

NRCS CURVE NUMBER EMPLOYED HYDROLOGIC HOMOGENEOUS REGIONALIZATION IN REGIONAL FLOOD FREQUENCY ANALYSIS

Binaya Kumar Mishra¹, Kaoru Takara² and Yasuto Tachikawa³

¹ Graduate Student, Department of Urban and Environmental Engineering, Kyoto University
Uji Campus, Kyoto, 611-0011, Japan, e-mail: mishra@flood.dpri.kyoto-u.ac.jp

² Professor, Disaster Prevention Research Institute, Kyoto University
Uji Campus, Kyoto, 611-0011, Japan, e-mail: takara@flood.dpri.kyoto-u.ac.jp

³ Associate Professor, Department of Urban and Environmental Engineering, Kyoto University
Katsura Campus, Kyoto, 615-8540, Japan, e-mail: tachikawa@mbox.kudpc.kyoto-u.ac.jp

ABSTRACT

Estimation of extreme flood for different return period is required in design of various hydraulic structures. Regional flood frequency analysis is an effective method for estimating such extreme flood. Delineation of hydrologic homogeneous regions is key step in success of regional flood frequency analysis. This study deals hydrologic regionalization of Nepalese territory. Cluster analysis, a multivariate technique, is generally used to identify objectively hydrological regions. This work addresses the difficulty of allocating suitable weight to different clustering attributes by employing NRCS runoff curve number. On superimposing monsoon rainfall pattern map over runoff curve number map, five hydrologic regions were proposed. L-moment based regional hydrologic homogeneity test led finalization of hydrological homogeneous regions.

Keywords: curve number, flood frequency, homogeneity test, L-moment

1. INTRODUCTION

Estimation of extreme flood for different return periods is required in design and planning of various water related structures. This extreme flood is popularly known as design flood. Design flood estimation methods can be broadly classified into two groups: streamflow-based methods and rainfall-based methods. This work is related to streamflow-based methods which base their analysis purely on stream-gauging data. The most common streamflow-based methods are the flood frequency analysis and regional flood frequency analysis. Flood frequency analysis is applicable to locations which possess long records of observed flood data. However, the gauged locations rarely coincide with the locations at which water-related structures are going to be constructed. In such situation, regional hydrological characteristics need to be used in estimation of design flood. Regional flood frequency analysis is such method which makes the use of regional hydrological characteristics for estimating design flood. Regional flood frequency analysis is popular method for estimating flood peaks within specified probabilities of exceedance at ungauged sites or enhancing estimation at gauged sites where historical records are short.

Among various methods (Cunnane, 1998) of regional flood frequency analysis, index flood method is the most popular. The index flood procedure includes three major steps: hydrologic homogeneous regionalization, selection of regional distribution function and estimation of index flood (scale factor). This study deals with the first step i.e. hydrologic homogeneous regionalization. Regionalization refers grouping of territory/basins having

similar flood generating mechanisms. In other word, all the sites in a hydrologic homogeneous region possess identical frequency distribution apart from a site-specific scale factor. Regionalization is required to affect the spatial transfer of hydrologic information.

There are no specific guidelines for identifying homogeneous regions. This is due to the complexity in understanding precisely the factors that have effect on the generation of floods. Several attempts have been made by different authors to identify hydrological homogeneous regions in different parts of the world. Earlier research work used subjective consideration on the attributes like location, topography of the basin in forming hydrologic homogeneous regions. More recent works used multivariate technique such as cluster analysis to identify objectively hydrologic homogeneous regions. The cluster analysis is based either on physiographic characteristics or on flood data characteristics. Burn and Goel (2000) used three physiographic characteristics: catchment area, length of the main stream and slope of the main stream in central Indian River basins. Regionalization using basin characteristics such as catchment area, rainfall and soil type was suggested as an attractive approach of regionalization by Wiltshire (1985). Mosley (1981) used clustering of flood statistics: specific flood (peak discharge generated by unit drainage area) and coefficient of variance (C_v) for delimitation of New Zealand river basins. Burn (1997) used seasonality statistic (average date of flood events) to avoid the dual use (region forming and homogeneity test) of statistics derived from flood magnitudes.

