# Rainfall entropy for delineation of water resources zones in Japan

Article *in* Journal of Hydrology · June 2001

DOI: 10.1016/S0022-1694(01)00355-9

Journal of Hydrology 246 (2001) 36–44

Journal of Hydrology

www.elsevier.com/locate/jhydrol

# Rainfall entropy for delineation of water resources zones in Japan

Toshihiko Kawachi[a,*], Takeo Maruyama[b], Vijay P Singh[c]

[a]*Water Resources Engineering, Division of Environmental Science and Technology Graduate School of Agriculture, Kyoto University, Kitashirakawa Oiwake-Cho, Sakyo Ku, Kyoto 606-8502, Japan*

[b]*Department of Human Sciences, Mukogawa Women's University, 6-46 Ikebiraki-Cho, Nishinomiya, Hyogo 663-8558, Japan*

[c]*Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA 70803-6405, USA*

Received 21 June 2000; revised 11 January 2001; accepted 6 February 2001

**Abstract**

Daily rainfall observed at a network of 1107 raingauges in Japan is analyzed using Shannon's entropy (informational) theory. The uncertainty of the over-a-year rainfall apportionment is quantitatively measured by entropy and an isoentropy map for the whole country is constructed. This isoentropy map is compared with Japan's well-known climatic division map, and is found to be capable of satisfactorily explaining the characteristics of nationwide rainfall. When used in conjunction with an isohyetal map of the average annual rainfall, it enables a relative assessment or categorization of the potential availability of water resources. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords*: Entropy; Rainfall; Uncertainty; Climatic division map; Water resources availability

## 1. Introduction

Rainfall is vital for assessment of the potential availability of water resources in Japan. Although rainfall is cyclic in nature, its distribution in both space and time is highly erratic, leading to unevenly distributed water resources. Development and management of water resources require not only the aggregate rainfall but also its variability. The homogeneity of water supply increases with decreasing spatial variability of rainfall. On the other hand, the less temporally variable the rainfall, the more perennial the water supply. When evaluating the availability of water resources (i.e., water resources potential or water supply potential) in a watershed or investigating the relative availability of local or regional water resources, the temporal variability of rainfall is of major concern.

Although the entropy theory has been applied to quantitatively assess uncertainties of hydrologic variables, and models of water resources systems and their parameters quantitatively (Harmancioglu et al., 1992; Singh, 1997; Harmancioglu and Singh, 1998), its application to delineation of rainfall zones and consequent assessment of water resources on a large scale does not seem to have been made. Caselton and Husain (1980) used the maximum information transmission to select stations in a hydrometric network. Husain (1989) applied an entropy-based methodology for selecting the optimum number of stations from a dense network and expanding a network using data from an existing sparse network by interpolation of information and identification of zones with minimum hydrologic information. Krstanovic and Singh (1992a) investigated information transfer between selected drought or flood sequences,

* Corresponding author. Tel.: +81-75-753-6348; fax: +81-75-753-6349.
*E-mail address:* tk@tk.org (T. Kawachi).

0022-1694/01/$ - see front matter 
PII: S0022-1694(01)00355-9

using marginal entropy, joint entropy and transinformation in long-term monthly rainfall series. Employing information transfer, Krstanovic and Singh (1992b,c) discussed the applicability of the entropy theory to evaluation of spatiotemporal variability of rainfall and the adequacy of the existing raingauges in Louisiana. Yang and Burn (1994) presented an entropy-based methodology for design of data collection systems. Using marginal entropy, Maruyama and Kawachi (1998) investigated the characteristics of local rainfall in Japan.

The objective of this study is to undertake a nationwide evaluation of the degree of variability (i.e., the degree of uncertainty of rainfall occurrence or temporal rainfall apportionment) of annual rainfall pattern in Japan, delineate entropy distribution of rainfall nationwide, contrast the entropy map with Japan's well-known climatic division map, and construct a water availability map by linking entropy with rainfall.

