A GRID-BASED STREAM-STRUCTURE MODEL FOR LARGE-SCALE WATERSHED (GSMLW)

LUO Qiang[1], Nobuyuki TAMAI[2] and Guangwei HUANG[3]

[1]Doctoral student of River and Environment Eng. Lab., Graduate school of Eng., the University of Tokyo (7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan)
[2]Prof. of the University of Tokyo (7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan)
[3]Assistant Prof. of the University of Tokyo (7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan)

This paper describes a computer model for organizing network system and outputting river parameter, for the usage of all kinds of river model. Making use of digital data input, GSMLW develops a "free, seriate tree" structure to handle the complicated river system and generate output data of river parameters. According to requirement, river parameters can vary with the curvature distance of river central lines or straight distance to the outlet.

There are three parts of this model, one is to find central lines of stream network, one is to identify influx points of tributaries, and the other is to compute river parameters and output data files. In the modeling, the critical technique is the "free, seriate tree" structure. This tree structure is adaptable to the random tree structure of a natural river system, and thus is efficient in river system modelling.

Key Words : *Watershed, Stream network, data processing, runoff routing, flood routing, hydraulics, hydrology*

1. INTRODUCTION

GIS is utilized in most watershed models, and it is sufficient and effective in mountainous area. But in plain or urban area, GIS shows its deficiency due to the human impact and the non-obviousness of the changing of elevation. River parameter determination is usually by hands, some time involving computer aided system. This method cannot operate efficiently in a large-scale watershed. Those watersheds with area larger than 1,000 Km^2 are usually called large-scale watersheds (Singh, 1995).

GSMLW is a sub-model of Grid-based Hydrological and Hydraulic Model for Large-scale Watershed. Making use of the digital output from Complementary Stream-network and Watershed Model (CSWM), a brief introduction of which is given in Section 2, GSMLW tries to organize a river system and generate digital data as input data for all kinds of hydrological and hydraulic model. The output of GSMLW contains all necessary river parameters, like stream density, slope, channel width, flood plain area, roughness, etc., along the central line of a river or as a function of straight distance to the estuary.

Since a stream network from CSWM comprises multi-dot lines without sequence and direction, they cannot be used for the purpose of mathematical simulation. Thus, in GSMLW, the first step is to obtain central lines of streams and make them in order. Then, the second step is to identify influx points of tributaries. Finally, to calculate river parameters and output data in certain sequence.

One of the eminent features of a natural river system, or steam network, is its random tree structure. To accomplish the above three steps by a computer, a special processing and data structure must be adopted in the program. Here, we introduce a tree structure called "free, seriate tree".

2. A BRIEF INTRODUCTION OF CSWM

CSWM is a model for generating stream networks and watersheds in large-scale basins covering plain and urban area, from DEM data and other image data sources.

(1) Modelling area

Japan Institute of Geography (JIG) covers the territory of Japan with blocks in size 40 minutes in latitude (approximate 74km) and 1 degree in longitude (approximate 90km), each block has a code number. The modeling area is Kanto region of Japan, which consists of 8 prefectures, covering more than 12 blocks in the JIG's mesh (3 in west-east direction, 4 in north-south direction). And according to the division by Kanto River Bureau (KRB), there are more than 20 watersheds in Kanto region. The GRID size is 250m.

(2) Validation of the model

The source image is obtained by scanning the sketch map of watersheds and river networks drawn by KRB. But this sketch map has no longitude and latitude lines, which are necessary for modeling and must be re-located artificially by calibration several critical points comparing with those in a topographic map. Then it is easy to produce the sketch boundaries with a PhotoShop software.

