ELSEVIER

Journal of Hydrology 246 (2001) 209–222

Journal of Hydrology

www.elsevier.com/locate/jhydrol

Regionalization of lumped water balance model parameters based on multiple regression

Y. Yokoo[a,*], S. Kazama[a], M. Sawamoto[a], H. Nishimura[b]

[a]Department of Civil Engineering, Graduate school of Engineering, Tohoku University, Aoba-yama 06, Aoba-ku, Sendai 980-8579, Japan
[b]Institute of Engineering Mechanics and Systems, University of Tsukuba, Japan

Received 18 July 2000; revised 26 January 2001; accepted 5 March 2001

Abstract

This study establishes a method for evaluating the coefficients of tank model. First, the model coefficients were optimized using the Standardized Powell Method at 12 watersheds. Then 16 characteristics were derived from geographical information on topography, soil type, geology, and land-use of the basins. Finally, a multiple linear regression model was applied to the relationship between the model coefficients and the basin characteristics. Trial application of the regression equations worked successfully at two different watersheds, suggesting that the coefficients of the tank model could be evaluated based solely on the geographical characteristics of the basin. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Dam catchment; Dairy runoff simulation; Base-flow; Parameter evaluation; Permeability

1. Introduction

Various runoff models have been proposed and used to predict the runoff of a river for water utilization and prevention of flood disasters. Every runoff model can be classified into lumped models and distributed models. The lumped models have a single conceptual model for a watershed, and the distributed models divide a watershed into sub-areas providing a model to them. Both type of the runoff models have been constructed based on both data sets of observed rainfall and runoff, but the collection of these data is not easy except for restricted watersheds such as dam basins. This problem always limits the number of catchments to be analyzed with runoff models.

In the case that there is no observed data, distributed models can be constructed based on geographical characteristics of the watersheds, but these are generally useful only for short-term runoff because they require extensive computer resources to run. Although this limitation implies that lumped runoff models should be used in simulating long-term runoff, lumped models cannot be constructed based on the geographical characteristics of the watersheds.

The coefficients of lumped runoff models are usually optimized based on observed data. The optimized coefficients not only characterize the model structure, but also reflect geographical features of a watershed, because a set of model coefficients has characteristic values in each watershed. Ishihara and Kobatake (1979) tried to evaluate the coefficients of the tank model, one of the lumped models proposed by Sugawara (1995), based on the geographical characteristics of the study areas, but all of the studies dealt with only short-term runoff.

---

* Corresponding author.
E-mail address: yokoo@kaigan.civil.tohoku.ac.jp (Y. Yokoo).

PII: S0022-1694(01)00372-9

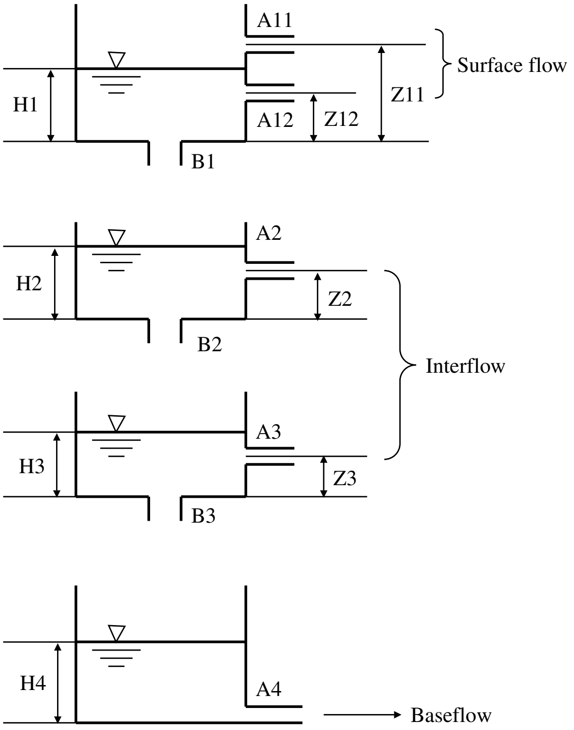

Fig. 1. Structure of tank model used in this study. *A* (1/day) represents the coefficient of runoff hole that is proportional to the runoff from each tank. *B* (1/day) is the coefficient of infiltration hole that is in proportion to the infiltration to the next tank. *Z* (mm) is the height of the runoff hole from each tank bottom in the four tanks, and *H* (mm) the storage height in a tank.

In this study, we propose constructing method of lumped runoff models based solely on geographical watershed features and attempt to simulate long-term runoff. The tank model was employed because the performance of the model was better than other lumped models. This paper reports the result of statistical investigation on the relationships between optimized coefficients of the tank model and geographical characteristics of 12 study areas in Japan.

2. Tank model

2.1. Overview

The tank model, originally proposed by Sugawara (1995), has been widely applied for runoff analysis and is considered a lumped runoff model. It is well known because of its simple structure, easy runoff calculation, and better performance in simulating runoff than other models. The tank model has many model coefficients that must be evaluated by trial and error. Recent advancements in computer technology have enabled easy optimization of the model coefficients, removing one drawback of this model.

2.2. Model structure

The tank model employed here consists of a series of four tanks (Fig. 1), because Sugawara (1972) often used four tank models for daily-runoff analyses and obtained good results in many Japanese watersheds. Referring the model coefficients evaluated by Sugawara analysis, we can compare the magnitude of model coefficients between our model and Sugawara's models.

In four tanks, the upper two tanks model rapid runoff near the ground surface, whereas the lower two tanks model delayed runoff that is far from the ground surface. In Fig. 1, *A* (1/day) represents the coefficient of runoff hole that is proportional to the runoff from each tank. *B* (1/day) is the coefficient of infiltration hole that is in proportion to the infiltration to the next tank. *Z* (mm) is the height of the runoff hole from each tank bottom in the four tanks, and *H* (mm) the storage height in a tank. Generally, *A*, *B*, and *Z* are dealt as 12 model coefficients in case of four tank model, because they are constant in runoff calculation. *H* is variable in runoff calculation and it is not dealt as tank model coefficients.

