Wildlife Damage Estimation and Prediction Using Blog and Tweet Information

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

Shohei Fujise
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Wildlife damage estimation and prediction using blog and tweet information is conducted. Through a regressive analysis with the truth data about wildlife damage which is acquired by the federal and provincial governments and the blog and the tweet information about wildlife damage which are acquired in the same year, it is found that some possibility for estimation and prediction of wildlife damage. Through experiments, it is found that R^2 value of the relations between the federal and provincial government gathered truth data of wildlife damages and the blog and the tweet information derived wildlife damages is more than 0.75. Also, it is possible to predict wildlife damage by using past truth data and the estimated wildlife damages. Therefore, it is concluded that the proposed method is applicable to estimate and predict wildlife damages.**

Keywords*—*Wildlife damage; Blog; Tweet; Big data analysis; Natural language recognition

I. Introduction

Wildlife damage in Japan is around 23 Billion Japanese Yen a year in accordance with the report from the Ministry of Agriculture, Japan. In particular, wildlife damages by deer and wild pigs are dominant (10 times much greater than the others) in comparison to the damage due to monkeys, bulbuls (birds), rats. Therefore, there are strong demands to mitigate the wildlife damage as much as we could. It, however, is not so easy to find and capture the wildlife due to lack of information about behavior. For instance, their routes, lurk locations are unknown and not easy to find. Therefore, it is difficult to determine the appropriate location of launch a trap. In Kyushu, Japan, wildlife damage is getting large and is one of severe problems for farmers as well as residents in the districts near the mountainous areas. The federal and provincial agricultural management organizations in the districts are surveying the wildlife damages every year. It is time consuming task and requires large budget. Also, it takes almost two years. Therefore, it is hard to make a plan for wildlife damage controls. It would be helpful to estimate and predict wildlife damages with some other methods. Meanwhile, blog and tweet information can gather with some software tools. Furthermore, it would be possible to extract some valuable information relating to wildlife damages. The method proposed here is to estimate and predict wildlife damages by using blog and tweet information. It can be done immediately after the end of the Japanese fiscal year. Therefore, wildlife damage prevention plan can be created by the end of the Japanese fiscal year.

The following section describes the proposed method for wildlife damage estimation and predictions followed by experimental data. Then, concluding remarks and some discussions are followed.

II. Literature and Related Work

According to the West, B. C., A. L. Cooper, and J. B. Armstrong, 2009, "Managing wild pigs: A technical guide. Human-Wildlife Interactions Monograph"[1], 1–551, there are the following wild pig damages, Ecological Impacts to ecosystems can take the form of decreased water quality, increased propagation of exotic plant species, increased soil erosion, modification of nutrient cycles, and damage to native plant species [1]-[5]. Agricultural Crops Wild pigs can damage timber, pastures, and, especially, agricultural crops [6]-[9]. Forest Restoration Seedlings of both hardwoods and pines, especially longleaf pines, are very susceptible to pig damage through direct consumption, rooting, and trampling [10]-[12]. Disease Threats to Humans and Livestock Wild pigs carry numerous parasites and diseases that potentially threaten the health of humans, livestock, and wildlife [13]-[15]. Humans can be infected by several of these, including diseases such as brucellosis, leptospirosis, salmonellosis, toxoplasmosis, sarcoptic mange, and trichinosis. Diseases of significance to livestock and other animals include pseudorabies, swine brucellosis, tuberculosis, vesicular stomatis, and classical swine fever [14], [16]-[18]. There also are some lethal techniques for damage managements. One of these is trapping. It is reported that an intense trapping program can reduce populations by 80 to 90% [19]. Some individuals, however, are resistant to trapping; thus, trapping alone is unlikely to be successful in entirely eradicating populations. In general, cage traps, including both large corral traps and portable drop-gate traps, are most popular and effective, but success varies seasonally with the availability of natural food sources [20]. Cage or pen traps are based on a holding container with some type of a gate or door [21]. The method and system for monitoring the total number of wild pigs in the certain district in concern is proposed [22]. All the aforementioned system is not so cheap. It requires huge resources of human-ware, hardware and software as well. Also, it is totally time consumable task. Usually, it takes two years to finalize the total number of wild animals and wildlife damages. Therefore, it is hard to plan the countermeasures for the wildlife damages.