## 2. Entropy of rainfall events

### 2.1. Shannon's informational entropy

Shannon (1948) developed the theory of informational entropy and introduced entropy as a measure of information, choice or uncertainty

$$H = -k \sum p_i \log p_i \qquad (1)$$

where $p_i$ is the probability of the $i$th outcome of a discrete random variable, $k$ is a positive constant, and $H$ is the entropy of the random variable. With constant $k$, a unit of measure, taken as $k = 1$ and the base of the logarithm as 2, Eq. (1) is simplified as

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad \text{(bit)} \qquad (2)$$

where 'bit' is used as the unit of measurement of entropy $H$, $n$ is the number of possible outcomes or events. If all $p_i$s are equal, i.e. $p_i = 1/n$, then the entropy is $H = \log_2 n$ (bit), which is a monotonically increasing function of $n$. For a given $n$, $H$ is maximum when all $p_i$s are equal. On the contrary, $H$ is minimum and is equal to zero when every $p_i$ but one is zero. This means that every outcome of the random variable is always the same and therefore one of $p_i$s becomes unity. Thus, the value of entropy, $H$, varies within the range of zero to $\log_2 n$, according to the shape of the distribution of probabilities $p_i$s. The entropy value decreases with the increasing number of constraints and increases with their decreasing number. Viewed in this manner, entropy can be regarded as a functional estimate of the uncertainty associated with the probability distribution.

### 2.2. Entropy as a measure of temporal rainfall apportionment

Uncertainties in a historical rainfall series can be quantitatively measured by using the entropy theory. Although uncertainties are associated separately with rainfall intensity and occurrence (i.e. temporal apportionment), the entropy theory has commonly been applied to determine the uncertainty of rainfall intensity or of rainfall amount. The probability $p_i$s are defined as occurrence frequencies of the discrete rainfall amounts spread over a given period. In this definition, the null entropy occurs when the rainfall intensity is uniform in time.

Here the focus is on the uncertainty of temporal rainfall apportionment and to measure its degree using entropy expressed in terms of the probability density of rainfall randomly apportioned over fragmented times. This is achieved as follows.

First, consider a historical rainfall series in a year. Let $r_i$ be the aggregate rainfall (daily rainfall) during the $i$th day in the year. For example, daily rainfall values on 1 January and on 31 December for the same year can be expressed as $r_1$ and $r_{365}$, respectively. The aggregate rainfall during the year (annual rainfall), $R$, can then be expressed by the summation of $r_i$ from $i = 1$ to 365 as

$$R = \sum_{i=1}^{365} r_i \qquad (3)$$

where the value of $r_i$ may be zero for some days and is finite for other days.

Second, we interpret the throughout-the-year rainfall variability using a random experiment. The experiment consists of a number of trials, where each trial is considered in a probabilistic sense. The experiment is conducted as follows. If the annual rainfall is $R$, say 1000 mm, then the experiment will consist of 1000 trials. The value of $R$ is rounded off to the next nearest integer. Each trial contains 365 days, 1 January–31 December. From each trial one day is randomly selected. That means that any day of the year has the same probability of being picked up. The selected day is assigned a value of rainfall of 1 mm. Suppose 3 January happens to be picked during this trial, then it will be assigned a rainfall value of 1 mm. Similarly, the next trial is performed and a day is picked randomly. The selected day will again be assigned a rainfall value of 1 mm. In this manner, 1000 trials are made and the total amount of rainfall associated with the selected days will be 1000 mm. If a particular day happens to be randomly selected 10 times, then it will be assigned a rainfall value of 10 mm (which is the sum of each selection). The days not picked up will be assigned a zero value. When the rainfall values, thus generated, are plotted against days, the result will be the rainfall series for the year. A rainfall series of $r_1, r_2, \ldots, r_n$ can thus be regarded as accumulated occurrence frequencies of unit rains for the $1, 2, \ldots, n$th days, respectively, and $r_i$ divided by a sample size of $R$ defines a relative frequency $p_i$:

$$p_i = \frac{r_i}{R} \tag{4}$$

This relative frequency, $p_i$, is regarded as an occurrence probability for the rainfall amount on the $i$th day, and therefore its distribution represents the probabilistic characteristic of the over-a-year temporal apportionment of annual rainfall, i.e. of uncertainty of rainfall occurrence. Substitution of Eq. (4) into Eq. (2) yields

