International Cartographic Association
Association Cartographique Internationale

Recognition of Risk Information – Adaptation of J. Bertin's Orderable Matrix for social communication

Keiichi Ishida[a]

[a] [Hosei University, Tokyo, Japan, keiichi.ishida.64@hosei.ac.jp]

Abstract: This paper aims to show capability of the Orderable Matrix of Jacques Bertin which is a visualization method of data analyze and/or a method to recognize data. That matrix can show the data by replacing numbers to visual element. As an example, using a set of data regarding natural hazard rankings for certain metropolitan cities in the world, this paper describes how the Orderable Matrix handles the data set and show characteristic factors of this data to understand it. Not only to see a kind of risk ranking of cities, the Orderable Matrix shows how differently danger concerned cities ones and others are. Furthermore, we will see that the visualized data by Orderable Matrix allows us to see the characteristics of the data set comprehensively and instantaneously.

Keywords: Visual analytics, Data mining, Decision making support, Social communication, Orderable Matrix, Jacques Bertin

1. Introduction

Risk information shall be assessed, for risk management purpose, based on perception of risk takers in order to control it. There are certain impact levels of risk causing unexpected results individually and/or socially. In case of risks relevant to social benefit, specialist's assessments are significantly important than ordinal peoples' one in terms of quality and quantity. However, looking at current circumstance such as nuclear power plant security, product liability law to protect consumers, informed-consent process in hospital, etc., people tend to think that such one-way explanations from specialists is not sufficient to determine what measure should be taken in next steps against concerned risks. The specialists' advice is sometimes too simplified (says yes or no) or too complicated (use many specialized terms) for other involved persons. So, one of expected situation is an establishment of decision making based on mutual understanding of risk information among all involved risk takers. And then, such required mutual under- standing of risk information shall be integrated into improvement of the risk management.

However, considering a gap of knowledge between specialists and ordinally people as non-specialists, it is not realistic to have totally same level of understanding for all concerned peoples. Therefore, a mutual understanding which is taken here in this paper is focusing on offering un intermediate understanding between them. It is not an issue on this paper how ordinally people can look thoroughly all detailed documents to be a specialist virtually. A question about communication is to avoid omissions of potentially understandable information as much as possible regarding concerned matters rather than to provide just one definitive result. For example, there is an index calling PML: probable maximum loss to measure damageability (%) of properties against earthquakes (or typhoon) in objected return period of year, i.e. PML for 500 years return period. This is normally calculated by a probabilistic simulation software based on multiple factors: location of earthquake epicenter, soil quality, ground shaking intensity, building foundation technic, construction type, number of story, shape of building, etc. One owner of property who spent a lot of money to build a solid one could have a similar PML result with another owner who did not care about it. A reason could be due to location of property how far from earthquake epicenters, soil quality under the properties or due to any other reasons. There are many different factors of PML calculation for one and the others. The result should be generated by a series of calculations, then the specialist might think that it is too complicated to explain correctly. On the other hand, it is sometimes difficult to believe the result which is simply provided as a percentage of damage without any explanation for the ordinal people. Or they wonder if all necessary considerations were appropriately taken into the result. Certainly, it is very hard to follow how the specialist establish each of calculations, and further, difficult to understand which factors would make different results between PML 5.1% and PML 4.6% of damage, for example. But the ordinally people can grasp to see what elements being essential factors without comprehension how each factor is used in complicated calculations.