[1] www.berrymaninstitute.org/publications,

III. PROPOSED METHOD

A. Methods for Acquisition of Blog and Tweet Information Relating to Wildlife Damages

There are some sites which allow acquisition of tweet and blog information. Fig.1 (a) shows one of the tweet information acquisition sites while Fig.1 (b) shows one of the blog information acquisition sites. For the tweet information acquisition site (https://dev.twitter.com/rest/public/search), the Search API is not complete index of all Tweets, but instead an index of recent Tweets. At the moment that index includes between 6-9 days of Tweets. Therefore, tweet information has to be acquired within 6-9 days after the event of wildlife appearance. It required some information collection robots. These examples are http://blog.ritlweb.com/ for blog information collection while http://twitter.com/ is for tweet information collections.

(a)Tweet

(b)Blog

Fig. 1. Examples of the tweet and the blog information acquisition sites

B. Methods for Extraction of Wildlife Damage Information from the Acquired Blog and Tweet Information

It has to be done to extract wildlife damage related information from the acquired blog and tweet information. The following set of three parameters have to be extracted, (1) the area name, (2) the types of wildlife damages, (3) the date of the wildlife damage reported. In order to extract sets of information, "Chasen" of sentence structure and words analysis software tool is used. It is morphological analysis tool. The extracted words and sentences acquired from the twitter and blog data collection sites are input to the "Chasen". Then noun and the other part of speech can be extracted as shown in Fig.2.

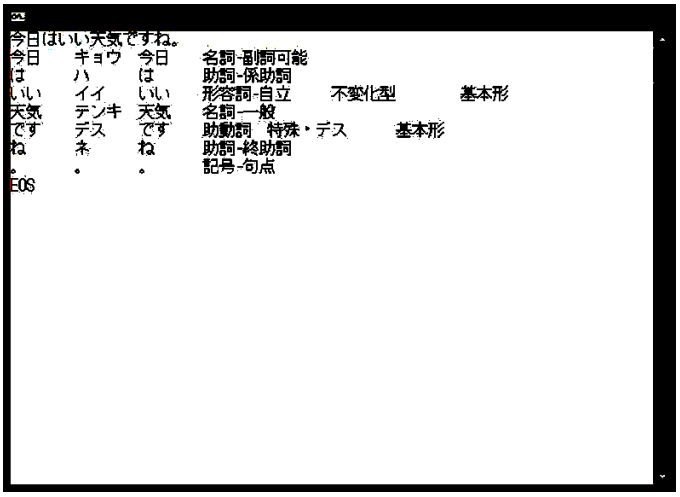


Fig. 2. Example of the screen shot of the Chasen analysis

The acquired sentence is "It is fine today" in Japanese and is appeared at the first line of the example. The first column of the second to the eighth lines "Today", "is" "Fine", "Weather", "it", and "is not it" show the words extracted from the acquired sentence. The second column shows their sounds while the forth column shows their part of speech. Thus, the words can be divided and extracted from the sentence together with their part of speech. Therefore, nouns can be extracted from the sentences. After that full text search is conducted to the extracted words.

Firstly, area names are extracted from the extracted words. In this regards, City name, Town name, and Village name in Kyushu provided by the federal and provincial governments are used in order to extract the area names. After that the names of the wildlife which is provided by the federal government of Agriculture, Forestry and Fishery ministry are extracted from the words. In this regards, combined words such as "prevention of bird damage" is recognized as the words of wildlife damage. The date of the tweet and blog information is easily extracted from the tweet and blob information because the information is dated information. Thus when, where, which wildlife can be extracted from the tweet and blog information.

C. Methods for Estimation of Wildlife Damage from the Acquired Tweet and Blob Information

The number of wildlife damage reports which are extracted from the acquired tweet and blog information in the year in concern must be proportional to the wildlife damages in that year. Therefore, linear regression would work for estimation of wildlife damage with the acquired tweet and blob information.

