Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 145 (2016) 729 – 737

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference on Sustainable Design, Engineering and Construction

Automatic logical inconsistency detection in the National Bridge Inventory

Zia Ud Din, M.S[a,]*, Pingbo Tang, Ph.D[b]

[a]Del E. Webb School of Construction, Arizona State University, PO Box 870204, Tempe, AZ 85287-0204, USA.
[b]Del E. Webb School of Construction, Arizona State University, PO Box 870204, Tempe, AZ 85287-0204, USA.

Abstract*

Studies about the data quality of National Bridge Inventory (NBI) reveal missing, erroneous, and logically conflicting data. Existing data quality programs lack a focus on detecting the logical inconsistencies within NBI and between NBI and external data sources. For example, within NBI, the structural condition ratings of some bridges improve over a period while having no improvement activity or maintenance funds recorded in relevant attributes documented in NBI. An example of logical inconsistencies between NBI and external data sources is that some bridges are not located within 100 meters of any roads extracted from Google Map. Manual detection of such logical errors is tedious and error-prone. This paper proposes a systematical "hypothesis testing" approach for automatically detecting logical inconsistencies within NBI and between NBI and external data sources. Using this framework, the authors detected logical inconsistencies in the NBI data of two sample states for revealing suspicious data items in NBI. The results showed that about 1% of bridges were not located within 100 meters of any actual roads, and few bridges showed improvements in the structural evaluation without any reported maintenance records.

Peer-review under responsibility of the organizing committee of ICSDEC 2016

Keywords: Logical inconsistency; NBI database; Bridges condition;

1. Introduction

The estimated average annual bridge collapse rate in the United States is between 87 and 222 with an expected value of 128 [1]. These bridge failures result in fatalities, injuries, and causing enormous economic loss. These

* Corresponding author. Tel.: +1-480-558-6233; fax: +0-000-000-0000 .
E-mail address: zziauddi@asu.edu

Peer-review under responsibility of the organizing committee of ICSDEC 2016
doi:10.1016/j.proeng.2016.04.093

disasters continue to occur despite the fact that the conditions of all 607,380 national bridges in the United States — are inspected at regular intervals not to exceed twenty-four months with some exception with FHWA approval, and the bridge condition is recorded in the National Bridge Inventory (NBI) [2,3]. The NBI data is used to identify and prioritize which bridges are in greatest need of preventive maintenance, rehabilitation, or replacement [4] [3]. NBI data provides the basis for the Federal Highway Administration's (FHWA) decisions about bridge safety and the resource allocations made to keep bridges safe. Unfortunately, unexpected structural failures of bridges indicate that that a lot of space exists for further improving the quality and reliability of the NBI data.

Studies about the data quality of National Bridge Inventory (NBI) have found flaws, oddities, and omissions in NBI [5] despite the fact that Federal Highway Administration (FHWA) has developed some error checkers. The state department of transportations (DOTs) are also required to check the data for errors before submittals [6]. These NBI data checkers can conduct cross checks, null checks, safety checks, and follow-up checks, etc. These checks ensure the data entered according to the coding guide [6,7]. However, some logical inconsistencies like wrong spatial coordinates of bridges, are not detectable using such data checkers. Moreover, the data produced from different inspection and data processing procedures can conflict with each other even engineers strictly follow the standard NBI procedures of generating bridge inspection reports[9]. For example, at the original data level, calibration errors occur when sensors drift over time; at the data processing workflow level, uncertainties of selected data processing parameters for filtering and analyzing data can propagate across data products produced along the workflow. In many cases, uncertainty accumulations are unavoidable due to the limitation of the measuring systems, missing data due to sensor malfunction or human errors, extra data due to duplication of data collection efforts [10]. Existing data quality checkers developed by the federal and state agencies cannot ensure comprehensive scrutiny of most of the uncertainties and logical inconsistencies due to original data errors and uncertainty accumulations.

In this paper, we developed a "Hypothesis testing" framework for detecting logical inconsistencies between NBI and external data sources. Specifically, we evaluated three hypotheses: 1) bridges should be on roads; 2) bridge condition should not improve without maintenance; 3) similar bridges (based on age, material, and traffic) in similar areas should show similar conditions. This paper uses NBI data of two states as a test case and discusses bridges identified by this hypothesis-testing approach as potentially logical inconsistencies in the data. We expect that the findings through such hypothesis-testing studies will help bridge maintenance agencies to quickly filter out the potential problems in bridge inspection data and take measures to resolve potential inconsistencies and improve the resource allocation through optimal utilization of labor and resources.