Research on regionalization of Nepalese rivers is extremely limited. McDonald and Partners (1990) divided Nepalese territory into 7 regions using information for *low flow estimation*. Sharma and Adhikari (2004) has developed relationships for estimating extreme (minimum and maximum) and long-term (average) flows using regression technique by considering the whole Nepalese territory as one homogeneous region. Except similar previous work of Mishra et al. (2008), no other study is found on hydrologic regionalization of Nepalese rivers for design flood estimation. In the previous regionalization work, the authors used specific flood, date of flood events and lat-long of river basins as clustering attributes.

A number of difficulties were experienced by the authors in the earlier attempt. One of the major difficulties is different measuring units of clustering attributes. For example, one attribute like soil type is expressed in term of infiltration rate (mm/h) whereas another attribute like catchment area is expressed in km^2 . Another problem is measuring scale (e.g., mm/hr or cm/hr) of clustering attributes. Because of different units/scale, a suitable weight needs to be allocated to each clustering attribute. Allocation of suitable weight to different clustering attributes is difficult and subjective work. A new approach of regionalization is presented here which addresses problems associated with different measuring units/scale of clustering attribute. Soil type, land cover, land slope and rainfall pattern are considered major physio-climatic attributes which govern flood generation. Effect of soil type, land slope and land cover can be taken into account by using NRCS runoff curve number (CN).

The present work identifies hydrologic homogeneous regions inside Nepalese territory. The work started with the estimation of runoff curve number for different combinations of soil type, land cover and land slope ranges. In next step, area-weighted CN was computed for spatially well-representing sample basins. These sample basins with drainage area smaller than 250 km^2 were generated from GTOPO30 digital elevation model (DEM). Hydrological regions were proposed by superimposing digitized monsoon rainfall map over CN associated basins map. Validation of proposed regions was made using the L-moment based regional homogeneity test. Theoretical background on NRCS runoff curve number and regional homogeneity test has been discussed in section 2 and 3 respectively. The following section deals study area, data analysis and region forming process. Results of the work and discussion on the obtained results have been discussed in section 7. The final section summarizes and concludes the present work.

2. NRCS RUNOFF CURVE NUMBER

The runoff curve number (CN), an index developed by USDA Natural Resource Conservation Service (NRCS), represents potential for storm water runoff within a drainage area. USDA developed runoff curve number from field experiments in small catchments for different combinations of hydrological soil group, land cover and soil moisture. Further research in different part of the world suggested adjustments in CN for different range of land slope and others. The curve number varies from 0 to 100; lower numbers indicate low runoff potential while larger numbers indicate increasing runoff potential. The CN is widely used for estimating direct runoff from a rainfall event. The runoff equation (Ritzema, 1994) is;

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \quad (1)$$

where Q is direct runoff (mm), P is rainfall (mm), S is potential maximum retention (mm), and I_a is initial abstraction (mm). I_a is usually taken 20% of S. The runoff curve number, CN, is then related to S as:

$$CN = \frac{25400}{254 + S} \quad (2)$$

For NRCS runoff curve number estimation, the available soil data needs to be expressed into hydrological soil group. There are four hydrologic soil groups: A, B, C and D. These soil groups are roughly indicate high, moderate, slow and very slow infiltration rates respectively. The land cover is defined like woodland, forest, cropland, urban area etc. in estimating runoff curve number.

Soil moisture condition in the drainage basin before runoff occurs is another important factor influencing the CN value. Antecedent moisture condition (AMC) is based on the 5-day antecedent rainfall i.e. the accumulated total rainfall preceding the runoff under consideration. AMC I, II and III roughly indicate dry, average and saturated condition of drainage basins.

A curve number value based on different combination of hydrological soil group, land cover, land treatment, impervious percentage and land slope is available in tabular form. This table is popularly known as look-up table. Initially the look-up table is available for average moisture condition (AMC II). Then, if necessary, these CN values are modified based on 5-day antecedent rainfall data.