D. Methods for Prediction of Wildlife Damage Information from the Acquired Blog and Tweet Information

Based on the well known time series analysis method, it is possible to predict using the past wildlife damage. If the estimated wildlife damage with tweet and blob information is used for the wildlife damage in year in concern together with the past wildlife damage, then it is possible to predict future wildlife damage. In this regards, the following linear prediction is used for this,

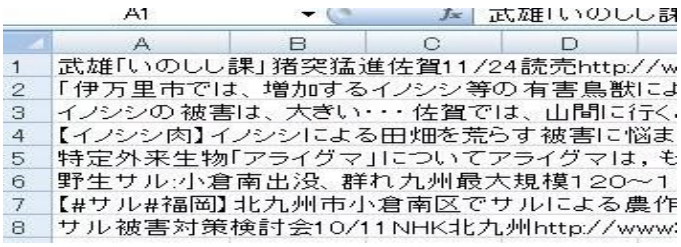
$$y - \bar{y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})y_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}(x - \bar{x}) \quad (1)$$

where x and y denote the past wildlife damage and the current wildlife damage, respectively. xbar and ybar denote mean of the past and the current wildlife damage, respectively.

IV. EXPERIMENTS

A. Examplessof the Acquied Blog and Tweet Information Relating to Wildlife Damages

One of the examples of the tweet and blog information relating to wildlife damage is shown in Fig.3 (a). Meanwhile, the extracted words of area names and the types of wildlife are shown in Fig.3 (b) while the results from the wildlife damage estimated from the acquired tweet and blog information is shown in Fig.3 (c), respectively. The summarized results of the number of wildlife damage which are reported by twitter and blog at every province, Fukuoka, Saga, Nagasaki, Ohita, Kumamoto, Miyazaki, and Kagoshima prefectures in Kyushu in 2013.



(a)Tweet and blob

(b)Area name and types of wildlife name

県名	鳥獣計	鳥類計		獣類計			
			カラス		イノシシ	サル	シカ
福岡	118	3	2	115	45	22	27
佐賀	82	8	1	74	45	5	19
長崎	103	2	2	101	58	12	17
熊本	18	0	0	18	3	0	4
大分	97	2	2	95	38	4	22
宮崎	50	2	2	48	18	6	24
鹿児島	0	0	0	0	0	0	0

(c)Example of the results of the number of wildlife damages (for every Provinces)

Fig. 3. Examples of the acquired tweet and blog information, the area name and the types of wildlife name as well as the summarized results from the wildlife damage in Kyushu in 2013

B. True Wildlife Damage Reported by the Regional Govermental Insititude of Kyushu Agricultural Management

True wildlife damage reported by the regional governmental institute of Agricultural Management in 2013 is shown in Table 1.

TABLE I. TRUE WILDLIFE DAMAGE REPORTED BY THE REGIONAL GOVERNMENTAL INSTITUTE OF AGRICULTURAL MANAGEMENT IN 2013

	Wildlife	Birds	Crow	Animals	Wild pig	Monkey	Deer
Fukuoka	91671	36306	19551	55366	35867	2578	7986
Saga	20486	6040	4559	14446	11023	1130	0
Nagasaki	28724	3171	2194	25553	23930	1	470
Kumamoto	45531	10398	8745	35133	28031	1311	4030
Ohita	21550	1078	749	20472	14474	1355	3229
Miyazaki	72978	4242	3363	68736	33396	7287	26066
Kagoshima	43950	8848	3219	35102	17070	2183	12878

The prefecture which shows the largest wildlife damage is Fukuoka followed by Miyazaki, Kumamoto, Kagoshima. Nagasaki, Ohita and Saga. The number of reports of wildlife damage, on the other hand, is shown in Table 2. The correlation coefficient between the total numbers of the reports and the total wildlife damage is just 0.013 as shown in Table 2.