$$H = -\sum_{i=1}^{n} (r_i/R)\log_2(r_i/R) \qquad \text{(bit)} \tag{5}$$

Eq. (5) is used to evaluate entropy and implies that the value of $H$ is independent of the sequential order of $r_i$ in the series. It is seen from Eq. (5) that $H$ takes on a zero value when $R$ falls only on one day of the year, and a maximum value ($\log_2 n$) when $R/n$ falls every day throughout the year (i.e. the closer the entropy $H$ approaches its maximum value, the more uniform the rainfall apportionment is (i.e. the less temporally variable the rainfall is)). $H$ can, thus, be a measure of rainfall variability in a scalar sense. It may be noted that the amount-based measure defined herein is distinct from the variability measure as a vector component of the timing-based seasonality measures (Magilligan and Graber, 1996; Burn, 1997) which consider the time of occurrence of hydrologic events, such as floods, rainfall, droughts, etc.

When yearly rainfall series for $m$ years are available at the same raingauge, a better estimate of the annual entropy can be obtained by averaging the entropy values as

$$\overline{H} = (1/m)\sum^{m} H \qquad \text{(bit)} \tag{6}$$

where $\overline{H}$ is the average entropy.

## 3. Methodology

### *3.1. Rainfall data*

Over the entire country of Japan, there are about 1300 raingauges that form a central part of the meteorological observation network, called automated meteorological data acquisition system (AMeDAS), under the Japan Meteorological Agency. Daily rainfall data for a period of 22 years from 1976 to 1997 are now available on CD-ROM (Japan Meteorological Agency, 1996–1998). All yearly rainfall series measured at these gauges cannot, however, be employed in the present study, because some of them have missing values due to the failure of recording and/or the extremely short duration of data acquisition. For a reliable estimation of $H$ and $\overline{H}$ at a raingauge, rainfall series and raingauges are screened, based on the following criteria:

(a) Any yearly rainfall series to be selected must have a complete set of daily rainfall data, thus having 365 and 366 consecutive data for the common and leap years, respectively.
(b) Any raingauge to be selected must have at least 8 years or more than 8 years of rainfall observations. The gauge must satisfy the above criterion (a).

The criterion (b) requiring $m \geq 8$ in Eq. (6) is based on the acceptance in Japan that the meteorological data consecutively observed over 8 years or more can be used for description of quasi-averaged yearly meteorology. Screening, using the above criteria, rendered only 1107 raingauges usable. These selected

1107 raingauges provide complete daily rainfall series for 8 years or more (Japan Meteorological Agency, 1996–1998). Thus, a total of 16,812 years of rainfall data was available which yields an average of 15.19 years a gauge and was employed for entropy-based calculations.

### 3.2. Calculation of rainfall entropy

The variability of rainfall is measured using the concept of entropy. The rainfall series with a 1-day resolution is considered to describe the throughout-the-year rainfall characteristics. The sequence of observed daily rainfall values in a year is described by a probability distribution of rainfall occurrence, and an annual entropy value is obtained at each raingauge station with the aid of Shannon's informational entropy (Shannon, 1948). The analysis is performed for all available yearly rainfall sequences. Then, an average of the entropies obtained over the years of interest is considered as the averaged annual entropy at the observation station. Note that the use of different averaging periods for different raingauges has insignificant effect on the results described below, since our focus is on the nationwide gross aspects of hydrological and water resources issues. The averaged annual entropies, thus obtained, for the observation stations densely scattered throughout the country are employed to construct an isoentropy map that delineates the rainfall characteristics. Considering the occurrence of these daily rainfall values within a year as a random outcome, the entropy as an objective measure of rainfall characteristics is computed, and isoentropy lines as well as isohyetal lines are constructed.