3. REGIONAL HOMOGENEITY TEST

Regional hydrologic homogeneity test is required for each proposed hydrologic regions. The L-moment based regional homogeneity test (Hosking and Wallis, 1997) needs computation of L-moment ratios: L-coefficient of variance (LC_v), L-skewness (LC_s) and L-kurtosis (LC_k) at each station for the available data series. Visually, regional homogeneity can be identified from the plot of LC_v versus LC_s or LC_s versus LC_k on L-moment ratio diagram (Rao and Hamed, 1997). If the plotted points are closer, the region can be expected as hydrological homogeneous. Numerically, there are two ways to check regional homogeneity. The first one is discordancy measure, intended to identify those sites that are unusual with the

group as a whole. The discordancy measure estimates how far a given site is from the centre of the group. If $u_i = [t^{(i)} \ t_3^{(i)} \ t_4^{(i)}]^T$ is L-moment ratios vector for the site i and superscript T as transpose of the vector, then discordancy (D_i) can be expressed as Eq. (3).

$$D_i = \frac{1}{3}(u_i - \bar{u})^T A^{-1}(u_i - \bar{u}) \quad (3)$$

where \bar{u} and A in a region for N gauge locations are defined as;

$$\bar{u} = \frac{1}{N} \sum_{i=1}^N u_i \quad (4)$$

$$A = \sum_i^N (u_i - \bar{u})(u_i - \bar{u})^T \quad (5)$$

A site i is declared to be discordant if D_i is large. Critical value of D_i is 3 for the region having 15 or more nos. of gauged sites. For the region with smaller nos. of gauge sites, critical value of D_i varies with nos. of gauge sites in the region.

The second regional homogeneity test is heterogeneity measure, intended to estimate the degree of heterogeneity in a group of sites and to assess whether they might be reasonably treated as homogeneous. This test compares the variability of L-moment ratios for the catchments in a region with the expected variability obtained from simulation. The heterogeneity measure is expressed as:

$$H = \frac{(V - \mu_v)}{\sigma_v} \quad (6)$$

where V is the weighted variance of the LC_v or LC_s or LC_k for the region, μ_v is mean and σ_v standard deviation of V obtained from simulation. Heterogeneity measure H_1 , H_2 and H_3 indicates variability in LC_v , LC_s and LC_k respectively. Heterogeneity measure H_1 is suggested as the most reliable parameter for testing regional homogeneity. A region is considered homogeneous if $H < 1$, possibly heterogeneous if $1 \leq H < 2$, and definitely heterogeneous if $H \geq 2$.

4. STUDY AREA

This study is intended to form hydrologic regions inside the Nepalese territory for carrying regional flood frequency analysis. Nepal, roughly a rectangle in shape with an area of 147,181 square kilometer, is situated between China in north and India in remaining three sides. It has a length of 885 km east–west and width of 145–248 km north–south. Within this relatively small latitudinal extent, altitude rises from 60 m in south to the 8848 m (Mount Everest, the world's highest peak) in North. Physiographically, the country is divided into five regions (Figure 1): Terai (Plain), Siwalik Hills, Middle Mountains, High Mountains and High Himalayas.

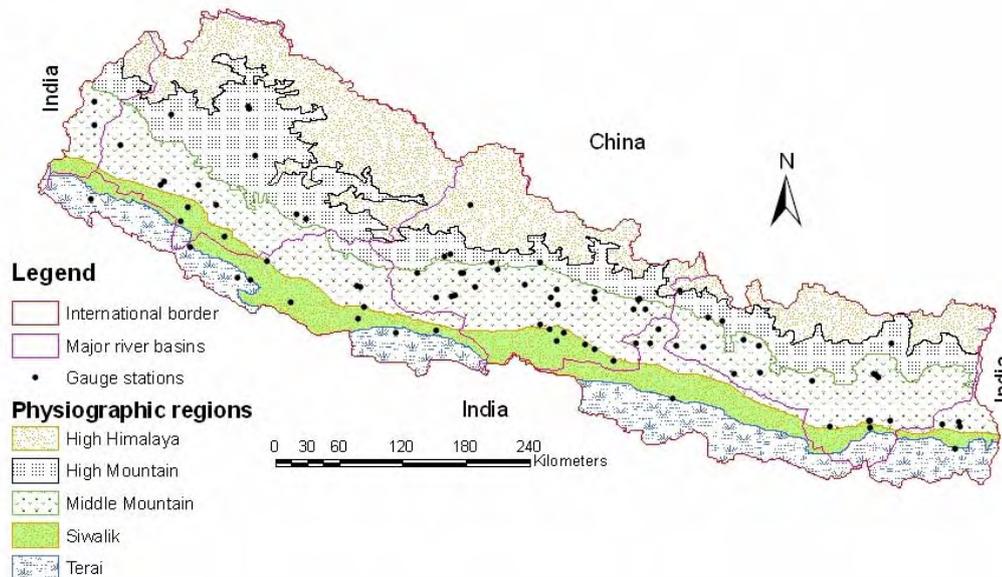


Figure 1 Physiographic regions of Nepal.