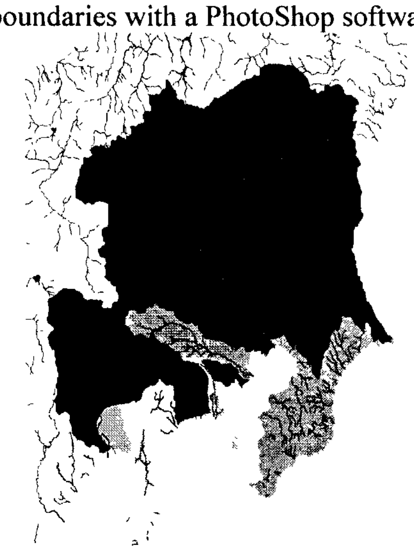

Fig. 1 Sketch watershed boundaries and stream networks

Because the sketch boundaries and the networks are from different sources, they may not fit each other perfectly (Fig. 1). Some rectification is necessary in the places where rivers belonging to one watershed overlay the other.

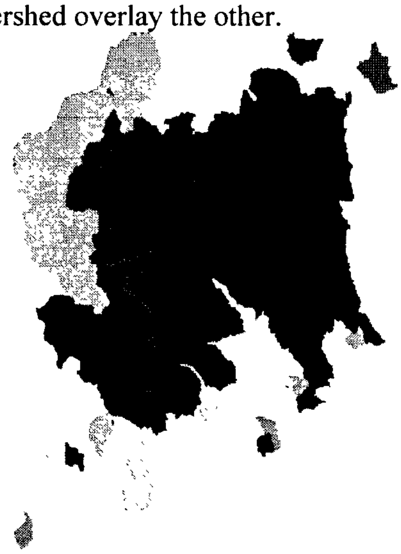

Fig. 2 GIS generated watersheds

Only by using one GRID function can the watersheds be generated: watershed(). But three important techniques must be mentioned here to obtain an effective output. First is to display the stream networks generated by GIS (not the final output). Second is to use selectpoint() function to select pour points. Third is that the area containing outlet point of a basin should be displayed in several enlarged windows separately, for the purpose of selecting pour points precisely. The output is shown in Fig. 2.

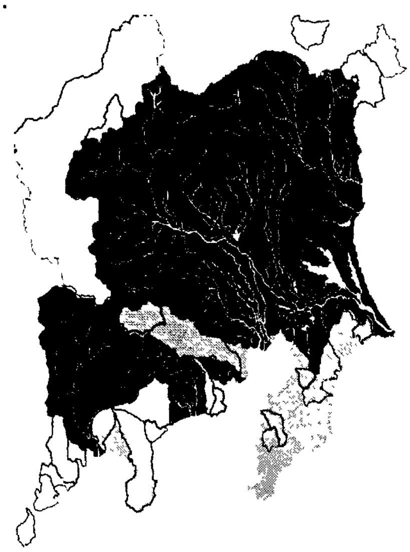

Fig. 3 Overlaying image of watersheds and stream networks

Even though the outputs from each source are quite different, some basins and some boundaries are similar. We can correct the sketch watersheds by overlaying the boundary lines of watersheds generated by GIS, as well as the stream networks, to the sketch watersheds. The overlaying image and the final output of watersheds are shown in Fig. 3 and Fig. 4 respectively.

The final outputs of CSWM are TIFF image files of watersheds and stream networks.

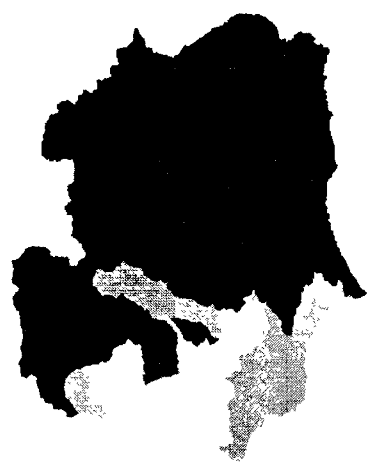

Fig. 4 Final output of watersheds

3. "FREE, SERIATE TREE" STRUCTURE

The main stream of a river can have an uncertain number of tributaries, and all these tributaries can also have an uncertain number of sub-tributaries in turn, ... and so on. This is the so-called random tree structure of a natural river system. This will cause severe trouble in computer programming by using a normal structure. One problem caused by it is that we cannot define a definite array because that we cannot decide the total number of tributaries and sub-tributaries. Even we can count this number manually and define an array, it will be a huge task to decide the processing and data output sequence of the network system. It is necessary to introduce a special structure.