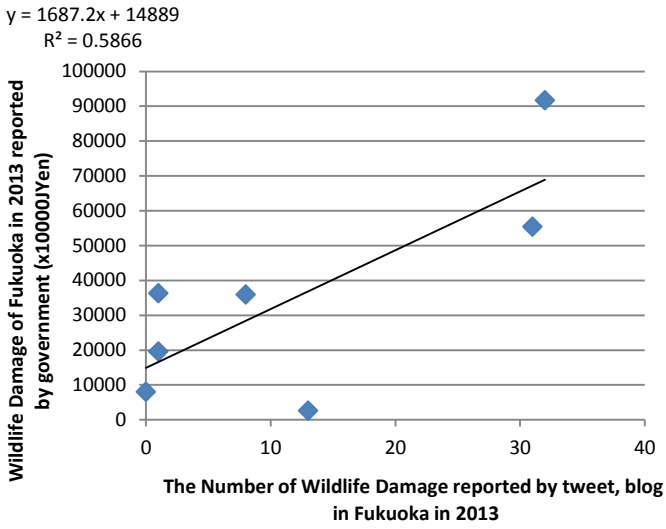
TABLE II. NUMBER OF REPORTS OF WILDLIFE DAMAGE AND TOTAL WILDLIFE DAMAGE IN KYUSHU IN 2013

	Fukuoka	Saga	Nagasaki	Kumamoto	Oita	Miyazaki	Kagoshima
Wildlife	32	7	42	12	29	8	1
Birds	1	0	0	1	0	1	0
Crow	1	0	0	0	0	0	0
Animals	31	7	42	11	29	7	1
Wild pig	8	6	35	2	3	0	0
Monkey	13	0	1	0	0	5	0
Deer	0	0	6	3	11	2	0
Wildlife	91671	20486	28724	45531	21550	72978	43950
No.of report	86	14	126	23	72	23	2

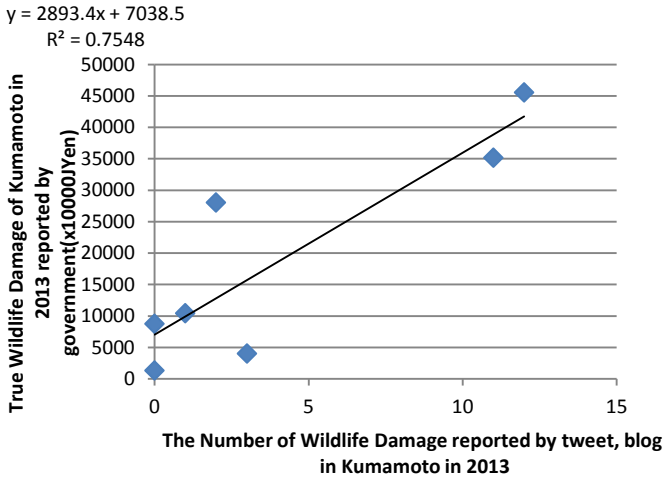
Although correlation coefficient is so poor, R=0.013, if the number of reports of wildlife damages of crow and birds, as well as monkey is deleted together with the number of report of Saga, Kumamoto and Kagoshima due to the fact that the number of reports are so small then the correlation coefficient between the total wildlife damage and the total number of the reports of wildlife damage through blog and tweet is increased R=0.538. Therefore, the relation between both is not so poor.