One question is to have an intermediate level of communication for a situation as mentioned above, and another question is a communication tool how easily show principal points. The text or language is commonly used for communication between one person to others. However, since letters and words are representative signs of phenomena or events, understandings of sender and receiver regarding concerned conversations are sometimes different. For example, the word "risk" has several meanings: pure risk which causes purely negative impacts, speculative risk which includes a possibility of

Proceedings of the International Cartographic Association, 1, 2017. This contribution underwent single-blind peer review based on submitted abstracts | https://doi.org/10.5194/ica-proc-1-55-2017 |

benefit and loss, also, other meanings might be used for certain situations. For accurate information transmitting in communication, the time for spending such communication is an important issue to share correct definitions of words and contents used in discussions. Therefore, not only text explanation, illustration is commonly used to de- scribe certain complex things and can also explain them to different language users. A brochure of seismic intensity scale issued by Japan Meteorology Agency (JMA) is an example (Fig.1). By using the illustration, this explains how strong a ground shaking about JMA scale 4 and 5-Lower (the highest scale is 7). Rather than taking time to read detailed explanations by text, it is easily to understand circumstances by illustrations. Also, a comparison of these 2 illustrations give certain idea of difference of earthquake intensity 4 and 5-Lower. The visualization is used because of its characteristic of information transmitting: instantaneous understanding.

2. Visualization – variables and Orderable Matrix

Fig. 1. Examples of JMA seismic intensity scale (source: JMA)

The visualization will be useful to share an idea of someone to others, however, this visual communication shall be used with consideration how visual materials to be accurately applied. Jacques Bertin's Orderable Matrix, which is referred as analysis tool in this paper, describes concerning numerical data set by visual image. This means a replacing of numbers in matrix table, i.e. 1, 2, 3…, to visual elements. The visual elements to be used must be relative to the original data figures and have a good visibility of such replaced visual element. We refer again J. Bertin's work which established graphic semiology contributed to graphic and cartographic language in order to provide a good graphical communication to peoples who see them without visual obstruction. He developed a notion of visual variables based on perception effects. As Fig.2 is showing, there are 6 visual variables on two dimension: Size, Value, Texture, Color, Orientation and Shape which are perceptively characterized.

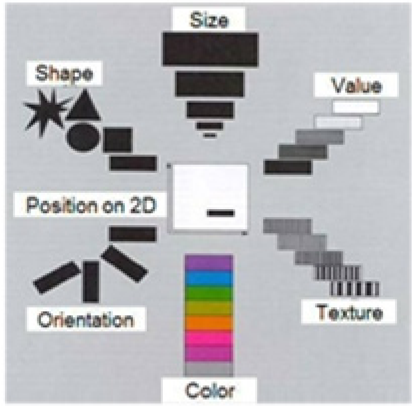

Fig. 2. Visual variables (source: J. Bertin)

Two variables: Size and Value give a clear selectivity regarding difference of numerical figures, whereas the other variables give same visibility. Fig.3 shows a comparison how each variable is identified its visibility and selectivity in a sample matrix. The Size of large or small scale and/or the Value of deep or light density give different perception in terms of quotative selectivity of numerical data. Then, large numbers shall be represented by large Size or dense Value, and small numbers shall be represented by small Size or light Value. This rule of representation is generally understandable for everyone who observes visualized data.

In case of the Texture, Color, Orientation and Shape, we can see a variable in a cell giving different "appearance", but there is no different "intense" of visibility. The Texture, Color, Orientation and Shape does not have "inherently" any a priori meaning of view strength. So, there is no natural meaning to evaluate figural data between red and green color, triangle and round shape, horizontally and vertically oriented bar..., etc. Furthermore, if we would like to use the Color or other variables to show numerical differences, additional definition shall be required which color represents which number. However, anyone does not have, a priori, any definitive combination of colors with numbers (no common sense that red is larger than green, for example). Therefore, such created definition requires obligatory observers unnecessary tasks: a conversion in mind from colors to relative data figures complying with the definition.