2.3. Runoff calculation

Dairy runoff quantities are calculated by the equations shown below.

$$R(x,n) = \begin{cases} A(x)\cdot(H(x,n) - Z(x)) & (H(x,n) > Z(x)) \\ 0 & (H(x,n) \le Z(x)) \end{cases} \tag{1}$$

$$I(x,n) = B(x)\cdot H(x,n) \tag{2}$$

$$H(x,n+1) = \begin{cases} H(x,n) - R(x,n)\cdot\Delta t - I(x,n)\cdot\Delta t + P(n+1)\cdot\Delta t & (x=1) \\ H(x,n) - R(x,n)\cdot\Delta t - I(x,n)\cdot\Delta t + I(x-1,n)\cdot\Delta t & (x \neq 1) \end{cases} \quad (3)$$

$$Q(n) = \sum_{X=1}^{4} R(x,n) \quad (4)$$

where

x:	number of tank counted from top
n;	number of days counted from the beginning of the runoff calculation (1/d)
Δt;	length of time step of runoff calculations
$A(x)$;	runoff coefficients of x-th tank (1/d)
$B(x)$;	infiltration coefficients of x-th tank (1/d)
$H(x,n)$;	water depth in x-th tank at n-th day (mm)
$I(x,n)$;	infiltration from x-th thank at n-th day (mm/d)
$P(n)$;	precipitation at n-th day (mm/d)
$Q(n)$;	total runoff of a basin at n-th day (mm/d)
$R(x,n)$;	runoff from x-th tank at n-th day (mm/d)
$Z(x)$;	height of runoff hole of x-th tank (mm).

3. Study areas and data set

3.1. Study area

The 12 catchments, used in this study, are located upstream of Hitokura dam, Iwase dam, Kotogawa dam, Midorikawa dam, Mukunashi dam, Nagase dam, Shimouke dam, Shintoyone dam, Shourenji dam, Tsubayama dam, Tsuruta dam, and Yubara dam (Fig. 2). These basins were selected because they have readily available and continuous data on rainfall and runoff, and they met the following conditions.

- Dam catchment that contains no other dams;
- Drainage area grater than 100 km^2;
- Little effect from snow.

The first condition was set to avoid the effects of other dams in the upper reaches of the catchment. The second condition was needed to take various kinds of land-use into each basin. The last condition allows for focus on the rainfall–runoff processes. General information on the study areas is shown in Table 1.

3.2. Hydrological data

Hydrological data used were daily rainfall and runoff from 1990 to 1992 at the dam basins obtained from ‘An Annual Report of the Management of the Multiple-Purpose Reservoir’ (Development Section, River Bureau, Ministry of Construction, Japan, 1994, 1995, 1996). The runoff data were calculated by the change of the water levels in the reservoirs. The rainfall data was obtained from rain gauges near the dams.

In general, the data show that rainfall volume tends to increase with elevation in the study areas, so that recorded data at a dam might be smaller than the net rainfall of the basin it impounds. Despite this underestimate, the rainfall measured at each dam site was treated as effective rainfall over the entire watershed. In the absence of detailed rainfall measurements, this probably underestimated rainfall can be considered acceptable because runoff ratios normally range from 0.5 to 0.7 in Japan.

4. Model construction using optimization theory

4.1. Optimization theory

We used a method developed by Powell (1964) to construct the tank model. This method is an efficient method to find the minimum of a function of several variables without calculating derivatives, and is called as a conjugate direction method.

By applying the Powell method to the construction of the tank model, the users meet two merits. One of the merits is that the users can optimize the model coefficients without trial and error, and the other is that they can obtain the four initial storage heights in the model and runoff can be simulated well soon after the beginning of the runoff calculation. The second merit is obtained because the Powell method allows dealing the 12 model coefficients and four initial storage heights in the tank model as 16 parameters to be optimized. The 16 parameters are optimized by calculating runoff through Eqs. (1)–(4)

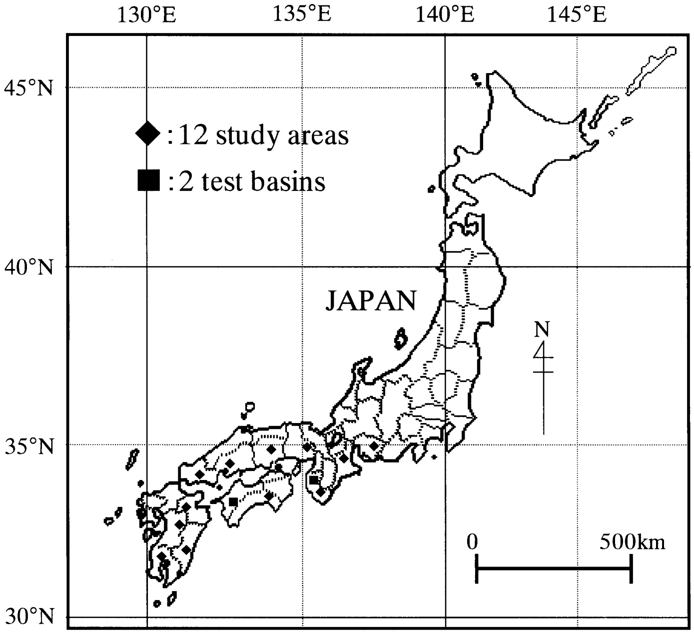

Fig. 2. Location of 12 study areas and two test basins. All of the areas were located in southern Japan because they are selected to have little effect of snowfall in the areas.

and minimizing the value RE in Eq. (5).

$$RE = \frac{1}{N}\sum_{1}^{N}\frac{Qci - Qoi}{Qoi} \tag{5}$$

where

RE;	mean rerative error
N;	number of days during the analysis (d)
Qci;	calculated runoff by Eqs. (1)–(4) (mm/d)
Qoi;	observed runoff (mm/d)