## 4. Results and discussions

### 4.1. Hydrological zoning

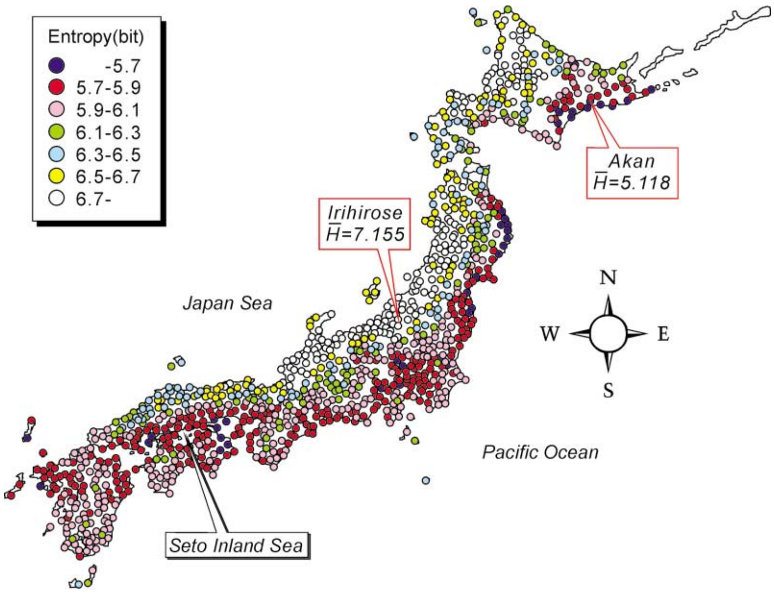
A colored dot map, depicting the nationwide distribution of averaged annual entropy $\overline{H}$, is shown in Fig. 1. This map consists of densely scattered circles each of which is drawn at the place of the raingauge, and different colored circles define the ranks of entropy. Basic statistics (mean, minimum, maximum, standard deviation (S.D.), and coefficient of variation (C.V.)) for $\overline{H}$, $R$ and $m$ selected for the calculation of entropy are summarized in Table 1. The maximum value of entropy is 7.155 bit and the minimum value is 5.118 bit; they occur at Irihirose of Niigata prefecture and Akan of Hokkaido prefecture, respectively, which are noted in Fig. 1.

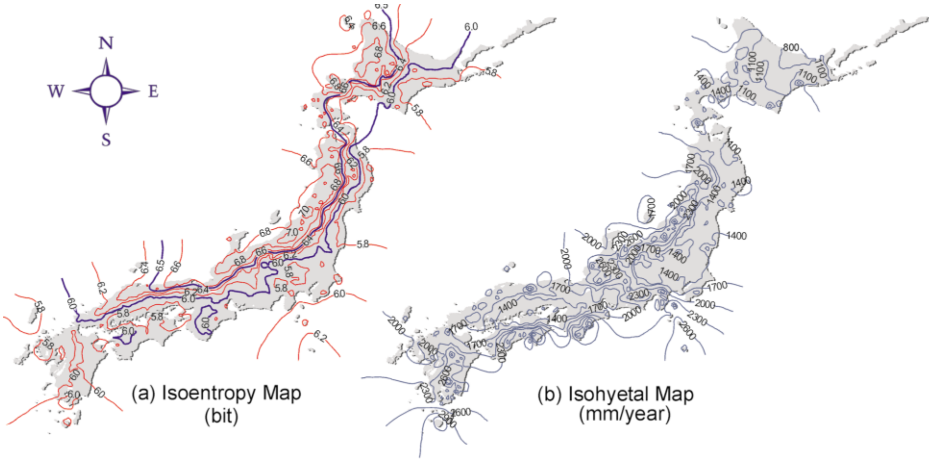
As shown in Fig. 1, entropy indicates significant zones over the whole country. This is more easily seen from the isoentropy map shown in Fig. 2(a) that delineates the entropy distribution with use of isoentropy lines. The distribution of the averaged annual rainfall, delineated by isohyetal lines, is given in Fig. 2(b).