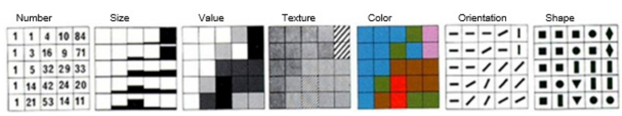

Fig. 3. Comparison of visual variables in matrix (source: J. Bertin)

J. Bertin's Orderable Matrix is a visualization of data which is a replacement of numeric data table by visual variable (Size and/or Value are concerned for this paper). And another characteristic is re-ordering of rows and columns in matrix. A matrix being permuted rows and columns improves visibility of the visualized table and makes it more readable view of concerned data set. Then, the Orderable Matrix let observer to grasp insights of data as a whole. An example is in Fig.4. Objective data comes

Proceedings of the International Cartographic Association, 1, 2017. This contribution underwent single-blind peer review based on submitted abstracts | https://doi.org/10.5194/ica-proc-1-55-2017 |

from a survey with respect to urban equipment for several areas where the people lives. And this tried to find out a categorization: city, suburban or country, depending on different equipment in each city in question. An Orderable Matrix just after replacing numerical figures to visual element is shown at (1) in Fig.4, and (2) is one another after re-ordering of columns and rows. By leading the Orderable Matrix of (2), the observer can easily and quickly remark categories of area: city, suburban or country by identification of which equipment make such categorization (highlighted by red frame). Areas categorized as city are C, H and K where we see properties: “high school”, “station” and “police station”. On the other hands, the county is represented by “only 1 class in school”, “no medicine” and “no water conveyance”. So, the Orderable Matrix makes observers relatively easily to see correspondences between township categories and urban equipment.

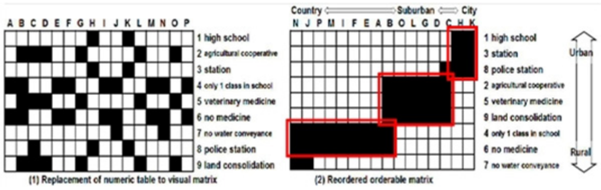

Fig. 4. Example of data handling in the Orderable Matrix (source: J. Bertin)

2.1 Application to risk information

A report published by a reinsurance company provides a global ranking for metropolitan cities potentially affected by 5 natural catastrophes: Earthquake, Tsunami, Wind storm, (or Typhoon), River flood and Storm surge. The ranking indicates Tokyo-Yokohama region as the first place as shown in Tab.1. About 57.1 million people is potentially affected. Some journals released articles, saying that Tokyo is most vulnerable cities or that Tokyo-Yokohama region is the riskiest place in the world to do business.

This statistic was made by a simple accumulation of potentially affected peoples by 5 natural catastrophes. In the same time, this report provides top 10 rankings for each of 5 events respectively in their annex as shown in Tab. 2. Undoubtedly, one summarized statistic table as Tab.1 is able to show data more simply than several tables as shown in Tab.2. There would be here a dilemma for presenter how to introduce the data: a brief idea with less figures of data or a deep understanding with much number of data. Naturally, it would be better to handle all available data and to see the data more accurately if we can see these data without inconvenient manner. The ranking tables for respective natural catastrophe are very interesting, but it shall take much time to get same idea of the summarized table as Tab.1, indeed. In other words, it is almost impossible to bring such summary in mind. We need normally to take a note which city is concerned which natural events, how serious damages are…, etc. One expectation to use the Orderable Matrix is to provide us an observation approach to see all necessary data without compromise. So, for an understanding of risk portfolio more comprehensively, the detailed ranking data sets is going to be analyzed by the Orderable Matrix.

Tokyo-Yokohama (JPN)	57.1
Manila (PHL)	34.6
Pearl-River Delta (CHN)	34.5
Osaka-Kobe (JPN)	32.1
Jakarta (IND)	27.7
Nagoya (JPN)	22.9
Kolkata (IND)	17.9
Shanghai (CHN)	16.7
Los Angeles (USA)	16.4
Tehran (IRN)	15.6

Tab. 1. Number of people potentially affected by 5 natural catastrophes (in million) (Source: Mind the risk)