(1) The concept of "free, seriate tree"

A "free, seriate tree" is a tree that has a variable number of branches, and each branch also has a variable number of sub-branches. The depth of each branch is not definite either. However, if starting from the root along the trunk, the first branch can be found, and then from the lowest point of this branch, its first sub-branch found, and so on, then all its sub-branch found. Then, back to the trunk, the second branch can be found, ..., finally all branches sub-branches can be found. In simple words, "free" means the indefiniteness of numbers of branches and their depth; "seriate" means the sequence from the root up to top or inverse. See Fig. 5.

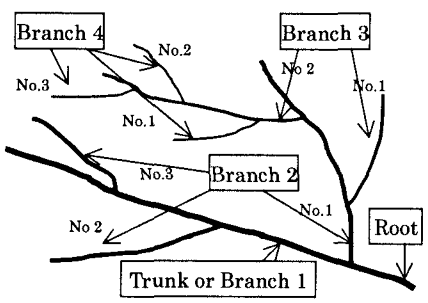

Fig. 5 A "free, seriate tree".

(2) The realization of "free, seriate tree" structure in a program

To realize the "free, seriate tree", a critical point is to introduce a stack array variable to store the previous address of proceeding. The stack variable is for the purpose of memorizing that at which point the program should continue after it completed the processing of one of the branches or sub-branches. Let us see the example of Fig. 5. When the program finished the processing of Branch 3 No.1, it should go back to Branch 2 No.1, and relocate its next processing to the flux point of Branch 3 No.2. So, before the program moves down to Branch 3 no.1, the present processing address must be pushed into the stack. Besides the stack array in this procedure, two additional variables are needed, one is a 'branch' variable, the other is a variable for storing intermediate input and output file names. Table 1 shows this processing procedure.

Table 1 Programming realization of "free, seriate tree" structure

		Stack				
Step	Branch	1	2	3	4	Processing
0	1	1	0	0	0	Branch+1, stack(Branch)+1
1	2	1	1	0	0	Branch+1, stack(Branch)+1
2	3	1	1	1	0	Branch+1, stack(Branch)+1
3	2	1	1	1	0	Branch-1
4	3	1	1	2	0	Branch+1, stack(Branch)+1
5	4	1	1	2	1	Branch+1, stack(Branch)+1
6	3	1	1	2	1	Branch-1
7	4	1	1	2	2	Branch+1, stack(Branch)+1
8	3	1	1	2	2	Branch-1
9	4	1	1	2	3	Branch+1, stack(Branch)+1
10	3	1	1	2	3	Branch-1
11	2	1	1	2	0	Stack(Branch)=0, Branch-1
12	1	1	1	0	0	Stack(Branch)=0, Branch-1
13	2	1	2	0	0	Branch+1, stack(Branch)+1
14	1	1	2	0	0	Branch-1
15	2	1	3	0	0	Branch+1, stack(Branch)+1
16	1	1	3	0	0	Branch-1
17	0	1	0	0	0	Stack(Branch)=0, Branch-1

Note: The "Branch" column indicates the current processing branch grades (numbers inside boxes in Fig.5). The "Stack" column shows the numbers of branch, and 1, 2, 3 and 4 in the head show the grade of a branch in the "Branch" column. The word "grade" is used to account the relationships amount branches, in Fig.5, the trunk has a grade of 1, all those branches having direct connections to branch grade 1 have grades of 2, those to branch grade 2 have grades of 3, and so on.