The average annual precipitation is around 1600 mm of which almost 80% or more occurs during the period of June-September. All the river systems of Nepal are tributaries to the Ganges river draining ultimately to the Bay of Bengal. The major tributaries generally flow towards south direction.

5. DATA AND PRELIMINARY ANALYSIS

The physiographic/climatic attributes: soil type, land cover, land slope and monsoon rainfall pattern have been considered in proposing hydrologic regions. Effect of soil type, land cover and land slope has been dealt using runoff curve number (CN). GTOPO30 digital elevation model (DEM) was used in generating sample basins for the study purpose. Area-weighted runoff curve number value was calculated for each spatially representing sample basins. Annual instantaneous flood data of 49 river basins have been used for regional homogeneity test. Discussion on these data used is dealt in the following subsections.

5.1 Soil type

Global soil data at resolution of 5 arc-min (approx. 10 km) has been prepared by FAO and can be freely downloaded through internet. This digital soil map is available in shape file format with geographic coordinate system. Digital soil data of the study area was extracted from the downloaded large data set. Textural classification associated with soil data was considered important here since it can be easily converted into hydrological soil group which is required in CN estimation. Textural classes reflect the relative proportions of clay (fraction less than 0.002mm), silt (0.002 - 0.05mm) and sand (0.05 - 2mm) in the soil. Based on their proportion, hydrological soil group map of Nepal (Fig.2) was prepared.

5.2 Land cover

The University of Maryland, Department of Geography generated global land cover

classification collection in 1998. The university made fourteen land cover categories from the analysis of AVHRR satellites imagery acquired between 1981 and 1994. This land cover data can be freely downloaded at various resolutions through internet. Land cover data of the study area at resolution of 1km was made available from the internet site maintained by Global Land Cover Facility (GLCF). Digital land cover map (Figure 3) required for CN estimation purpose was prepared after simple modification in original classification.

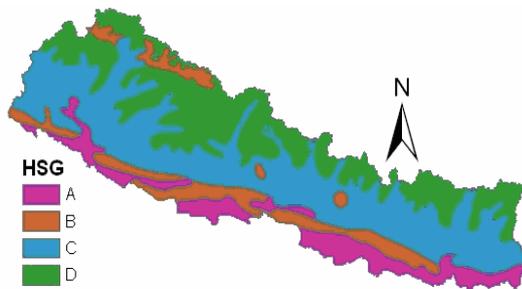


Figure 2 Hydrological soil groups of Nepal.

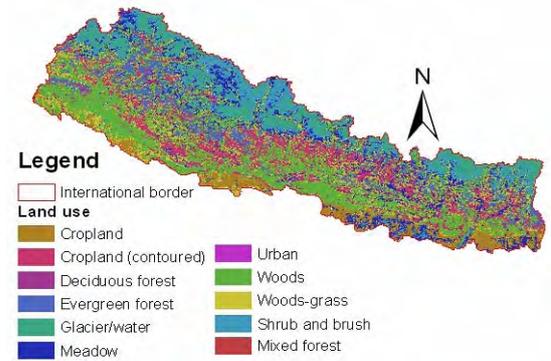


Figure 3 Land cover map of Nepal.

5.3 Rainfall distribution

Rainfall pattern i.e. distribution of rainfall over time and space is important in hydrologic regionalization. In general, annual rainfall amount decreases slightly from east to west and increases with elevation from south to north on windward slopes. Monsoon rainfall which occurs during the months of June to September has more importance in regionalization for design flood estimation. About 80 or more percent of the annual rainfall occurs during this season and the rainfall regime covers the whole country except the northern Himalayan region. Such concentration of rainfall during a few months results large floods and landslides. The study made three classes of monsoon rainfall pattern from mean monsoon precipitation map available at http://www.fao.org/ag/agL/swlwpnr/reports/y_sa/z_np/nmp134.htm. These classes (Figure 4) were defined for rainfall value less than 1000 mm, 1000 to 1500 mm and more than 1500 mm.