Earthquake ranking			
	Metro area	Area (km²); population (mn)	People potentially affected(mn)
1	Tokyo-Yokohama (JPN)	16300 37.1	29.4
2	Jakarta (IDN)	11600 33.1	17.7
3	Manila (PHL)	2900 20.9	16.8
4	Los Angeles (USA)	14400 15.4	14.7
5	Osaka-Kobe (JPN)	13600 18.6	14.6
6	Tehran (IRN)	11000 15.1	13.6
7	Nagoya (JPN)	15600 11.6	9.4
8	Lima (PER)	2600 8.9	8.9
9	Taipei (TWN)	2100 8.1	8.0
10	Istanbul (TUR)	4100 11.5	6.4

Storm ranking			
	Metro area	Area (km²); population (mn)	People potentially affected(mn)
1	Pearl River Delta (CHN)	20600 42.4	17.2
2	Tokyo-Yokohama (JPN)	16300 37.1	14.1
3	Manila (PHL)	2900 20.9	12.6
4	Osaka-Kobe (JPN)	13600 18.6	7.8
5	Taipei (TWN)	2100 8.1	5.4
6	Shantou (CHN)	5200 10.0	5.1
7	Nagoya (JPN)	15600 11.6	4.3
8	Mumbai (IND)	2600 20.6	4.3
9	Chennai (IND)	1600 8.5	4.0
10	Tainan-Kaohsiung (TWN)	3100 5.1	4.0

River flood ranking			
	Metro area	Area (km²); population (mn)	People potentially affected(mn)
1	Pearl River Delta (CHN)	20600 42.4	12.0
2	Shanghai (CHN)	8000 17.6	11.7
3	Kolkata (IND)	3200 19.1	10.5
4	Jakarta (IDN)	10600 33.1	10.0
5	Delhi (IND)	5700 21.9	8.9
6	Tokyo-Yokohama (JPN)	16300 37.1	8.9
7	Bangkok (THL)	3500 9.5	7.1
8	Mexico City (MEX)	4500 19.6	6.1
9	Cairo (EGY)	3500 17.7	5.5
10	Tianjin (CHN)	2600 5.8	5.5

Storm surge ranking			
	Metro area	Area (km²); population (mn)	People potentially affected(mn)
1	Pearl River Delta (CHN)	20600 42.4	5.3
2	Osaka-Kobe (JPN)	13500 18.6	3.0
3	Mumbai (IND)	2600 20.6	2.6
4	Tokyo-Yokohama (JPN)	16300 37.1	2.3
5	Amsterdam-Rotterdam (NDL)	10600 5.4	1.8
6	Nagoya (JPN)	15600 11.6	1.7
7	Shanghai (CHN)	8000 17.6	1.4
8	Kolkata (IND)	3200 19.1	1.4
9	Ho Chi Minh (VNM)	2000 9.1	1.3
10	New York-Newark (USA)	11900 16.5	1.1

Tsunami ranking			
	Metro area	Area (km²); population (mn)	People potentially affected(mn)
1	Tokyo-Yokohama (JPN)	16300 37.1	2.4
2	Nagoya (JPN)	15600 11.6	2.4
3	Osaka-Kobe (JPN)	13500 18.6	1.8
4	Shantou (CHN)	5200 10.0	0.7
5	Kolkata (IND)	3200 19.1	0.6
6	Dhaka (BGD)	1800 12.9	0.4
7	Izmir (TUR)	1400 2.7	0.4
8	Jiaxing (CHN)	2700 3.1	0.3
9	Guayaquil (ECU)	900 2.4	0.3
10	Chennai (IND)	1600 8.5	0.2

Tab. 2. Number of people potentially affected for five perils (in million) (Source: Mind the risk)

2.2 Overlooking of data without omission

The data previously showing in Tab. 2 contains not only the number of people potentially affected by the different natural catastrophes, but also the number of people living in concerned agglomeration areas. Thus, this data can lead observers to see damage impacts by a percentage of people affected by 5 different natural catastrophes to total people living in cities.