C. Estimation of Wildlife Damage from the Number of Reported Tweet and Blog for Every Province

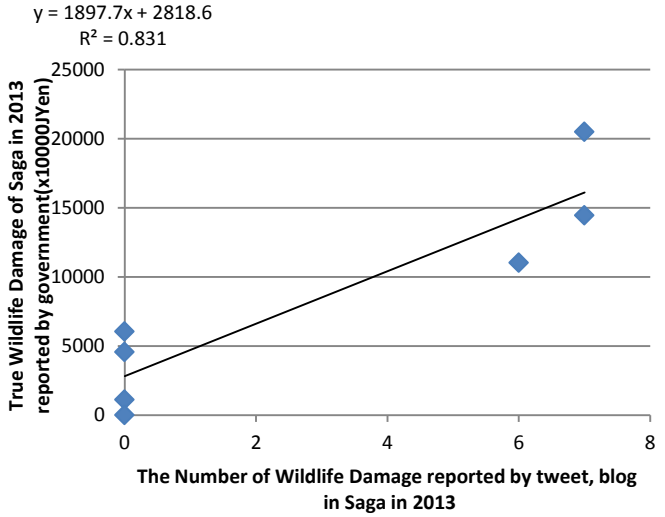
Through the linear regressive analysis, it can be done to estimate wildlife damages using the reported tweet and blog information. The results from the regressive analysis are shown in Fig.5. At the top left corners of the figures in Fig.4, there are regressive equations and the R^2 values. The R^2 values range from 0.5657 to 0.9693 while slope (gain) coefficients range from 607.17 to 30686. On the other hand, the number of reports of tweet and blog (Horizontal axis of the graphs in Fig.4) range from 1 to 42. The uncertainty of the regressive analysis is totally dependent to the number of reports. Therefore, the regressive analysis results of Saga, Kagoshima, Miyazaki are not so reliable. Then the ranges of the R2 values and gain coefficients are (0.5866 – 0.9693), and (607.17 – 2893.4), respectively.