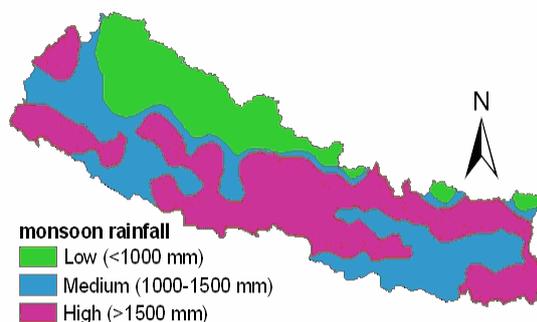


Figure 4 Monsoon rainfall map of Nepal.

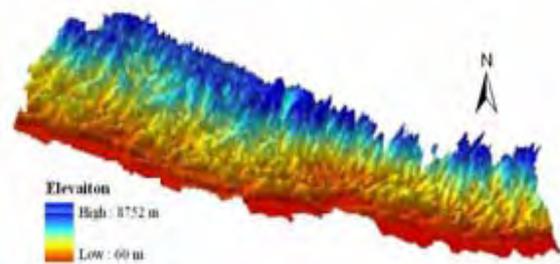


Figure 5 GTOPO30 DEM of Nepal.

5.4 Digital elevation model

Digital elevation model (DEM) of the study area (Figure 5) was clipped from large GTOPO30 DEM of Asia. This can be freely downloaded from the internet site maintained by

global data centre of NASA. This 30 arc-sec (approx. 1 km) resolution DEM was used to generate spatially representing sample basins and stream network using Arc GIS. The stream network derived from this DEM was compared with the stream network available in various reports available at internet and found acceptably okay. Also, the areas of river basins for the gauged outlets generated from DEM were compared with the areas made available by department of hydrology and meteorology (DHM), Nepal for further validation of downloaded DEM.

5.5 Flood data

Regional hydrologic homogeneity test need to be performed on each of the proposed hydrologic region. It needs adequate number of gauge station with long record length of flood data. Department of Hydrology and Meteorology (DHM), Nepal is one and only government organization responsible for collecting and distributing flood data in Nepal. Regularly observed instantaneous annual maximum stream flow data were collected for 49 stream gauging stations in Nepal. Selection of these stream gauging stations were based on basins' boundary position and record length of observed data. Boundaries of all these stations are restricted inside a proposed hydrologic region. Most of these stations are situated in middle mountain region. Out of these 49 considered river basins, 46 have observed length of more than 10 years. Because of inadequate gauging stations in lower region, 3 stations were selected though their record length is smaller.

Analysis of flood data was started by visual inspection. These flood data have been obtained from the 1998, 2003 and 2004's DHM publications. In addition, some recent flood data were collected in digital format from the authority of DHM, Nepal. In case of overlapping mismatched data, the latest publication was considered as correct flood data. The flood data series being homogeneous and stationary are the basic assumptions in flood frequency analysis. Test of homogeneity and stationary for the flood data series was performed as discussed by Mann and Whitney (1947). All the data series were found homogeneous and stationary at 5% significance level. Presence of outliers in the data causes difficulties when fitting a distribution to the data. The G-B test (Grubbs and Beck, 1972, Rao and Hamed, 2000) was used to detect outlier. Approximate relationship proposed at 10% significance level by Pilon and Harvey (1993) was used in calculating G-B statistic. The study found 1 no. of outliers at 14 stations (station indices: 267, 404.6, 438, 439.8, 445.3, 465, 530, 570, 589, 602, 620, 627.5, 650 & 660), 2 nos. at station index 447.9 and 3 nos. at station index 241.

6 REGION FORMING PROCESS

The region forming started with development of NRCS runoff curve number map for hydrological soil group, land cover and land slope. Since information on rainfall amount prior to the instantaneous flood data series could not be obtained, average soil moisture was considered while estimating the runoff curve number. Arc CN-runoff tool (Zhan and Huang, 2004), a rainfall-runoff model on Arc GIS platform, can generate efficiently runoff curve number map. The tool Arc CN-runoff requires with a look-up table (Table 1) consisting specific number between 0 and 100 for different soil type, land cover and land slope combination. Area-weighted average runoff curve number was calculated for each of 650 sample basins. The present study did not consider sample basins from Himalayan region considering its insignificance in monsoon flood. Hydrologic regions were proposed by superimposing monsoon rainfall map over sample basins associated with CN values.

Table 1 CN association with HSG and land cover.