The Powell method does not effectively work in applying it directly to the evaluation of the tank model coefficients, because H4 (the initial storage height of the 4th tank) has the value of 10^2 order and A4 (the runoff coefficient of the 4th tank) has the value of 10^{-3} order causing a difference of as much as 10^5 order. Kadoya and Nagai (1980) suggested a method to standardize the 16 parameters by their initial values to solve this problem. Their method is called the Standardized Powell (SP) method and is applied in many cases, so that we employed this method in evaluating the tank model coefficients.The SP method and the Powell method work well in optimizing the coefficients when initial values of the parameters are selected close to the optimum values, but both tend to be trapped into nearly optimized values that are different from the theoretically optimum values. We solved this problem by preparing two sets of initial parameters before the optimizations. The two sets of initial parameters, named ‘initial parameters 1’ and ‘initial parameters 2’, are shown in Table 2. The two sets of parameters were selected depending on the geological features of each basin. The first set, ‘initial parameters 1’, was used for basins with high vertical permeability such as volcanic rocks or coarse-grained sediments. The second one, ‘initial parameters 2’, was used for basins with low vertical permeability such as plutonic rocks or metamorphic rocks.

4.2. Result of optimization

Using the SP method and the two types of initial parameters, tank model coefficients were evaluated for each year in each basin, for a total of 36 runs. The evaluation of the model coefficients for each year was made because it was better to maximize

Table 1
General information of 12 study areas. These are dam catchemts, which are greater than 100 km^2 and have no other dams and little effect of snowfall

Name of dam basin	Effective storage capacity (1000 m^3)	Basin area (km^2)	Representative gradient
Hitokura	30800	115.1	0.04894
Iwase	41000	354.0	0.05359
Kotogawa	23042	324.0	0.02762
Midorikawa	35200	359.0	0.06269
Mukunashi	6270	160.0	0.03354
Nagase	41470	295.2	0.08467
Shimouke	52300	185.0	0.06600
Shintoyone	40400	136.3	0.05797
Shourenji	23800	100.0	0.05309
Tsubayama	39500	396.5	0.04460
Tsuruta	77500	805.0	0.03552
Yubara	86000	255.0	0.04350

the number of samples for the analysis of multiple linear regression later on. In optimizing the model coefficients, we made a standard that model coefficients were optimized when the coefficients of correlation became greater than 0.8 between calculated runoff and observed runoff.

Runoff calculations were made using 36 sets of optimized coefficients to draw a hydrograph for each set of model coefficients. Figs. 3 and 4 show the examples of the 36 hydrographs. In the figures, ‘grad.’ is the regression coefficients of simulated runoff to observed runoff, and ‘R^2’ the coefficients of determination. Obviously, base-flow was simulated well on the whole when ‘grad.’ had a large value, while peak runoff was simulated well when ‘R^2’ has a large value. The target of this study was long-term runoff,

Table 2
Two sets of initial parameters for the optimization of tank model coefficients. ‘initial parameters 1’ was used for basins with high vertical permeability such as volcanic rocks or coarse-grained sediments, and ‘initial parameters 2’ was used for basins with low vertical permeability such as plutonic rocks or metamorphic rocks

Vertical permeability	Initial parameters 1 (high)	Initial parameters 2 (low)
$A11$	0.400	0.400
$A12$	0.200	0.200
$B1$	0.150	0.100
$A2$	0.100	0.400
$B2$	0.050	0.050
$A3$	0.020	0.060
$B3$	0.030	0.020
$A4$	0.003	0.003
$Z11$	40.000	40.000
$Z12$	15.000	15.000
$Z2$	20.000	20.000
$Z3$	10.000	10.000
$H4$	200.000	200.000
$H3$	40.000	40.000
$H2$	2.000	2.000
$H1$	1.000	1.000

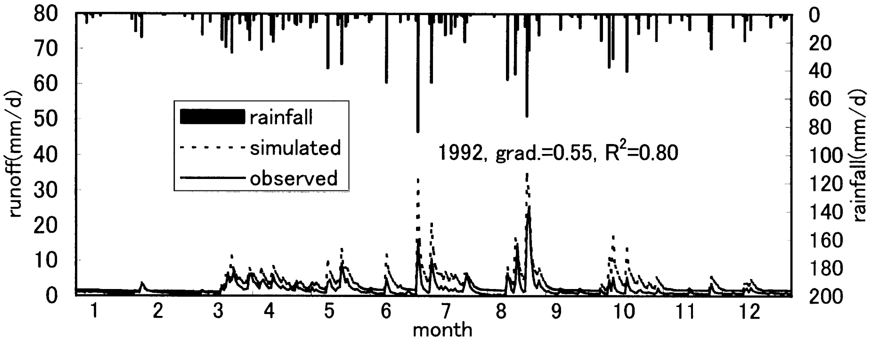

Fig. 3. Hydrograph drawn with the calculated runoff using the model coefficients optimized by SP method in Hitokura dam catchment for the data in 1992. 'grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

and we evaluated that the results with large value of 'grad.' was better than those with small value of 'grad.'. In this sense, runoff simulation in Fig. 3 was the best result and the simulation in Fig. 4 was the worst. The 36 sets of the model coefficients optimized here were used in Section 5.

5. Model construction using geographical information

5.1. Geographical information

A variety of geographical information was compiled by Geographical Survey Institute, Ministry of Construction, Japan. The data are mainly based on topographic maps, land-use maps, and geology maps. These maps were digitized and the data numerically recorded. Most of the data are systematically recorded by giving header codes that indicates latitude and longitude. In the real scale, they are recorded in a nearly squared cell of 1 km^2.

This study used three types of information closely related with rainfall–runoff phenomena: land-use, soil type, and surface geology. The subsurface geology in each watershed was not taken into consideration in this study, because subsurface geology data was not currently available for these areas, nor is it common data for other areas.

5.2. Extraction of geographical characteristics

This section focuses on geological characteristics, which can be related with tank model coefficients. Geographical characteristics of each basin were drawn from digital survey data using Eq. (6).