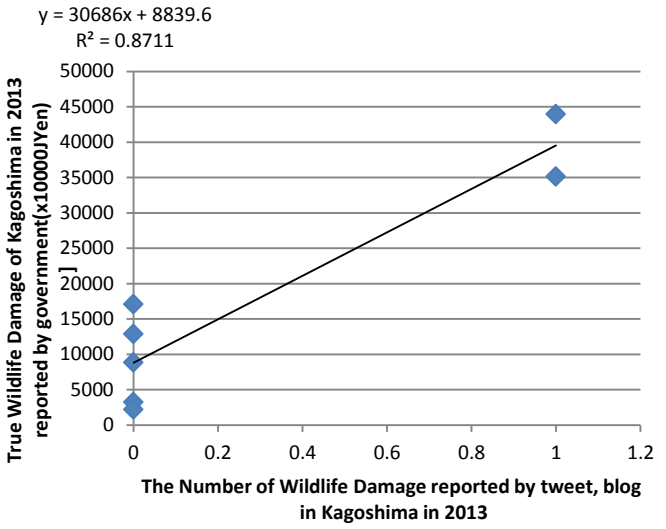
(a)Fukuoka



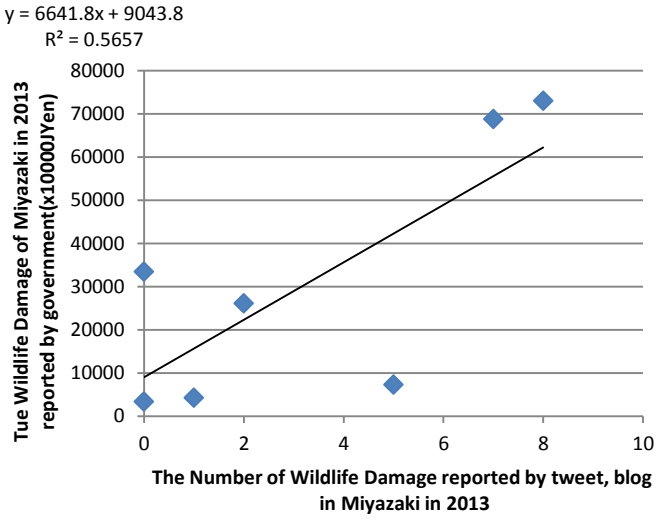
(b)Kumamoto



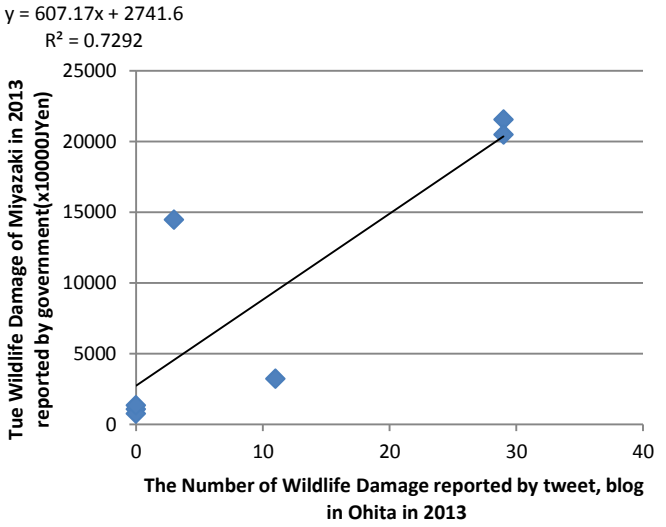
(c)Saga



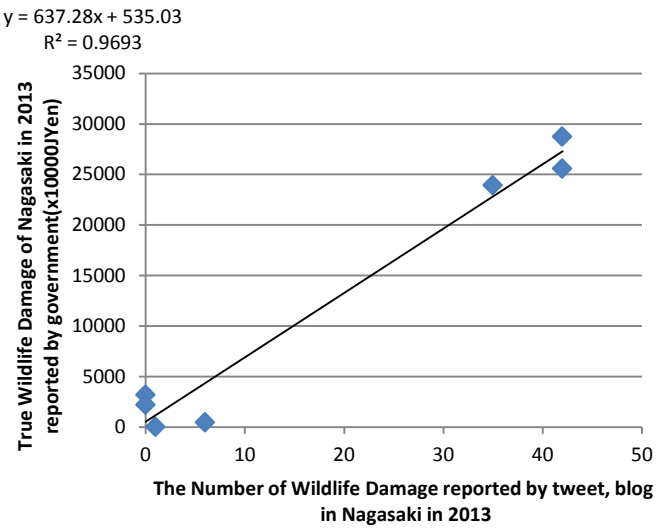
(d)Kagoshima



(e)Miyazaki



(f)Ohita



(g)Nagasaki

Fig. 4. Estimate wildlife damages for every province using the reported tweet and blog information

D. *Predictions of Wildlife Damage from the Number of Reported Tweet and Blog for Every Province*

The newest true wildlife damage data is 2014 which is provided by Kumamoto prefecture. There is no other prefecture of which true wildlife damage of 2014 is reported. Therefore, the wildlife damage of 2014 is predicted by using the past data of wildlife damage (2008 to 2013) based on the linear prediction which is expressed in equation (1). Table 3 shows the results from the predicted wildlife damage (in the second row of Table 3). The correlation between the wildlife damage from the true report of prefecture Kumamoto and predicted wildlife damage from the report of blog and tweet information is 0.996. By taking into account the compensation of mean and standard deviation of the predicted wildlife damage (adjusted), the difference between true wildlife damage and the predicted wildlife damage from the acquired blog and tweet information ranges from -1158 to 2944 in unit of 10,000 Japanese Yen.

TABLE III. PREDICTED WILDLIFE DAMAGE BY USING THE PAST DATA FOR 6 YEARS, 2008 TO 2013

Kumamoto	Wildlife	Birds	Crow	Animals	Wild pig	Monkey	Deer
True report	45531	10398	8745	35133	28031	1311	4030
Predicted	65270.8	14856	9520.9	52210.4	37813.5	4761.7	7181.1
Adjusted	44689.56	9399.2	5664.63	35547.28	25469.45	2333.19	4026.77
Difference	841.44	998.8	3080.37	-414.28	2561.55	-1022.19	3.23

From the relation between year and wildlife damage in Kumamoto in unit of 10,000 Japanese Yen, the wildlife damage can be calculated with the number of the tweet and the blog. Red colored number in Table 4 shows the calculated wildlife damage and the blue colored number indicates the predicted wildlife damage derived from the linear prediction with the true wildlife damage for five years (2008 – 2012) and the estimated wildlife damage in 2013. Through a comparison between true wildlife damage and the predicted one is approximately 6.0 %. Therefore, it is capable to predict wildlife damage in the next year with the past true wildlife damage reported by the local prefectural government and the relation between wildlife damage and the number of report by twitter and blog.