4. FINDING CENTRAL LINES AND IDENTIFYING INFLUX POINTS

Some parts of a river obtained from CSWM consist of several points, and the network is arranged in cell-based grids, not in the sequence from upstream to downstream. Before the stream network can be used for mathematical simulation, the central lines must be obtained and arranged in order. It is necessary to identify the influx point of a tributary, because of 1) the existence of sub-tributaries, 2) to know the outlet of a tributary, and 3) to make the finding of the central line of a tributary possible.

Terms like "points", " Δx ", " Δy " and "lines" in the following narrating are referred to image dots, or distances and lines composed by them.

(1) Finding central lines of a river system

The criterion for finding a central line is that the line should cross the middle points of a river. If the Δx, Δy of a river are larger than 3 (grids), the middle point should lie at $(x + \Delta x/2,\ y + \Delta y/2)$. If Δx, Δy are less than 2, the closest point to the previous point should be selected as the next middle point. All points of central line are arranged according to the "free, seriate tree".

(2) Identifying influx points

A circle is drawn at a point of the central line (Fig. 6). If the number of intersection(s) that the circle encounters central line(s) is more than 2, then the central point has a tributary or tributaries. This also acts as the ending condition of a tributary if the number is 1. However, the estuary point of the main stream must be marked manually before all processing.

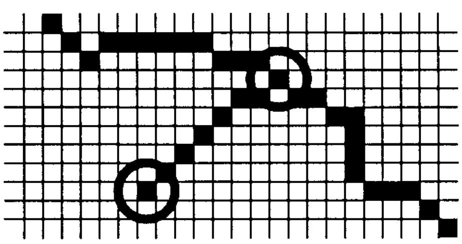

Fig. 6 To Identify an influx point and an ending point

5. RIVER PARAMETER COMPUTATION AND FINAL DATA OUTPUT

Until now, the results of processing are coordinates of central lines of the stream network, which are stored in intermediate files. These central lines are arranged according to the "free, seriate tree", or in the sequence from the estuary or outlet to upstream. Many river parameters, like stream density, bed elevation, slope and channel width, are given according to the curvature distance of the central line or straight distance to the estuary. From these intermediate results, GSMLW can easily calculate river parameters of every point of central lines of the stream network.

Most river models, e.g., runoff routing model, handle a river system by following a sequence like this, if we take the river system as a tree, starting from the highest leaves of the most upper sub-branch of the tree, and then the lower one, gradually to the branches, the trunk, finally the root. Thus, the structure of the final data outputs should be in accordance with this requirement. It's actually an upside-down "free, seriate tree". In the example of Fig. 5, the sequence of the output data is like this: Branch 2 No. 3 $\Rightarrow$ Branch 2 No. 2 $\Rightarrow$ Branch 4 No. 3 $\Rightarrow$ Branch 4 No. 2 $\Rightarrow$ Branch 4 No. 1 $\Rightarrow$ Branch 3 No. 2 $\Rightarrow$ Branch 3 No. 1 $\Rightarrow$ Branch 2 No. 1 $\Rightarrow$ Branch 1 (the trunk).

6. MODEL CALIBRATION

The model is calibrated by a river network in a small watershed, the Abegawa River. The depth of the tree of this stream network is only 3, and the branches and sub-branches are limited. Fig. 7 shows the central lines obtained by GSMLW and the original network. One can see the model works correctly in finding central lines.

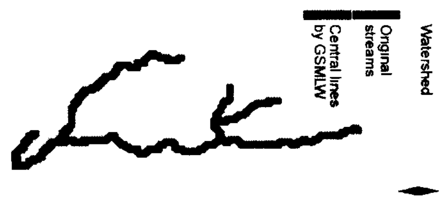

Fig. 7 Abegawa River

Another function of the model is to organize a river system in a "free, seriate tree" structure. Fig. 8 shows the arrangement order of sub-branches, branches and the main channel of the Abegawa River system. That's just the sequence of a "free, seriate tree".