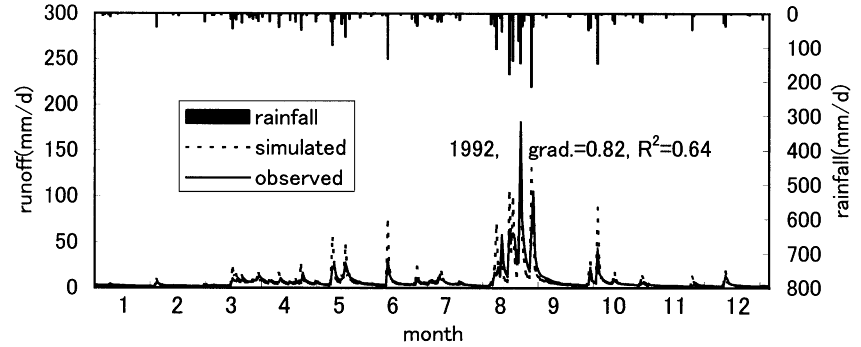

Fig. 4. Hydrograph drawn with the calculated runoff using the model coefficients optimized by SP method in Nagase dam catchment for the data in 1992. 'grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

Table 3
Classification of geology type after Nakano (1976). Italicized entries have been added to Nakano's original scheme (H, Holocene epoch; Pl, Pleistocene epoch; Tn, Tertiary period (Neogene); Tp, Tertiary period (Paleogene); M, Mesozoic era; Pa, Paleozoic era)

	Degree of contribution: high (simplified name: GTA)	Degree of contribution: medium (simplified name: GTB)	Degree of contribution: low (simplified name: GTC)
Unconsolidated deposit	Gravel (H, Pl, Tn, Tp, M, Pa)	Sand and mud (H, Pl, Tn, Tp, M, Pa)	Mud (H, Pl, Tn, Tp, M, Pa)
	Sand (H, Pl, Tn, Tp, M, Pa)	Gravel and sand and mud (H, Pl, Tn, Tp, M, Pa)	Clay (H, Pl, Tn, Tp, M, Pa)
	Gravel and sand (H, Pl, Tn, Tp, M, Pa)	Sand and mud and silt (H, Pl, Tn, Tp, M, Pa)	Peat (H, Pl, Tn, Tp, M, Pa)
	Clastics (H, Pl, Tn, Tp, M, Pa)		
Sedimentary rocks	Conglomerate (H, Pl)	Conglomerate (Tn, Tp)	Conglomerate (M, Pa)
	Sandstone (H, Pl)	Sandstone (Tn, Tp)	Sandstone (M, Pa)
	Alternation of conglomerate and sandstone (H, Pl)	Alternation of conglomerate and sandstone (Tn, Tp)	Alternation of conglomerate and sandstone (M, Pa)
	Limestone (H, Pl)	Limestone (Tn)	Limestone (Tp, M, Pa)
		Mudstone (H)	Mudstone (D, Tn, Tp, M, Pa)
		Alternation of sandstone and mudstone (H, Pl)	Alternation of sandstone and mudstone (Tn, Tp, M, Pa)
		Alternation of conglomerate, sandstone, and mudstone (H, Pl)	Alternation of conglomerate, sandstone, and mudstone (Tn, Tp, M, Pa)
		Diabasic tuff (H, Pl, Tn)	Diabasic tuff (Tp, M, Pa)
		Siliceous rock (H, Pl, Tn, Tp)	Siliceous rock (M, Pa)
Volcanic rock	Agglomerate (H, Pl)	Agglomerate (Tn, Tp)	Agglomerate (M, Pa)
	Tuff breccia (H, Pl)	Tuff breccia (Tn, Tp)	Tuff breccia (M, Pa)
	Tuff (H, Pl)	*Tuff (Tn, Tp)*	*Tuff (M, Pa)*
	Andesite (H, Pl)	Andesite (Tn, Tp)	Andesite (M, Pa)
	Welded tuff (H, Pl)	Welded tuff (Tn)	Welded tuff (Tp, M, Pa)
		Rhyolite (H, Pl)	Rhyolite (Tn, Tp, M, Pa)
		Basalt (H, Pl)	Basalt (Tn, Tp, M, Pa)
	Volcanic ash sand (H, Pl, Tn, Tp, M, Pa)		
	Pyroclastic material (H, Pl, Tn, Tp, M, Pa)		
	Pumice (H, Pl, Tn, Tp, M, Pa)		
	Pumice flow deposit (H, Pl, Tn, Tp, M, Pa)		
	Volcanic breccia (H, Pl, Tn, Tp, M, Pa)		
Plutonic rock		Porphyry (H, Pl, Tn, Tp, M, Pa)	*Diabase (H, Pl, Tn, Tp, M, Pa)*
		Gabbro (H, Pl, Tn, Tp, M, Pa)	
		Granite (H, Pl, Tn, Tp, M, Pa)	
		Diorite (H, Pl, Tn, Tp, M, Pa)	
Metamorphic rock		Crystalline schist (H, Pl, Tn, Tp)	Crystalline schist (M, Pa)
		Phyllite (H, Pl, Tn, Tp)	Phyllite (M, Pa)
		Black schist (H, Pl, Tn, Tp, M, Pa)	
		Green schist (H, Pl, Tn, Tp, M, Pa)	
		Gneiss (H, Pl, Tn, Tp, M, Pa)	
		Serpentinite (H, Pl)	*Serpentinite (Tn, Tp, M, Pa)*
		Amphibolite (H, Pl, Tn, Tp, M, Pa)	
	Mylonitic rock (H, Pl, Tn, Tp, M, Pa)		
	Hornfels (H, Pl, Tn, Tp, M, Pa)		

Table 4
Classification of soil type after Nakano (1976). Italicized entries have been added to Nakano's original scheme

