index	-0.50
Potential evaporation seasonality	0.46
Minimum temperature	0.38
Mean temperature	0.37
Agricultural land exept rice paddy	-0.37
Brown forest soil (Andosols)	-0.36
Andosols (type A)	-0.36

Table 5. Top 10 basin characteristics for Spearman' rank correlation coefficient ($p < 0.001$) in Q50.

Description	ρ
Aridity index	-0.68
Mean annual precipitation	0.63
Mean frequency of daily precipitation	0.59
Podozolic soil (dry)	0.48
Mean daily snow water equivalent	0.44
Mean annual snow water equivalent	0.42
Sunshine duration	-0.39
Podozolic soil (wet)	0.39
Agricultural land exept rice paddy	-0.37
Rocky area	0.37

Table 6. Top 10 basin characteristics for Spearman' rank correlation coefficient ($p < 0.001$) in Q90.

Description	ρ
Aridity index	-0.55
Mean annual precipitation	0.49
Mean frequency of daily precipitation	0.42
Podozolic soil (dry)	0.42
Podozolic soil (wet)	0.37
Rocky area	0.34
Mean daily snow water equivalent	0.33
Volcanic rock	0.33
Plutonic rock	-0.32
Mean annual snow water equivalent	0.31

The rank correlation coefficients between the Q characteristics and basin characteristics are shown in Table 3 – 6. Note that these rank correlation coefficients were described up to the top 10 showing high correlation (Table 3 – 6). According to Table 3 – 6, the precipitation indices and aridity index showed high correlation for the all Q characteristics. In addition, trends of the rank correlation were similar to the performance of ANNs (Fig. 3).

6. DISCUSSIONS

Beck et al. (2015) estimated QMEAN with $R^2 = 0.88$. Barbarossa et al. (2018) also estimated mean monthly streamflow with $R^2 = 0.91$. Compare to these studies, the results of this study showed a bit low performance for QMEAN ($R^2 = 0.72$). The difference in these performances may derive from the difference in the number of target basins. Although there was 448 target basins in this study, Beck et al. (2015) and Barbarossa et al. (2018) applied 4,079 and 6,600 basins respectively. In addition, the scale of catchment size in this study was different from these previous studies. Since this study targeted on only Japanese basins with no dams and hydropower intakes upstream, the median basin size was 72.3 km^2 (Table 1). On the other hand, the range of basin size in Beck et al. (2015) showed from 10 to 10,000 km^2 in global scale. Although Barbarossa et al. (2018) did not described the basin size in detail, they also targeted on global scale basins. Parajka et al. (2013) investigated on a relationship between hydrological model performances in ungauged basins and the basin scales, and then found a clear pattern of an increase of the performance with the basin scale. This pattern fit naturally into the difference in the performance between this study and previous studies (Beck et al., 2015, Barbarossa et al., 2018). Parajka et al. (2013) also suggested the following two reasons for the pattern. First, there is a trend for an increasing number of raingauges within a catchment with the basin size increasing. The second is the aggregation effect of runoff. As the basin scale increases, some of the hydrological variability is averaged out due to an interplay of space–time scale processes, which will improve hydrological simulation. In addition, we observed a high impact of agricultural land use on QMEAN (Table 3). This anthropogenic impact may make a difficulty to estimate QMEAN. Furthermore, there was also a snow effect on QMEAN (Table 3). Although Beck et al. (2015) conducted some regression analysis between Q characteristics and basin characteristics, QMEAN in Beck et al. (2015) did not show a significant correlation with their snow index. Incidentally, Barbarossa et al (2018) did not apply any snow indices to their ANNs. It is widely known that snow particles are under-caught at raingauges due to wind (Wolff et al., 2015). The snow under-catch can obstruct to know true water balance. Therefore, QMEAN in this study has a possibility to have received the impact of snow under-catch.

Yokoo and Oki (2010) investigated on a relationship between Q characteristics and basin characteristics for 14 basins (100 km^2 >) in Japan. They found that climatic impacts on Q decreased as the flow regime became low. This trend was consistent with this study (Table 3 - 6) and Beck et al (2015). This fact indicates a difficulty to estimate low flow regime thorough using only climatic indices. Musiake et al. (1981) revealed a geological impact on low flow regime in Japan. Yokoo and Oki (2010) also observed that the lowest flow index significantly correlated with geological indices, whose correlation coefficients were higher than that of climatic indices. Although Q90 in this study correlated with some geological indices (Table 6), the correlation coefficients were lower than that of climatic indices. Note that Yokoo and Oki (2010) selected their target basins in no snow area to eliminate effects of snowmelt on flow regimes. Actually, Japan is one of the heaviest snowfall countries in the world (Kazama et al., 2008). Thus, we can see some effects of snow indices even on Q90 (Table 6). This fact indicates that extreme snowmelt contributions obstruct to detect geological effects on low flow regimes. Nevertheless, we found that the accuracy of Q90 had the highest sensitivity for increasing training data. This will be a hope to improve the performance in low flow regimes.