Now, all data is integrated into one table as shown in Tab.3. Then, this numerical data table is replaced by an Orderable Matrix as Fig.5. The width of columns represents the number of people living listed city areas: areas with wider columns represent that number of people lives in that area are represented much people than ones with narrow width. By this arrangement, the observers can see a potential social exposure against natural catastrophes. The damageability (number of people affected by natural event divided by a total number of people living in concerned city) is presented by different sizes of block in each cell of matrix. So, Size of black blocks in cells show importance of damage due to each event: bigger blocks represent higher damageability. Finally, based on perception of such importance of cells, rows and columns are permuted to have a better view how different risks are in each area as shown in Fig.6. The arrangement of permutation is done by a principal rule of that similar view intensity of cell shall be gathered. Then, certain groups of cells are emergent in the Orderable Matrix. As shown in Fig.6, we see 4 groups which is highlighted in red frame and some exceptions shown in bleu. So, the variable of black blocks being initially scattered one part and other parts in the matrix, but now, these are re-ordered to be ready to read what are main drivers to produce characteristic of this data set and what correspondences are observed between natural catastrophes and cities.

City	Tokyo-Yokohama	Jakarta	Manila	Los Angeles	Osaka-Kobe	Tehran	Nagoya	Lima	Taipei	Istanbul	Pearl River Delta	Shanbu	Mumbai	Shanghai	Kolkata	Delhi	Bangkok	Mexico City	Tianjin	Cairo
Country	JPN	IDN	PHL	USA	JPN	IRN	JPN	PER	TWN	TUR	CHN	CHN	IND	CHN	IND	IND	THL	MEX	CHN	EGY
Population	37.1	33.1	20.9	15.4	18.6	15.1	11.6	8.9	8.1	11.5	42.4	10.0	20.6	17.6	19.1	21.9	9.5	19.6	5.8	17.7
Earthquake	29.4	17.7	16.8	14.7	14.6	13.6	9.4	8.9	8.0	6.4										
Wind storm	14.1		12.6		7.8		4.3		5.4		17.2	5.1	4.3							
River flood	8.9	10.0									12.0			11.7	10.5	8.9	7.1	6.1	5.5	5.5
Storm surge	2.3				3.0		1.7				5.3		2.6	1.4	1.4					
Tsunami	2.4				1.8		2.4					0.7			0.6					

Tab. 3. Potential cities affected by 5 natural catastrophes (in million) (based on the original data from Mind the risk)

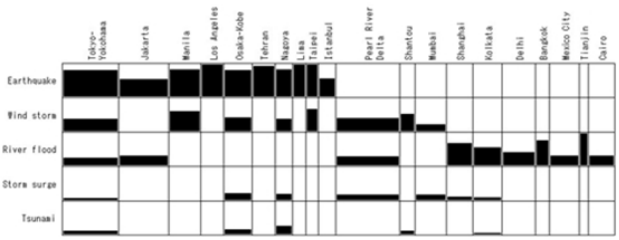

Fig. 5. Orderable matrix

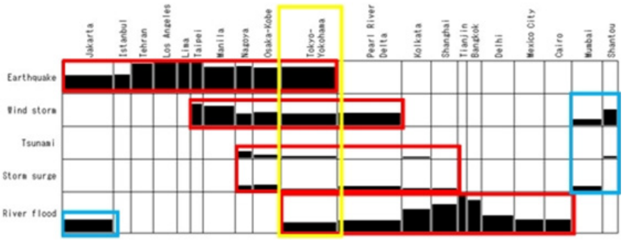

Fig. 6. Reordered orderable matrix

Firstly, we should understand that each different natural catastrophe does not cause same level of damageability to the cities. As a holistic level of view, it is understood that the Earthquake, Wind storm and River flood give more strong impact to the cities than Tsunami and Storm surge. And we see one combination of risk on Tsunami and Storm surge by grouping of cells. Both risks are shown in pair at same area in the matrix. It is understood as natural cause that areas near shoreline are exposed to marine related risks.