Degree of contribution: high (simplified name: STA)	Degree of contribution: medium (simplified name: STB)	Degree of contribution: low (simplified name: STC)
Immature soil (the others)	Immature soil (coarse residual)	Rock land
Andosol (the others)	Andosol (wet)	Lithosol
Brown forest soil	Andosol (wet, coarse)	Residual immature soil
Brown forest soil (dark red)	Brown forest soil (yellow brown)	Brown forest soil (dry)
	Brown forest soil (red brown)	Brown forest soil (dry, yellow brown)
	Brown forest soil	Brown forest soil (dry, red brown)
	Brown forest soil (wet)	*Brown forest soil* (*dry*, *red-yellow brown*)
	Podzolic soil (dry)	*Brown forest soil* (*wet*, *yellow brown*)
		Podzolic soil (wet)
		Red soil
		Red-yellow soil
		Yellow soil
		Dark red soil
		Gley soil (fine)
		Gley soil
		Gley soil (coarse)
		Peat soil
		Black peat soil

$$\text{Basin factor} = \frac{\text{Number of cells satisfying a certain criterion}}{\text{Total number of cells in the basin}} \times 100(\%) \tag{6}$$

This basin factor shows the characteristics of surface geology, soil, and land-use by area ratio. If 'a certain criterion' is forest area, Eq. (6) derives the area ratio of forest in a watershed. A total of 16 basin factors have been identified from three types of geographical information: three on surface geology, three on soil type, eight on land-use pattern, and two from basin maps.

5.2.1. Surface geology

This study employed a geologic classification suggested by Nakano (1976). It was made to classify empirically surface geology into three groups based on the hydraulic properties of the surface geology. Nakano's classification was modified for this study because Nakano's study does not classify all the surface geology types recorded in the geographical information. Table 3 shows the resulting classification of surface geology. Rock types in the first group have high vertical permeability, whereas the second group shows intermediate permeability, and the third group low. Area ratios of the three geology types were calculated using Eq. (6) and used as 'Basin Factors' in the multiple linear regression analysis.

5.2.2. Soil

Nakano (1976) also suggested a classification for soil types. Nakano's study does not classify all the soil types recorded in the geographical information and we modified it as shown in Table 4. As with surface geology, we divided the soil types into three groups by permeability and calculated the area ratios of the three soil types by Eq. (6) and dealt as 'Basin Factors' in multiple linear regression analysis.

5.2.3. Land-use

We grouped the land-use patterns according to the classification names in geographical information. The original classification in the geographical information is shown in Table 5. The classification by their names is a popular way as introduced with runoff coefficients by Chow et al. (1988). In Table 5, some items were quite similar to each other and a simplification was made as shown in Table 6. Eq. (6) was used again to

Table 5
Original classification of land-use type in Japanese geographical information

Simplified name	Type of land-use
LU1	Rice paddy
LU2	Field
LU3	Orchard
LU4	Shrub
LU5	Forest
LU6	Wasteland
LU7	Buildings (A)
LU8	Buildings (B)
LU9	Road, railroad
LU10	Vacant/naked land
LU11	Lake
LU12	River (A)
LU13	River (B)
LU14	Coast
LU15	Sea

calculate the area ratios of the eight types of land-use in Table 6.

5.2.4. Other factors

There are some geographical features that can be easily identified even from basin maps—this study used two such factors. The first factor was 'basin area' (BA) that influences hydrological features of a basin. The other factor was 'representative gradient' (RG), which was calculated by the following procedures. First, the difference of the elevations between the maximum elevation point on basin boundary and the elevation of the dam top was initially obtained from the basin map. Then the horizontal distance was then measured between the two points. The value RG was finally calculated as the ratio of these two values.

Table 6
Simplified classification of land-use type used in this study. The number of the land-use patterns was reduced from 15 to eight

Simplified name	Type of land-use
LU1	Rice paddy
LU2	Field
LU3 and LU4	Orchard, shrub
LU5	Forest
LU6 and LU10	Vacant/waste/naked land
LU7 and LU8	Building
LU9	Road, railroad
LU11 and LU12 and LU13	River, lake, sea

5.3. Multiple linear regression analysis

It would be ideal if all of the coefficients of tank model could be estimated by simple equations using geographical character of a basin. The analysis of multiple linear regression has been aimed at achieving such description and is useful for the investigation of numerical relationships between independent and dependent variables, and it enables to estimate dependent variables using one-dimensional equations that consists of several independent variables. This study employed multiple linear regression where basin factors and model coefficients were set as independent and dependent variables, respectively. The resulting equations would be optimizing equations for tank model construction if runoff quantities could be simulated successfully using the coefficients determined by the regression equations.

5.3.1. Condition of analysis

The tank model in Fig. 1 has 12 model coefficients, which are A (the coefficients of runoff), B (the coefficients of infiltration), and Z (the height of the runoff hole from each tank bottom in the four tanks) for each of four tanks. The storage heights, H, are not considered as model coefficients, because they are variable in runoff calculation whereas A, B, and Z are constant. From geographical information, 16 basin factors were derived from surface geology, soil type, land-use, gradient, and basin area. This meant that multiple linear regression analysis could not be used because the number of independent variables was more than that of the dependent variables. In addition, equations of multiple linear regression should have few independent variables in order to apply the equations effectively. Accordingly, the number of independent variables was reduced by the following conditions:

- all basin factors were used for the coefficients in the 1st and the 2nd tank;
- all basin factors except land-use factors were used for the coefficients in the 3rd tank;
- all basin factors except land-use and soil type factor were used for the coefficients in the 4th tank;
- only independent variables having F-values greater

Table 7
Result of multiple regression analysis (1). Names of the dependent variables are at the end of the left hand sides. The partial regression coefficients (P.R.C.) and *F*-values were shown on the right side of each dependent variable for the constant term and independent variables in each regression equation