TABLE IV. COMPARISON OF THE WILDLIFE DAMAGES BETWEEN TRUE AND THE PREDICTION

Year	2008	2009	2010	2011	2012	2013	2014
True report	61468	70013	84516	54495	51975	45531	47235
True+Estimated	61468	70013	84516	54495	51975	58509.25	51000

Fig.5 shows the true and the predicted wildlife damages as a function of year. Therefore, it may say that wildlife damage in the next year can be predicted with the past true data of wildlife damage and the relation between the number of reports by twitter and blog.

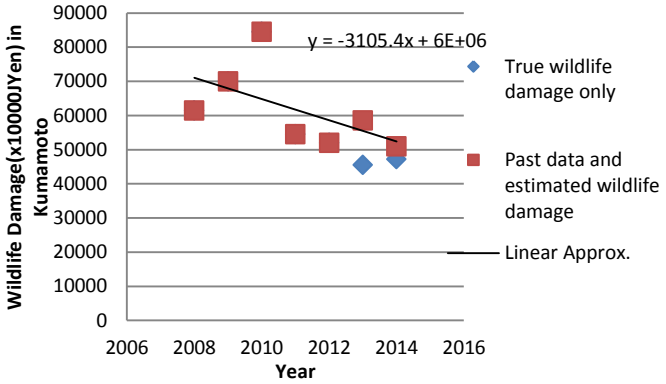


Fig. 5. True and the predicted wildlife damages as a function of year

V. CONCLUSION

Method for wildlife damage estimation and prediction using blog and tweet information relating to wildlife appearances is proposed in this paper. Through regressive analysis with the truth data about wildlife damage which is acquired by the federal and provincial governments and the blog and tweet information about wildlife damage which are acquired in the same year, it is found that some possibility for estimation and prediction of wildlife damage. Through experiments, it is found that R^2 value of the relations between the federal and provincial government gathered truth data of wildlife damages and blog tweet information derived wildlife damages is more than 0.75. Also, it is possible to predict wildlife damage by using past truth data and the estimated wildlife damages. Therefore, it is concluded that the proposed method is applicable to estimate and predict wildlife damages.

It is also found that the correlation between the wildlife damage from the true report of prefecture Kumamoto and predicted wildlife damage from the report of blog and tweet information is 0.996. By taking into account the compensation of mean and standard deviation of the predicted wildlife damage (adjusted), the difference between true wildlife damage and the predicted wildlife damage from the acquired blog and tweet information ranges from -1158 to 2944 in unit of 10,000 Japanese Yen. Therefore, future wildlife damage can be predicted by using the reports from blog and tweet information in some extent.

Further investigations are required for increasing the cases of wildlife damages for improving prediction accuracy.

ACKNOWLEDGMENT

The author would like to thank Prof. Dr. Hiroshi Okumura of Saga University for his valuable suggestions and comments for this study.

REFERENCES

[1] Patten, D. C. 1974. Feral hogs — boon or burden. Proceedings of the Sixth Vertebrate Pest Conference 6:210–234.

[2] Singer, F. J., W. T. Swank, and E. E. C. Clebsch. 1984. Effects of wild pig rooting in a deciduous forest. Journal of Wildlife Management. 48:464–473.

[3] Stone, C. P., and J. O. Keith. 1987. Control of feral ungulates and small mammals in Hawaii's national parks: research and management strategies. Pages 277–287 in C. G. J. Richards and T. Y. Ku, editors. Control of mammal pests. Taylor and Francis, London, England, and New York and Philadelphia, USA.

[4] Cushman, J. H., T. A. Tierney, and J. M. Hinds. 2004. Variable effects of feral pig disturbances on native and exotic plants in a California grassland. Ecological Applications 14:1746–1756.

[5] Kaller, M. D., and W. E. Kelso. 2006. Swine activity alters invertebrate and microbial communities in a coastal plain watershed. American Midland Naturalist 156:163–177.

[6] Bratton, S. P. 1977. The effect of European wild boar on the flora of the Great Smoky Mountains National Park. Pages 47–52 in G. W. Wood, editor. Research and management of wild hog populations. Belle W. Baruch Forest Science Institute, Clemson University, Georgetown, South Carolina, USA.