The index of 57.1 million potential victims in Tokyo-Yokohama area shown in Tab.1, which is told as the most hazardous city in the medias, has been decomposed by each objected natural catastrophe, and been compared with other city areas. The danger of Tokyo area is clarified how differently danger among different cities. Tokyo is an area exposed by the all catastrophes. Osaka and Nagoya follows situation of Tokyo. This should be understood how Japan is natural hazard prone area. This relative comparison in partial level view of analysis in the matrix could not be observed by Tab.1 which is a simple accumulation of number of potentially affected peoples for all catastrophes.

When we look at the matrix in detailed level, high hazard of Tokyo does not mean that the other areas are relatively secured from the catastrophes. Some of cities would have destructive damage due to certain perils. In case of that black blocks occupies totally in cells, this means almost 100% loss. For example, Los Angeles, Lima and Taipei should be concerned cities of total loss against earthquake. And river flood causes serious damage to cities: Kolkata, Shanghai, Jakarta, Delhi, Mexico City, Cairo Bangkok and Tianjin. The Orderable Matrix give a recognition comprehensively how serious each natural catastrophe causes in cities.

The awareness through analysis above mentioned might lead to think a hypothesis for next step to prevent damages due to concerned perils, not necessary to be fearful of the risk ranking. Further, it would be interesting to look for a complement data to be in blanks of table: Tab.3 which was left out from the original top 10 data set described in Tab.2. And this could allow to see exactly natural catastrophe exposures of each city.

2.3 Study on visualization

The data visualization by Orderable Matrix let us to observe risk exposure against natural catastrophes in some metropolitan cities as described above. However, since there are other visual variables to be able to use and also there are some other possibilities of re-ordering of rows and columns in the Orderable Matrix as suggested by J. Bertin, other ways of visual establishment can be available. A matter for study here is to see different nuance giving from other versions of the Orderable Matrix, even if observed and/or analyzed information does not changed result quality. Fig.7 is one variation to show same data of Fig.6. A difference of Fig.7 compare to Fig.6 is to introduce only 5 levels of different size of black block, rather than using totally matched between visual variable size and data figures.

This would allow to see a data set simply than before.

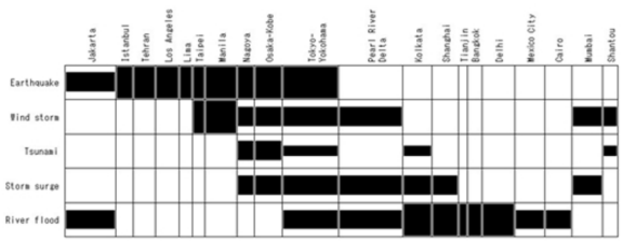

Fig. 7. Variation on Orderable Matrix, in case of introducing 5 level of visual variable of Size

Fig.8 is an Orderable Matrix with using a variable of Value (a gray scale gradation from black to white). And this is simpler version than the original because of exclusion of the notion about number of living people, then, the all widths of columns are same. We are purely focusing on damageability of cities.

With this version, another ordering of column could be suggested. For example, Jakarta at left side of the matrix could be at right side after Cairo as Fig.9. But again, there is no major influence on the analysis that we made before. Any data set coming from reality contains exceptions, therefore, it is possible to have always some difference of matrix view. However, one objective of using the Orderable Matrix is not to have one definitive re-ordering method, but to be a valuable tool for social communication.