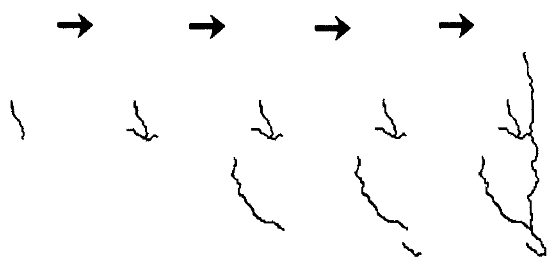

Fig. 8 Arrangement order of Abegawa Rive system

7. MODEL VERIFICATION – AN APPLICATION IN THE TONE RIVER

The Tone River is the most complicated river system in Kanto region of Japan. The basin area is about 16,000Km^2, and more than half of which is located in the Kanto Plain or urban area. This feature makes both GIS and human hands inappropriate, in generating watershed and stream network, in organizing river system and obtaining river parameters. Substitutional methods must be taken.

The watershed boundary and stream network are generated by CSWM (Fig. 9). System organization and parameters identification are resorted to GSMLW. Two of the main river parameters output by GSMLW are shown in Fig. 10 to Fig. 11.

Fig. 9 Watershed and river system of Tone River

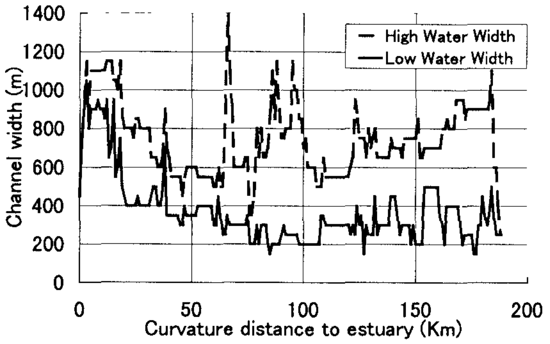

Fig. 10 Main channel width

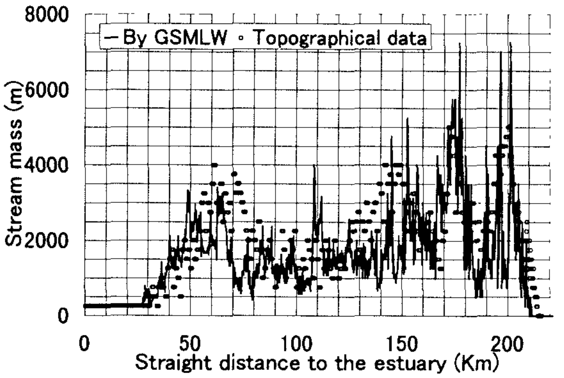

Fig. 11 Steam mass ($\Delta x = 250m$)

Figure10 is the channel width, and the data are from Upper Tone River Work Office and Lower Tone River Work Office. GSMLW assigns these data to the relevant Grids of the stream network.

Figure 11 is the stream mass, the integral of stream density over the drainage area, which serves as verification. The solid line is the result of GSMLW, the hollow dot line is drawn according to data obtained from a topographical map. Since the geological data are measured manually and the accuracy is not so high, peaks are cut off. However, the general outlines of these two lines agree quite well.

8. CONCLUSION

In coping with a stream network of a large-scale watershed, GSMLW is efficient in organizing network system and obtaining data files of river parameters. "Free, seriate tree" structure is the critical technique in GSMLW, which is adaptable to the random tree characteristic of natural rivers, and thus solved the difficulty in the programming of a natural river system in a large-scale watershed.

REFERENCES

1) Mark, D.M.: Network Models in Geomorphology, *Modelling in Geomorphological Systems*, John Wiley, 1988.
2) Roger J.M. De Wiest: Geohydrology, John Wiley and Sons, Inc., 1965
3) Singh, Vijay P. and Fiorentino, M.: *Geographical Information Systems in Hydrology*, Kluwer Academic Publishers, 1996
4) Singh, Vijay P.: *Computer Models of Watershed Hydrology*, Water Resources publication, 1995

(Received September 30,1998)