*A*11	Variable	Constant term	BA	GTC	STB		
	P.R.C	2.85E − 01	− 2.52E − 04	1.41E − 03	− 2.07E − 03		
	F-value	1.39E + 02	1.87E + 01	1.35E + 01	7.99E + 00		
*A*12	Variable	Constant term	GTC	STB	GTB		
	P.R.C	7.90E − 02	1.17E − 03	− 1.00E − 03	7.24E − 04		
	F-value	1.86E + 01	2.12E + 01	1.15E + 01	5.93 + 00		
*B*1	Variable	Constant term	STC	LU7 and LU8			
	P.R.C	1.69E − 01	− 8.31 − 04	1.13E − 02			
	F-value	4.88E + 02	3.70E + 01	7.32E + 00			
*A*2	Variable	Constant term	STC	LU9			
	P.R.C	7.16E − 02	2.49E − 03	7.02E − 01			
	F-value	1.65E + 01	3.87E + 01	9.11E + 00			
*B*2	Variable	Constant term	STB	GTC			
	P.R.C	6.95E − 02	4.62E − 04	− 2.11E − 04			
	F-value	2.29E + 02	6.16E + 00	4.68E + 00			
*A*3	Variable	Constant term	STC	RG	STA	BA	GTA
	P.R.C	5.51E − 02	7.16E − 04	− 1.04E + 00	8.06E − 04	− 7.21E − 05	− 4.00E − 04
	F-value	1.70E + 01	3.28E + 01	3.16E + 01	3.00E + 01	2.17E + 01	9.66E + 00
*B*3	Variable	Constant term	STC	STA	GTA	RG	
	P.R.C	4.30E − 02	− 3.01E − 04	− 2.60E − 04	2.13E − 04	1.51E − 01	
	F-value	1.06E + 02	3.61E + 01	2.21E + 01	1.66E + 01	6.55E + 00	
*A*4	Variable	Constant term	GTA	RG	GTC		
	P.R.C	1.45E − 03	− 9.82E − 06	9.84E − 03	4.13E − 06		
	F-value	5.11E + 01	1.02E + 01	7.45E + 00	4.68E + 00		
*Z*11	Variable	Constant term	LU6 and LU10	RG	LU1		
	P.R.C	6.34E + 01	6.59E − 01	− 2.59E + 02	− 4.76E − 01		
	F-value	1.69E + 02	1.43E + 01	1.14E + 01	6.17E + 00		
*Z*12	Variable	Constant term	LU6 and LU10	STB			
	P.R.C	1.49E + 01	2.30E − 01	2.86E − 02			
	F-value	8.52E + 02	1.47E + 01	2.64E + 00			
*Z*2	Variable	Constant term	RG				
	P.R.C	2.79E + 01	− 5.09E + 01				
	F-value	4.68E + 02	4.43E + 00				
*Z*3	Variable	Constant term	STB	GTC			
	P.R.C	1.20E + 01	4.96E − 02	− 2.01E − 02			
	F-value	6.34E + 02	6.58E + 00	3.95E + 00			
*H*4	Variable	Constant term	BA	GTA			
	P.R.C	1.51E + 02	1.89E + 00	1.19E + 01			
	F-value	1.29E + 00	2.09E + 01	8.94E + 00			

than 2.0 were used by using stepwise method;

- no more than five independent variables were used in each regression.

5.3.2. Result of analysis

The following equations were obtained from multiple linear regression analysis. Model coefficients are on the left side as dependent variables, and basin factors are on the right side as independent variables. The names of basin factors correspond to the simplified names in Tables 3, 4, and 6.

$$A11 = 2.85\times10^{-1} - 2.52\times10^{-4}\mathrm{BA} + 1.41 \times10^{-3}\mathrm{GTC} - 2.07\times10^{-3}\mathrm{STB}$$

$$A12 = 7.90\times10^{-2} + 1.17\times10^{-3}GTC - 1.00 \times10^{-3}\mathrm{STB} + 7.24\times10^{-4}\mathrm{GTB}$$

Table 8
Result of multiple linear regression analysis (2). 'R', 'R^2', and 'R^*' are the multiple correlation coefficient, the coefficient of determination, and the multiple correlation coefficient adjusted for the degrees of freedom R^*, respectively

	R	R^2	R^*
$A11$	0.703	0.494	0.668
$A12$	0.676	0.457	0.637
$B1$	0.756	0.572	0.739
$A2$	0.797	0.635	0.783
$B2$	0.432	0.186	0.370
$A3$	0.880	0.775	0.859
$B3$	0.800	0.641	0.771
$A4$	0.722	0.521	0.690
$Z11$	0.599	0.358	0.546
$Z12$	0.556	0.309	0.517
$Z2$	0.340	0.115	0.299
$Z3$	0.430	0.185	0.368
$H4$	0.768	0.589	0.751

$$B1 = 1.69 \times 10^{-1} - 8.31 \times 10^{-4}\mathrm{STC} + 1.13 \times 10^{-2}\mathrm{LU7\&LU8}$$

$$A2 = 7.16 \times 10^{-2} + 2.49 \times 10^{-3}\mathrm{STC} + 7.02 \times 10^{-1}\mathrm{LU9}$$

$$B2 = 6.95 \times 10^{-2} + 4.62 \times 10^{-4}\mathrm{STB} - 2.11 \times 10^{-4}\mathrm{GTC}$$

$$A3 = 5.51 \times 10^{-2} + 7.16 \times 10^{-4}\mathrm{STC} - 1.04 \times 10^{0}\mathrm{RG} + 8.06 \times 10^{-4}\mathrm{STA} - 7.21 \times 10^{-5}\mathrm{BA} - 4.00 \times 10^{-4}\mathrm{GTA}$$

$$B3 = 4.30 \times 10^{-2} - 3.01 \times 10^{-4}\mathrm{STC} - 2.60 \times 10^{-4}\mathrm{STA} + 2.13 \times 10^{-4}\mathrm{GTA} + 1.51 \times 10^{-1}\mathrm{RG}$$

$$A4 = 1.45 \times 10^{-3} - 9.82 \times 10^{-6}\mathrm{GTA} + 9.84 \times 10^{-3}\mathrm{RG} + 4.13 \times 10^{-6}\mathrm{GTC}$$

$$Z11 = 6.34 \times 10^{1} + 6.59 \times 10^{-1}\mathrm{LU6\&LU10} - 2.59 \times 10^{2}\mathrm{RG} - 4.76 \times 10^{-1}LU1$$

$$Z12 = 1.49 \times 10^{1} + 2.30 \times 10^{-1}\mathrm{LU6\&LU10} + 2.86 \times 10^{-2}\mathrm{STB}$$

$$Z2 = 2.79 \times 10^{1} - 5.09 \times 10^{1}\mathrm{RG}Z3 = 1.20 \times 10^{1} + 4.96 \times 10^{-2}\mathrm{STB} - 2.01 \times 10^{-2}\mathrm{GTC}$$