7. CONCLUSIONS

We developed ANNs to estimate Q characteristics in absent of discharge data inputting basin characteristics (e.g., climate, topography, land use) throughout Japan. The performance was the highest in QMEAN ($R^2 = 0.72$) and the lowest in Q99 ($R^2 = 0.18$). We found that the performance of the QMEAN showed a bit lower than that of previous studies (Beck et al., 2015; Barbarossa et al., 2018). The basin scale may cause this fact since a large basin scale increases raingauge density and averages out some of the hydrological variability. In addition, our results indicated that the impacts of agricultural land use and snow under-catch made difficulties to estimate QMEAN. We also evaluated relationships between Q characteristics and basin characteristics. The result showed precipitation and aridity index related with all Q characteristics dominantly. In addition, these climatic impacts became small in low flow regimes. Although some previous studies (Musiake et al., 1981; Yokoo and Oki, 2010) found clear geological effects in low flow regimes, we observed the limited trend. Given there was a snow effect in Q90 and Japan is one of the heaviest snowfall countries, we can point out that extreme snowmelt contributions obstructs to detect geological effects on low flow regimes. Nevertheless, we found that the performance of Q90 had the highest sensitivity for increasing training data, which indicates the potential for the improvement in low flow regimes.

REFERENCES

Barbarossa, V., Huijbregts, M. A. J., Beusen, A. H. W., Beck, H. E., King, H., and Schipper, A. M. (2018). Erratum: FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. *Scientific Data*, *5*(March), 180078. https://doi.org/10.1038/sdata.2018.78

Beck, H. E., de Roo, A., and van Dijk, A. I. J. M. (2015). Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology*, *16*(4), 1478–1501. https://doi.org/10.1175/JHM-D-14-0155.1

Biswal, B., and Marani, M. (2010). Geomorphological origin of recession curves. *Geophysical Research Letters*, *37*(24), 1–5. https://doi.org/10.1029/2010GL045415

Japan Dam Foundation (2012). Dam almanac 2012, Japan Dam Foundation, Japan (in Japanese).

Kamiguchi, K., Arakawa, O., Kitoh, A., Yatagai, A., Hamada, A., and Yasutomi, N. (2010). Development of APHRO_JP, the first Japanese high-resolution daily precipitation product for more than 100 years. *Hydrological Research Letters*, *4*, 60–64. https://doi.org/10.3178/hrl.4.60

Kazama, S. ans Sawamoto, M. (1997). Water balance in a heavy snow region. Snow Engineering : Recent Advances, Balkema, pp. 543–548.

Kazama, S., Izumi, H., Sarukkalige, P. R., Nasu, T., Sawamoto, M. (2008). Estimating snow distribution over a large area and its application for water resources. *Hydrological Processes*, 22, 2315–2324. doi: 10.1002/hyp.6826.

Kominami, Y., Hirota, T., Inoue, S., and Ono, H. (2015). Development of snow water equivalent estimation model for mesh agricultural meteorological data. *Seppyo*, 77(3), pp.233-246 (in Japanese).

LeCun, Y., Bottou, L., Genevieve, B. O. and Müller, K. L. (1998) Efficient BackProp. *Neural Networks: Tricks of the Trade*, pp 9-48.

Lehner, B., Verdin, K., Jarvis, A. (2008). New global hydrograhy derived from space borne elevation data. *Eos, Transactions American Geophysical Union*, 89(10), pp.93-94.

Milly, P. C. D. (2001). A minimalist probabilistic description of root zone soil water. *Water Resources Research*, *37*(3), 457–463. https://doi.org/10.1029/2000WR900337

Musiake, K., Takahasi, Y., and Ando, Y. (1981). Effects of basin geology on river-flow regime in mountainous areas of Japan. *Proceedings of the Japan Society of Civil Engineers*, 309, pp. 51–62 (in Japanese).

Rumelhart, D. E., Hinton, G. and Wiliams, R. J. (1986). Learning representations by back-propagating errors, *Nature*, 323, pp.533-536.

Takahasi, Y., Ando, Y., Ito, T., and Ito, K. (1983). Study on base flow recessions in mountainous basins. Proceedings of Japan Society of Civil Engineers, 337, pp. 75-82 (in Japanese).

Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., Sivapalan, M., and Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins-Part 1: Runoff-hydrograph studies. *Hydrology and Earth System Sciences*, *17*(5), 1783–1795. https://doi.org/10.5194/hess-17-1783-2013

Porporato, A., Daly, E., and Rodriguez-Iturbe, I. (2004). Soil water balance and ecosystem response to climate change. *American Naturalist*, *164*(5), 625–632. https://doi.org/10.1086/424970

Price, K. (2011). Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review. *Progress in Physical Geography*, *35*(4), 465–492. https://doi.org/10.1177/0309133311402714

Wolff, M. A., Isaksen, K., Petersen-ØVerleir, A., Ødemark, K., Reitan, T., and Brækkan, R. (2015). Derivation of a new continuous adjustment function for correcting wind-induced loss of solid precipitation: Results of a Norwegian field study. *Hydrology and Earth System Sciences*, *19*(2), 951–967. https://doi.org/10.5194/hess-19-951-2015

Yokoo, Y. and Oki, T. (2010). Effects of climate, topography, soil, geology and land use on flow regimes in Japanese mountainous watersheds. *Annual Journal of Hydraulic Engineering*, 54, pp. 461-474 (in Japanese).

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., and Briggs, P. R. (2004). A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, *40*(2), 1–14. https://doi.org/10.1029/2003WR002710

Zomer, R. J., Bossio, D. A., Trabucco, A., Yuanjie, L., Gupta, D. C. and Singh, V. P. (2007). Tree sand water: small holder agroforestry of irrigated lands in northern India, Colombo, SriLanka. International *Water Management Institute*, 45.