[7] Lucas, E. G. 1977. Feral hogs — problems and control on National Forest lands. Pages 17–22 in G. W. Wood, editor. Research and management of wild hog populations. Belle Baruch Forest Science Institute, Clemson University, Georgetown, South Carolina, USA.

[8] Thompson, R. L. 1977. Feral hogs on National Wildlife Refuges. Pages 11–15 in G. W. Wood, editor. Research and management of wild hog populations. Belle W. Baruch Forest Science Institute, Clemson University, Georgetown, South Carolina, USA. Kohei Arai, Preliminary Assessment of Radiometric Accuracy for MOS-1 Sensors, International Journal of Remote Sensing, Vol.9, No.1, pp.5-12, Apr.1988.

[9] Schley, L, and T. J. Roper. 2003. Diet of wild boar Sus scrofa in Western Europe, with particular reference to consumption of agricultural crops. Mammal Review 33:43–56.

[10] Whitehouse, D. B. 1999. Impacts of feral hogs on corporate timberlands in the southeastern United States. Pages 108–110 in Proceedings of the Feral Swine Symposium, June 2–3, 1999, Ft. Worth, Texas, USA.

[11] Mayer, J. J., E. A. Nelson, and L. D. Wike. 2000. Selective depredation of planted hardwood seedlings by wild pigs in a wetland restoration area. Ecological Engineering, 15(Supplement 1): S79–S85.

[12] Campbell, T. A., and D. B. Long. 2009. Feral swine damage and damage management in forested ecosystems. Forest Ecology and Management 257:2319–2326

[13] Forrester, D. J. 1991. Parasites and diseases of wild mammals in Florida. University of Florida Press, Gainesville, Florida, USA.

[14] Williams, E. S., and I. K. Barker. 2001. Infectious diseases of wild mammals. Iowa State University Press, Ames, Iowa, USA.

[15] Sweeney, J. R., J. M. Sweeney, and S. W. Sweeney. 2003. Feral hog. Pages 1164–1179 in G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, editors. Wild mammals of North America. Johns Hopkins University Press, Baltimore, Maryland, USA.

[16] Nettles, V.F., J. L. Corn, G. A. Erickson, and D. A. Jessup. 1989. A survey of wild swine in the United States for evidence of hog cholera. Journal of Wildlife Diseases 25:61–65.

[17] Davidson, W. R., and V. F. Nettles, editors. 1997. Wild swine. Pages 104–133 in Field manual of wildlife diseases in the southeastern United States. Second edition. Southeastern Cooperative Wildlife Disease Study, Athens, Georgia, USA.

[18] Davidson, W. R., editor. 2006. Wild swine. Pages 105–134 in Field manual of wildlife diseases in the southeastern United States. Third . Southeastern Cooperative Wildlife Disease Study, Athens, Georgia, USA.

[19] Choquenot, D. J., R. J. Kilgour, and B. S. Lukins. 1993. An evaluation of feral pig trapping. Wildlife Research, 20:15– 22.

[20] Barrett, R. H., and G. H. Birmingham. 1994. Wild pigs. Pages D65–D70 in S. Hyngstrom, R. Timm, and G. Larsen, editors. Prevention and control of wildlife damage. Cooperative Extension Service, University of Nebraska, Lincoln, Nebraska, USA.

[21] Mapston, M. E. 1999. Feral hog control methods. Pages 117–120 in Proceedings of the Feral Swine Symposium, June 2–3, 1999, Fort Worth, Texas, USA

[22] Kohei Arai, Indra Nugraha Abdullah, Kensuke Kubo, Katsumi Sugaw, Methods for Wild Pig Identifications from Moving Pictures and Discrimination of Female Wild Pigs based on Feature Matching Method, (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.7, 41-46, 2015

AUTHORS PROFILE

Kohei Aarai He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission-A of ICSU/COSPAR since 2008. He wrote 33 books and published 510 journal papers. He is Editor-in-Chief of International Journal of Advanced Computer Science and Applications as well as I ternational Journal of Intelligent Systsems and Applications. http://teagis.ip.is.saga-u.ac.jp/index.html