Land cover	Soil Type			
	A	B	C	D
Meadow	30	58	71	78
Woods - Grass (Fair)	43	65	76	82
Woods (Fair)	36	60	73	79
Deciduous Forest	36	60	73	79
Evergreen Forest	40	66	77	85
Mixed Forest	38	63	75	82
Urban	68	80	88	94
Cropland	49	69	79	84
Cropland (terraced)	65	74	82	86
Shrub / Brush Tundra	48	67	77	83
Glaciers/Stream/Lake	0	0	0	0

The L-moment based regional hydrologic homogeneity test was applied on each proposed regions. To apply the homogeneity test, L-moments/ratios were computed for collected flood data at each station. Heterogeneity measure, particularly H_1 , was checked for each region. The region for which heterogeneity measure was found smaller/nearer to 2 was accepted as homogeneous region. For the region having heterogeneity measure far beyond 2, discordancy measure (D_i) was taken into consideration to make adjustment in proposed regions. When heterogeneity measure and discordancy measure reduced to limiting value, the region was declared hydrological homogeneous.

7 RESULTS AND DISCUSSION

Out of total 650 sample basins, 450 nos. were found with CN varying from 73 to 80. The remaining basins were found with smaller curve number ranging from 50 to 65. This observation led 2 divisions of Nepalese territory. The basins with higher CN were found mostly in mid-mountain and higher mountain whereas sample basins with smaller CN values were found in low-mountain and plain region.

Monsoon rainfall map was superimposed over the proposed CN based regional map for further regionalization. Superimposition of monsoon rainfall map over CN-regional map led additional 3 hydrologic homogeneous regions. In this way, the study proposed total of 5 hydrologic regions and proceeded for validation of the regions.

For validation, hydrologic homogeneity test was performed in each region using historical flood data series. Heterogeneity measures H_1 , H_2 , H_3 before and after removing the discordant stations are given in Table 2. A value of $H_1 = 8.9$ was found when considered all stations of Nepalese territory. The situation improved not much even after removing discordant sites. This value is far beyond the acceptable limit indicating hydrological non-homogeneity of Nepal. The regions 2, 3 and 5 were found with acceptable value of heterogeneity measure and with no discordant sites in first attempt. The value of H_1 was found as 5.72 and 4.77 for regions 1 and 4 respectively before removing the discordant sites. When discordant sites were removed and some adjustments were made in the proposed region, the heterogeneity measure reduced near acceptable limit for region 1 and region 4. Figure 6 shows the finally identified 5 nos. of hydrologic homogeneous regions.

Table 2 Regional heterogeneity measures for Nepalese river basins.

Region	% area	All				Discordant sites removed			
		Nos. of sites (NS)	H ₁	H ₂	H ₃	NS	H ₁	H ₂	H ₃
All	100	49	8.9	2.53	-0.45	44	6.36	1	-1.67
1	16.24	9	5.72	1.41	-0.44	7	1.13	-0.07	0.11
2	14.67	2	-0.41	-0.7	-0.62	2	-0.41	-0.7	-0.62
3	33.20	10	1.96	1.1	0.09	10	1.96	1.1	0.09
4	24.63	26	4.77	1.22	-0.14	24	2.09	0.08	-0.96
5	11.26	2	-1.08	0.69	0.01	2	-1.08	0.69	0.01

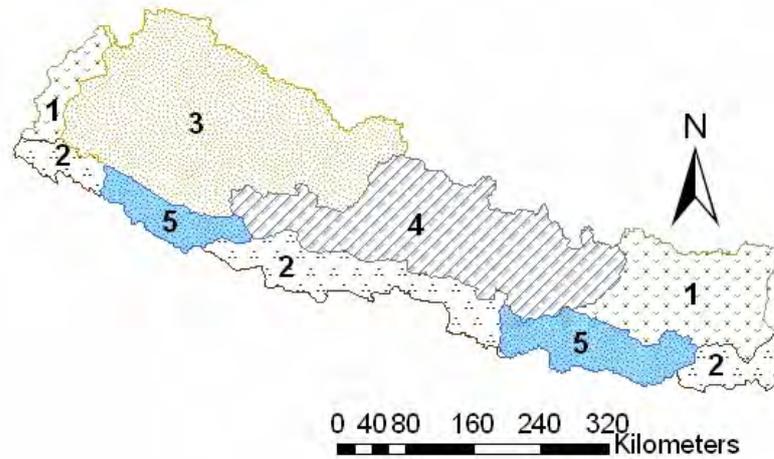


Figure 6 Hydrological homogeneous regions of Nepal.