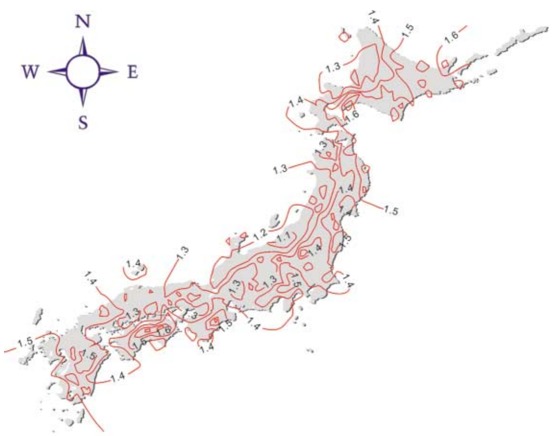
The spatial distribution of C.V. of daily rainfall is shown in Fig. 3. The coefficient of variation of daily rainfall at each raingauge was computed in the same way as for entropy. Eliminating the zero values, the mean daily value of rainfall was obtained for each raingauge for each year. Then C.V. was computed for each year. By summing the C.V. values obtained for each year and dividing the sum by the number of years of record, the annually averaged C.V. value was obtained for each raingauge. These values were obtained for constructing the C.V. map. The average value of C.V. in Japan is 1.39 and exceeds the value of 1.0. The averaged coefficient of variation of daily rainfall exhibited a significant spatial variability over the country. The values of C.V. show a complicated distribution as compared with the entropy-based map. The C.V. contours break too often and exhibit many separated islands. Thus, the C.V. map does not lend itself to a clear classification of the region. As a result, it does not have a clear match with the climate division map.

Fig. 2 also shows that the average of the entropies of over eight years at each observation point is regionally and clearly classified in Japan. Of particular interest is the agreement between Fig. 2(a) and Japan's well-known map of climatic division shown in Fig. 4. This mid-scale climatic map developed by Suzuki (1962) draws the incidence of daily rainfall during the winter season, based on the fact that Japan's climatic property can be explained satisfactorily by the spatial distribution of daily rainfall, especially during the northwest monsoon season. The lines, thus depicted, represent the rainfall boundaries that divide the whole country into three climatic regions I, II and III. Two extremities I and III are the regions with and without perceptible winter rainfall; these are referred to as the reverse and obverse of Japan, respectively.

The transitional region II is that where rain or snow




may or may not fall depending on the atmospheric pressure gradient and wind direction. This map is sufficiently comparable to the isoentropy map shown in Fig. 2(a) in as much as the former is also a result of considering the temporal variability of rainfall. This comparison implies that with a sufficiently high degree of accuracy the boundary between regions I

Fig. 1. Entropies at 1107 rainfall stations.

Fig. 2. (a) Isoentropy map and (b) isohyetal map.

Fig. 3. Contours of C.V.

Table 1
Basic statistics for $\overline{H}$, $R$ and $m$

| | Mean | Minimum | Maximum | S.D. | C.V. |
|---|---|---|---|---|---|
| Entropy (bit): $\overline{H}$ | 6.178 | 5.118 | 7.155 | 0.387 | 0.0626 |
| Annual rainfall (mm/year): $R$ | 1696.6 | 652.3 | 4390.6 | 573.6 | 0.338 |
| Number of yearly rainfall series: $m$ | 15.19 | 8 | 22 | 3.44 | 0.226 |

and II is identified as $\overline{H} = 6.5$ bit, and the boundary between II and III, inclusive of the isolated closed ones, is identified as $\overline{H} = 6.0$ bit. Such an identification is climatologically meaningful not only for quantitatively measuring the rainfall boundaries but also demonstrating the validity of the isoentropy map obtained. In addition, it justifies indeed the inference by Suzuki (1962) that the entire boundary shown in Fig. 4, obtained by using a yearly rainfall series for the year 1955, would be indistinguishably moved even if the series for another year or several years were taken into account.

### 4.2. Correlation between entropy and rainfall

Of further interest is the relation of entropy to rainfall. Depicting the averaged annual entropy on the ordinate and the averaged annual rainfall on the abscissa, Fig. 5 is constructed, which is a scattered diagram, plotting values at all the raingauges selected. Glancing at the diagram one observes that the averaged annual entropy and the averaged annual rainfall are less mutually related, with a correlation coefficient of $r = 0.1886$. This demonstrates that besides the aggregate of rainfall, its temporal apportionment can be a significant aspect of rainfall events.