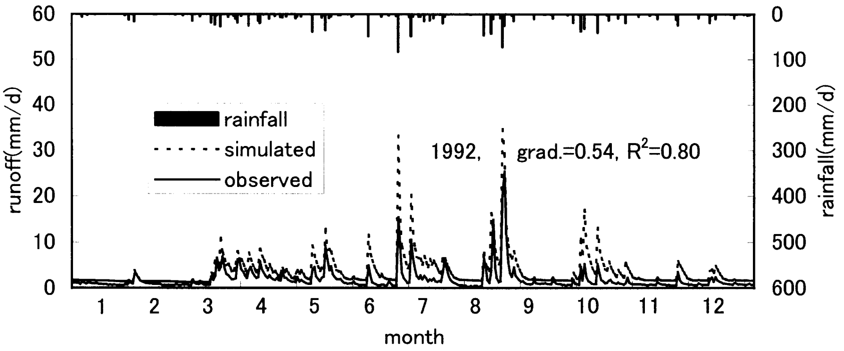

Fig. 5. Hydrograph drawn with the calculated runoff using the model coefficients evaluated using geographical information in Hitokura dam catchment for the data in 1992. 'Grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

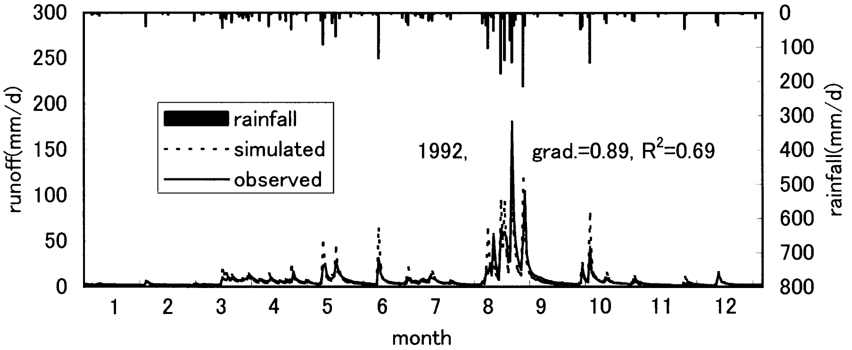

Fig. 6. Hydrograph drawn with the calculated runoff using the model coefficients evaluated using geographical information in Nagase dam catchment for the data in 1992. 'Grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

$$H4 = 1.51 \times 10^2 + 1.89 \times 10^0 \mathrm{BA} + 1.19 \times 10^1 \mathrm{GTA} (H3 = 35.0,\ H2 = 25.0,\ H1 = 15.0) \quad (7)$$

In Eq. (7), $H1$, $H2$, $H3$, and $H4$ are the initial storage heights in runoff simulation. The initial storage heights were not considered as tank model coefficients and they should not be evaluated before runoff simulation. However, $H4$ was evaluated through multiple linear regression analysis as well as tank model coefficients, because the initial storage height in the 4th tank could be strongly controlled by the geographical characteristics of the basins. It is quite useful to know the initial storage height in 4th tank before analysis, as it significantly influences the simulated base-flow. The $H1$, $H2$, and $H3$ listed in the equations were the mean initial values of 36 sets of $H1$, $H2$, and $H3$ optimized by the SP method and were to be used for the application of these equations.

F-values of the independent variables are shown in Table 7, and Table 8 gives the multiple correlation coefficient R, the coefficient of determination R^2, and the multiple correlation coefficient adjusted for the degrees of freedom R^*. The regression equations for $B2$, $Z2$ and $Z3$ had small R^*, and it was apparent that evaluations of the dependent variables did not work well for these equations. However, all the equations were treated as regression models that evaluate the tank model coefficients, because not all the tank model coefficients are sensitive to simulated runoff and the regression models could work well on the whole. Evaluation of these regression models is made in Section 5.6.

5.4. Consideration on equation of multiple linear regression

The independent variables are mathematically significant in the equations of multiple linear regression, because F-values of the independent variables were greater than 2.0. To understand the physical meaning of each equation, the relationship between

Table 9
General information of two test basins. These are dam catchemts, which are greater than 100 km^2 and have no other dams and little effect of snowfall

Name of dam basin	Effective storage capacity (1000 m^3)	Basin area (km^2)	Representative gradient
Fudagawa	19200	228.8	0.03930
Nomura	12700	168.0	0.03988

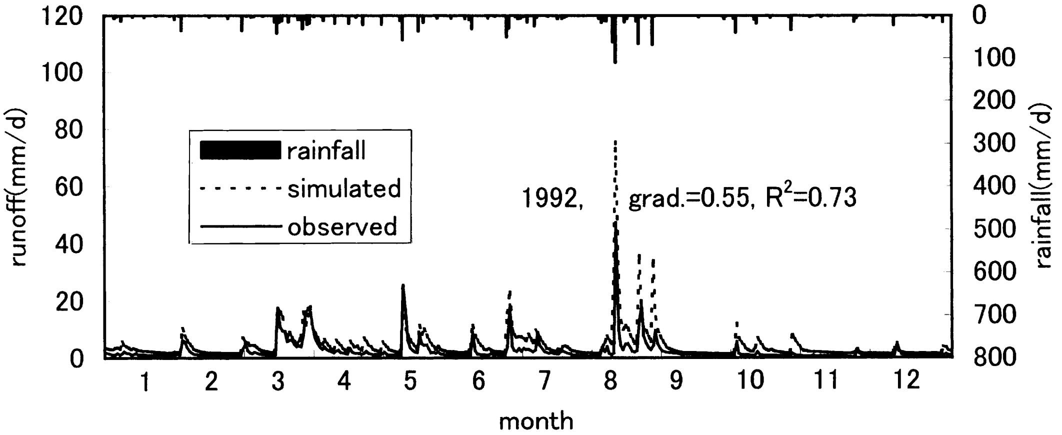

Fig. 7. Hydrograph drawn with the calculated runoff using the model coefficients evaluated using geographical information in Nomura dam catchment for the data in 1992. 'Grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

model coefficients and basin factors was investigated by evaluating the signs in front of the independent variables in Eq. (7). The following describes results of investigations into the relationship between the model coefficients and the basin factors.