8 SUMMARY AND CONCLUSIONS

Difficulties associated with mostly used regionalization technique i.e. clustering technique were discussed. Use of runoff curve number (CN) was proposed to address difficulties like attributes' measuring units/scale and suitable weight in cluster analysis. To estimate the runoff curve number, land cover map prepared from AVHRR satellite imagery and soil data prepared by FAO were brought in use. Hydrologic regions were proposed by superimposing monsoon rainfall map over runoff curve number map. The L-moment based regional hydrologic homogeneity test led some adjustment in proposed regions and finally 5 hydrologic regions were identified inside the Nepalese territory.

The main objective to delineate hydrological homogeneous regions for flood frequency analysis was achieved despite the complexity in understanding precisely the factors that have effect on generation of floods. Heterogeneity measures for each region were found within or near critical values. This led to conclude the runoff curve number employed regionalization approach as effective approach of hydrologic regionalization. The present regionalization work will be able to predict design flood better as compared to the previous work. In the next stage, the authors will estimate regional distribution function and index flood for each delineated region to accomplish the regional flood frequency analysis. To justify this new approach of hydrologic regionalization, the approach will be applied in country/territory other than Nepal.

ACKNOWLEDGMENTS

We would like thank to MEXT, Japan for providing financial support in carrying research work. We would also like to thank Innovative Disaster Prevention Technology and Policy Research Lab, Kyoto University, Japan for funding/technical support.

REFERENCES

- Burn, D.H. (1997): Catchment similarity for regional flood frequency analysis using seasonality measures, *Journal of Hydrology*, Vol. 202, pp. 212-230.
- Burn, D.H. and Goel, N.K. (2000): The formation of groups for regional flood frequency analysis, *Hydrological Sciences Journal*, Vol. 45, No. 1, pp. 97-112.
- Cunnane, C. (1998): Methods and merits of regional flood frequency analysis, *Journal of Hydrology*, Vol. 100, pp. 269-290.
- Grubbs, F. and Beck, G. (1972): Extension of sample sizes and percentage points for significance tests of outlying observations, *Technometrics*, Vol. 14, No. 4, pp. 847-854.
- Hosking, J.R.M. and Wallis, J.R. (1997): *Regional Frequency Analysis: An approach based on L-moments*, Cambridge University Press, Cambridge, UK.
- Mann, H.B. and Whitney, D.R. (1947): On a test of whether one of two random variables is stochastically larger than the other, *The annuals of mathematical statistics*, Vol. 18, pp. 50-60.
- McDonald, M. and Partners Ltd (1990): *Hydrology and Agro-metrology Manual, M3, Design Manuals for Irrigation Projects in Nepal*.
- Mishra, B.K., Takara, K. and Tachikawa, Y (2008): Regionalization of Nepalese river basins for flood frequency analysis, *Annual Journal of Hydraulic Engineering, JSCE*, Vo. 52, pp. 91-96.
- Mosley, M.P. (1981): Delimitation of New Zealand Hydrologic Regions, *Journal of Hydrology*, Vol. 49, pp. 173-192.
- Pilon, P.J. and Harvey, K.D. (1993): *Consolidated frequency analysis, version 3.1, reference manual*, Ecosystem Services and Evaluation Directorate, Environment Canada.
- Rao, A. R. and Hamed, K.H. (2000): *Flood Frequency Analysis*, CRC Press LLC, Florida.
- Rao, A.R. and Hamed, K.H. (1997): Regional frequency analysis of Wabash river flood data by L-moments, *Journal of hydrologic engineering*, Vol. 2, No. 4, pp 169-179.
- Ritzema, H.P. (Editor-in-chief) (1994): *Drainage principles and applications*, International Institute for Land Reclamation and Improvement, Wageningen, The Netherlands.
- Sharma, K.P. and Adhikari, N.R. (2004): *Hydrological Estimations in Nepal*, Kathmandu, Nepal.
- Wiltshire, S.E. (1985): Grouping basins for regional flood frequency analysis, *Hydrological Sciences Journal*, Vol. 30, No. 1, pp. 151-159.
- Zhan, X. and Huang, M. (2004): ArcCN-Runoff; an ArcGIS tool for generating Curve number and runoff maps, *Environmental Modelling and Software*, Vol. 19, No. 10, pp. 875-879.