### 4.3. Categorization of water resources availability

When coupled, entropy and rainfall on a yearly basis can become a measure of the throughout-the-year potential availability of water resources. To explain it qualitatively, the whole plotted area of Fig. 5 is divided into four parts, each delineated with two intersecting dividing lines that pass through the means of the respective two variables. Then, in

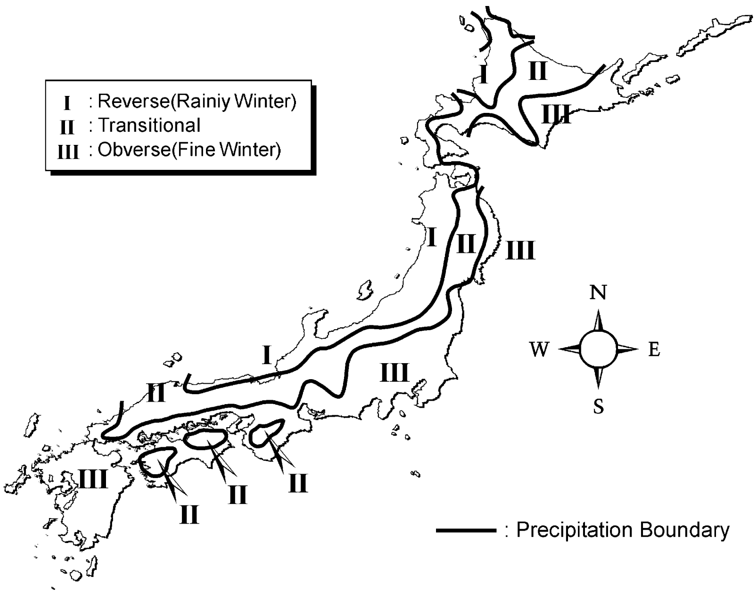



Fig. 4. Mid-scale climatic division (after Suzuki (1962)).

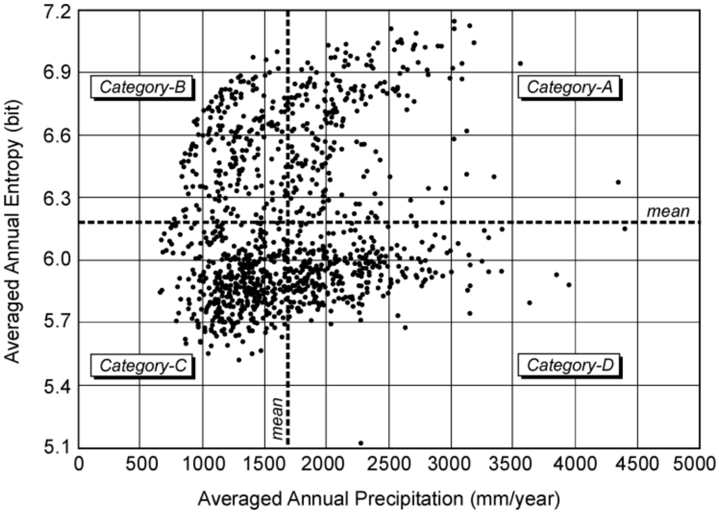



Fig. 5. Relation between entropy and precipitation.

terms of water resources availability, the respective quadrants, A, B, C and D, can be comparatively categorized as follows.

*Category-A*. Water resources are abundantly and perennially available. The inhabitants on the areas in this category can use the water resources willfully, being limited only by the permissible water withdrawal from the natural water cycle.

*Category-B*. The water resources availability is identified as moderate, since the rainwater is poor but perennial. For spells of high water demand beyond available capacity, water resources development and management are needed to temporarily raise the availability.

*Category-C*. The water resources availability is absolutely so low that permissible water use may be extremely constrained. To meet perennial or intensive water demands, what is needed is to intervene in the water cycle and appropriately increase disposable water resources by construction of water storage facilities.