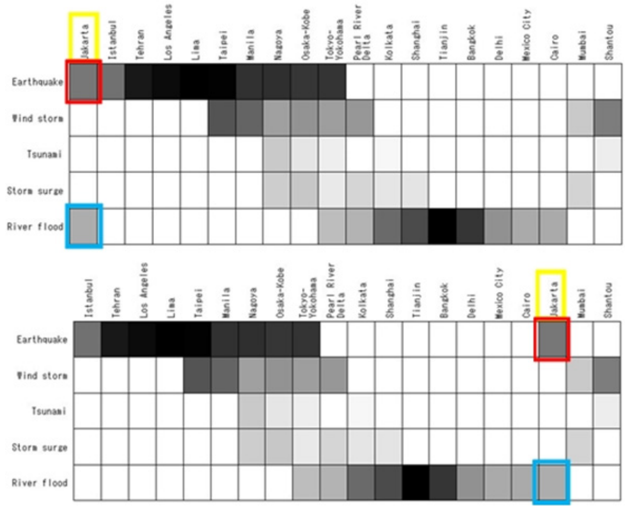

Fig. 9. Variation on Orderable Matrix: re-postponement of item

3. Conclusions

This paper has highlighted usability of the Orderable Matrix of Jacques Bertin which allows to analyze visually a data set regarding natural catastrophes. Actually, the Orderable Matrix has been able to show how danger concerned cities are. This could not be found out from the top 10 index of 5 different natural catastrophes which are presented separately by numerical data table. The panoramic view of all data sets arranged by Orderable Matrix allow us to see visually a portfolio of natural hazards instantaneously and comprehensively. For observation of it, there is no individual intervention of observers, whereas an reading of 5 different numerical data tables shall require certain tactics how to read and how to understand integrally all data, totally depending on observation methods with personal capability of data reading. So, the Orderable Matrix itself would be considered as a recognition tool to allow each observer being able to have same quality of understandings among them. This will be, in other words, to have same level of risk perception for all persons involved. Then, from this common view of risk, we could expect a decision commonly shared with everyone for next step of risk management.

Nowadays, large data can be handled easily with currently developed IT infrastructure, however, analysis method of such data is a major issue. The Orderable Matrix could solve for some aspect of this issue. Regarding natural hazard, there are other information such as historical disasters, recovery period from disasters, economical damage... etc., to be added to the Orderable Matrix that we worked in this paper. And the large Orderable Matrix might give principal elements to describe characteristics of large data set based on the panoramic visualization of data. To prove it, we need to examine this possibility of the Orderable Matrix with more complex and large data, and then, the usability of the Orderable Matrix will be reinforced again for social communication where we find always certain conflicts be- tween specialist and non-specialist. This is the radical question on the Orderable Matrix: how essential information shall be selected and how that information is arranged to show according to requirements of people who observe it. That would also be a problematic of cartography.

4. References

Fukazawa, N. (2005). Risk perception and human behaviors. Kobundo-shuppansha, 223p.

Fekete, JD., Dragicevic, P. and Perin, C. (2014) Revisiting Bertin Matrices: New Interactions for Crafting Tabular Visualizations. IEEE. p2082-2091.

Hirota, S., Masuda, S. and Sakagami, T. (2002). Physiologically described world of risk. Keikodaigaku-shuppankai, 253p. Neisser, U (1978) [translation by Kozaki, T]. Cognition and reality. Sceince-sha, 246p.

Bertin, J (1977). La graphique et le traitement graphique de l'information. Flammarion. p11, 252p. Bertin, J (1966). Sémiologie graphique. Paris : Mouton, 432p.

Kamei. K (2001). Management strategies of French company and risk management. Horitsubunka-sha, p512, 544p. Morita, T. (1982). Zu no Kigougaku. Heibonsha, 277p.

Omi, M., Yosizawa, T., Takao, A., Amari, K. and Kubo, E. Elements of risk and insurance. P82, 320p. de Saussure, F. (1992) Cours de linguistique générale. Paris : Payot, 520p.

Japan Meteorological Agency. http://www.jma.go.jp/jma/indexe.html

SwissRe. Mind the risk. (2014). http://www.swissre.com/library/expertisepublication/Mind_the_risk_a_global_ranking_of_cities_under_threat_from_natural_disasters.html. Accessed 26 October 2016.

Reuters. (2016). http://www.reuters.com/article/swissre-risks-idUSL5N0HD2RT20130918. Accessed 26 October 2016.

USA Today. (2016). http://www.usatoday.com/story/weather/2013/09/18/worlds-most-disaster-prone-cities-tokyo/2833667/. Accessed 26October 2016.

PML simulation software providers: AIR Worldwide, CoreLogic and RMS are commonly used. (AIR Worldwide