5.4.1. Coefficient of 1st tank

According to equations of $A11$ and $A12$, the ratio of surface-flow to total runoff is relatively high in large basins consisting mainly of geology type C that has low vertical permeability. The equation of $B1$ indicates that a basin widely covered with soil type C has a tendency to have small infiltration into deeper layers. The equations of $Z11$ and $Z12$ show that the land-use type 6 and 10 have large storage capacity for surface-flow and the capacity becomes larger if the basin has small gradient.

5.4.2. Coefficient of 2nd tank

The equation of $A2$ and $B2$ showed that rapid interflow is large if a basin has larger area covered with soil type C and if the basin is mainly over the geology type C. According to the equation for $Z2$, a steep basin has small storage capacity for rapid interflow.

5.4.3. Coefficient of 3rd tank

According to the equations of $A3$ and $B3$, delayed interflow can be small and degree of contribution to

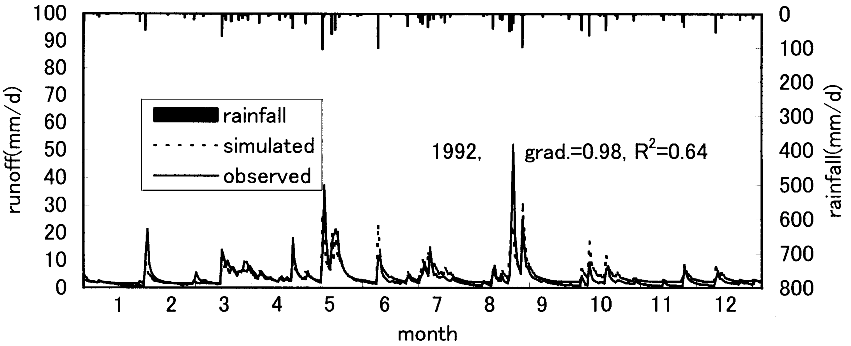

Fig. 8. Hydrograph drawn with the calculated runoff using the model coefficients evaluated using geographical information in Fudagawa dam catchment for the data in 1992. 'Grad.' is the regression coefficients of simulated runoff to observed runoff, and 'R^2' the coefficients of determination.

base-flow becomes large in a basin mainly consists of geology type *A*. The equation of *Z*3 indicates that a basin mainly consists of geology type *C* has small storage capacity and small degree of contribution to base-flow.

5.4.4. Coefficient of 4th tank

The equations of *A*4 and *H*4 show that geology type *A* tends to retain groundwater and to discharge quite slowly as base-flow. Moreover, the equation for *H*4 indicates that a large basin tends to recharge groundwater.

5.5. Comparison with existing optimization

After calculating the model coefficients estimated by the equations of multiple linear regression, runoff quantities were simulated with the coefficients for each year in each basin. Figs. 5 and 6 show the results of the simulations. The hydrographs are reproduced quite well with magnitudes and shapes similar to those of Figs. 3 and 4.

5.6. Performance of this method

The results of Figs. 5 and 6 were very promising to a degree similar to those obtained by the tank models optimized by SP method. This was because they were the results of the basins that were used in multiple linear regression analysis. To know the limits of the application, runoff simulations were made similarly using geographical information and Eq. (7) in two test basins (Fig. 1), Nomura and Fudagawa dam basins, which were not considered in the multiple linear regression analysis. General information on the two test basins are shown in Table 9. Figs. 7 and 8 are the results of this verification. Although peak runoff quantities were not simulated well, base-flow were well simulated on the whole. Therefore, it was recognized that this modeling method could be useful in applying tank models for long-term and daily analyses in Japan.

6. Conclusion

A multiple linear regression analysis was made between 12 sets of tank model coefficients optimized by SP method and basin characteristics derived from geographical information. Applications of the regression equations showed that tank model could be successfully applied using geographical characteristics of a watershed with little effect of snow. These results also showed that tank models could be constructed using the regression relationships obtained even if there were no runoff data. The method could help to remove a big hurdle in developing a runoff model.

Acknowledgements

This study was supported by JSPS Research Fellowship for Young Scientists and Grant-Aided for Scientific Research (Grant-in-Aid for JSPS Fellows) in part. Information on watersheds were given from branch offices of Ministry of Construction, Japan, branch offices of The Water Resources Development Public Corporation (WARDEC), Japan and prefecture offices in Japan.

References

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology. McGraw-Hill, New York, NY, USA pp. 498.

Development Section, River Bureau, Ministry of Construction, Japan, 1994. An Annual Report of the Management of the Multiple-Purpose Reservoir, Chugoku Construction Public Utility Association (in Japanese).

Development Section, River Bureau, Ministry of Construction, Japan, 1995. An Annual Report of the Management of the Multiple-Purpose Reservoir, Chugoku Construction Public Utility Association (in Japanese).

Development Section, River Bureau, Ministry of Construction, Japan, 1996. An Annual Report of the Management of the Multiple-Purpose Reservoir, Chugoku Construction Public Utility Association (in Japanese).

Ishihara, Y., Kobatake, S., 1979. Runoff model for flood forecasting. Bulletin of the Disaster Prevention Research Institute, Kyoto University, Vol. 29, pp. 27–43.

Kadoya, M., Nagai, A., 1980. Method of runoff analysis (12th). Japan Society of Irrigation, Drainage and Reclamation Engineering 48 (12), 935–943.

Nakano, H., 1976. Forest Hydrology. Kyouritsu Pub. Co, Japan (in Japanese).

Powell, M.J.D., 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. The Computer Journal 7, 155–162.

Sugawara, 1972. Method of Runoff Analysis. Kyouritsu Pub. Co, Japan (in Japanese).

Sugawara, M., 1995. Tank model. In: Singh, V.J. (Ed.). Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, CO, USA.