*Category-D*. The water resources availability is considered as moderate, in a different sense from that of *Category-B*. This category is characterized by concentrated heavy rainfall, so that water control is needed for flood mitigation. Excess water, withheld from wastefully running off and stored in reservoirs, can be used effectively to reduce the temporal variation of water resources and consequently augment water supply for water users.

The aforementioned categories are illustrated separately in Fig. 6 in a form of a location map of the raingauges included in the respective categories. From this map, it can be readily seen that raingauges in the same category are clustered, and therefore the spatial distribution of the water resources availability is zoned with discernible boundaries. Fig. 6 serves as an improved mid-scale climatic division map, compounded by rainfall and its temporal apportionment. At the same time, although this map is a result of simple clustering with the means of the variables, it is worthy of becoming a water availability map, explaining long-established water use practices well. For example, the strip identified as *Category-A* includes the nation's most prominent rice-growing districts, such as Niigata, Yamagata and Akita prefectures, with heavy snowfall that increases rainfall in winter and consequently decreases the degree of its throughout-the-year variability. This less time-varying rainfall provides one of the natural environments necessary for good rice production, with melted snow and off-winter rainfall supplying sufficient water for soil paddling and rice planting in spring and for rice growing, respectively. On the other hand, the area belonging to *Category-C* has long been suffering from serious water shortages because of the nonuniform

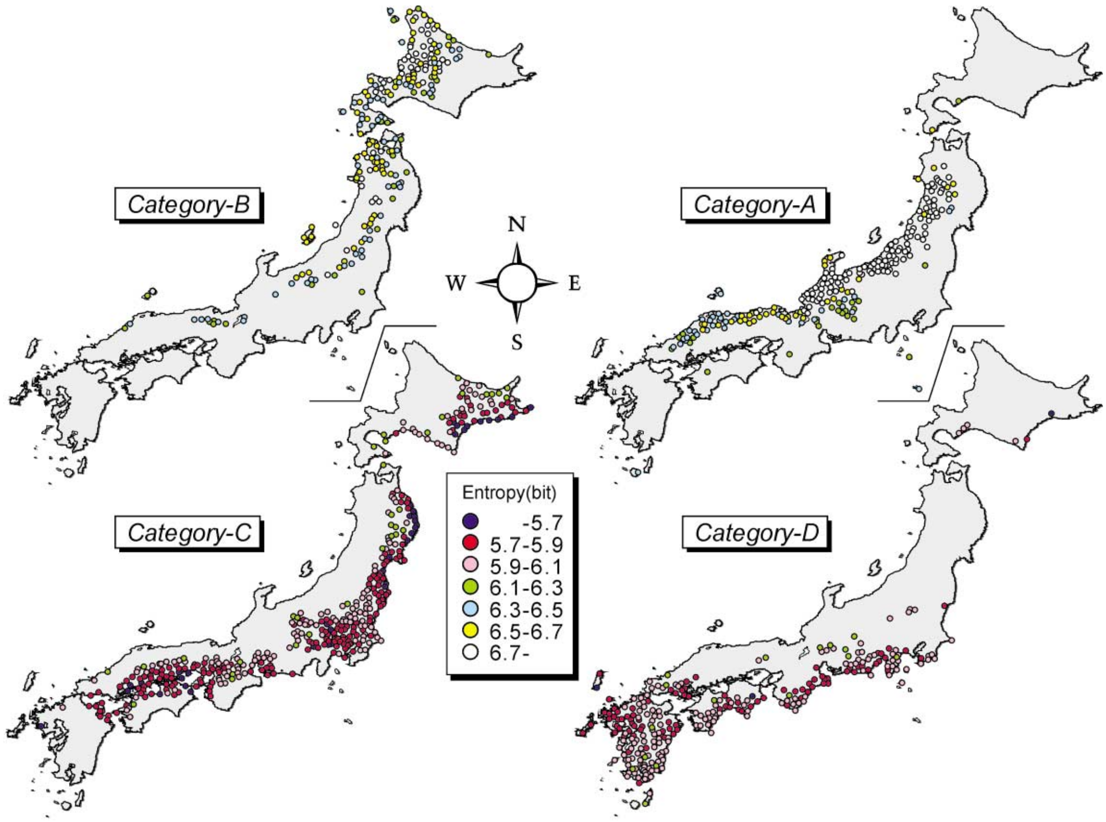



Fig. 6. Explanatory categories of water resources availability.

and absolutely scarce rainfall. To escape from such a predicament and to control the ravages of droughts, a huge number of small-scale irrigation tanks have been constructed as a self-reliant storage system all over the area. Especially on the fringes of the Seto Inland Sea that embraces the prefectures as Osaka, Hyogo, Okayama, Hiroshima and Kagawa, there exist 110,835 tanks, about half of the total 246,158 tanks in Japan, and even nowadays much effort is being made to eradicate the dearth of water.

## 5. Conclusions

The following conclusions are drawn from this study: (1) For yearly rainfall, entropy delineates a plausible climatic map that qualitatively explains rainfall characteristics. (2) An entropy-based climatic map (isoentropy map), drawn for the whole country of Japan, is valuable in comparison with the known classical climatic division map, and for coupling with the isohyetal map. (3) Coupling of entropy with annual rainfall enables a relative assessment or categorization of the potential availability of water resources at local, areal or regional levels, considering both the aggregate and the temporal variability of rainfall. (4) Coupling the entropies associated with uncertainties of rainfall intensity and temporal rainfall apportionment can help describe the water resource availability in a coherent manner.

## References

Burn, D.H., 1997. Catchment similarity for regional flood frequency analysis using seasonality measures. Journal of Hydrology 202, 212–230.

Caselton, W.F., Husain, T., 1980. Hydrological networks: information transmission. Journal of Water Resources Planning and Management, ASCE 106 (WR 2), 503–520.

Harmancioglu, N.B., Singh, V.P., 1998. Entropy in environmental and water resources. In: Herschy, R.W., Fairbridge, R.W. (Eds.). Encyclopedia of Hydrology and Water Resources. Kluwer Academic Publishers, Dordrecht, pp. 225–241.

Harmancioglu, N.B., Singh, V.P., Alpaslan, N., 1992. Versatile uses of the entropy concept in water resources. In: Singh, V.P., Fiorentino, M. (Eds.). Entropy and Energy Dissipation in Water Resources. Kluwer Academic Publishers, Dordrecht, pp. 91–117.

Husain, T., 1989. Hydrologic uncertainty measure and network design. Water Resources Bulletin 25 (3), 527–534.

Japan Meteorological Agency (1996–1998). AMeDAS Data of Japan [CD-ROM], Japan Meteorological Business Support Center.

Krstanovic, P.F., Singh, V.P., 1992a. Transfer of information in monthly rainfall series of San Jose. In: Singh, V.P., Fiorentino, M. (Eds.). Entropy and Energy Dissipation in Water Resources. Kluwer Academic Publishers, Dordrecht, pp. 155–173.

Krstanovic, P.F., Singh, V.P., 1992b. Evaluation of rainfall networks using entropy I. Water Resources Management 6, 279–293.

Krstanovic, P.F., Singh, V.P., 1992c. Evaluation of rainfall networks using entropy II. Water Resources Management 6, 295–314.

Magilligan, F.J., Graber, B.E., 1996. Hydroclimatological and geomorphic controls on the timing and spatial variability of floods in New England, USA. Journal of Hydrology 178, 159–180.

Maruyama, T., Kawachi, T., 1998. Evaluation of rainfall characteristics using entropy. Journal of Rainwater Catchment Systems 4 (1), 7–10.

Shannon, C.E., 1948. A mathematical theory of communication. The Bell System Technical Journal XXVII, 379–423 see also 623–656.

Singh, V.P., 1997. The use of entropy in hydrology and water resources. Hydrological Processes 11, 587–626.

Suzuki, H., 1962. Classification of Japanese climates. Geographical Review of Japan 35, 205–211 (in Japanese).

Yang, Y., Burn, D.H., 1994. An entropy approach to data collection network design. Journal of Hydrology 157, 307–324.