2. Literature Review

Data are discrete objective facts, which are used to make informed decisions. Therefore, data quality affects the soundness of the decisions [9,10]. The first sub-section below focuses on reviewing general spatiotemporal data quality checking in various engineering domains. The second sub-section focuses on existing data quality checking methods designed in the domain of civil engineering, some of which are specifically for NBI data quality checking.

2.1. Quality checking and uncertainty analysis of spatiotemporal data

Data quality is the reliability of data to inform and evaluate decisions. Data of poor quality would not be suitable for the intended purpose [13]. Therefore, in today's world of massive electronic datasets, data quality problems can create significant economic and political inefficiencies [14]. Frequently mentioned dimensions of data quality are accuracy, reliability, consistency, precision, usefulness, timeliness, fineness, understandability, conciseness, and usefulness [15]. Researchers analyzed data quality along all these dimensions using data mining methods, which are algorithms for discovering latent knowledge, intelligible patterns, and meaningful insights from data [16]. Data mining draws on concepts from statistics, machine learning, database systems, and high-performance computing to accomplish its tasks. Another subset of data mining is spatiotemporal mining; it extracts spatial, temporal, and spatiotemporal relationships or other potentially useful patterns from datasets [17]. Spatiotemporal datasets have spatial and temporal components. The spatial component defines a universal reference space for all objects, usually geographical location on Earth's surface (indexed by latitude and longitude), whereas the temporal component is time-series data in which successive values in a series represent measures over time for a spatial location [18].

To check the quality of data of civil infrastructures, Buchheit et.al [19] devised a two-level solution to check the quality of data. At the first level of this two-level solution, the proposed solution detects outliers at an aggregated data level (e.g., The NBI data of a particular state). At the second level of the approach, the proposed solution explores data further at an individual data element level (e.g., Individual attributes present in the NBI data). After the identification of causes of the outliers, data cleaning can improve the quality of the data to a level that is appropriate for its intended uses. However, Buchheit's model focuses on identifying inconsistencies such as missing data and calibration errors, but it does not address the problems of data in spatial domain – for example, bridges with suspicious spatial coordinates - and temporal areas – for example, bridge with upgraded rating without any improvement activity on publically available record. Therefore, it cannot check logically inconsistent data along spatial and temporal dimensions. Hence, detecting and correcting of such inconsistent data is necessary to make data reliable.

The set of data points that are considerably different than the remainder of the data may be anomalous or outliers [20]. Identification of outliers can lead to the discovery of hidden but useful knowledge [21]. Methods for detecting the outliers include statistical methods, such as clustering, pattern-based detection, and association rules [22]. Statistical methods measure mean, standard deviation or quartile ranges, and conduct regression analysis of data [23]. Moreover, clustering method tries to combine the data into groups in such a way that the intra-cluster similarity is maximized while the inter-cluster similarity is minimized [20,21]. For example, in NBI data, bridges with similar construction materials can be clustered into groups. Pattern-based methods use patterns in the data set to identify records that do not conform to the pattern. For example, in NBI data, bridges showing different condition as compared to their contemporaries and neighbors. While, association rules find dependencies between values in a record. For example, a particular construction material's performance in a cold region can show some association between construction material, and temperature and any record that does not follow the rule definition is a suspect [26].

Engineers collect data about the condition of bridges and document their performance at regular intervals not exceeding twenty-four months with exceptions of some bridges with the FHWA approval and record bridge condition data in NBI [2,3]. The decision-makers rely on the NBI database to identify and prioritize the needs for maintenance, rehabilitation, replacement of the parts or the whole bridge [4]. Therefore, high-quality monitoring data are crucial for informed decision making, design analysis, and bridge performance studies [10].

2.2. NBI data quality checking

Bridge management agencies have been exploring ways of improving bridge data quality for reliable maintenance planning. Therefore, FWHA suggests checks on the NBI data to check the data consistency as a part of data quality assurance Program [27]. Likewise, state DOTs are also working to improve the data quality using visualization and data quality checking rules. For Example, the New Mexico Department of Transportation has developed a software application (Special Application – Bridge Information System or SABIS) to process NBI source files available from the Federal Highway Administration website [28]. The software populates state-specific databases and provides a simplified means for reviewing detailed information on public bridges in the inventory and for generating reports of commonly requested information. Another example is Delaware DOT's NBI Database Quality Assurance Program called NBIchk [29].

None of the programs mentioned above capture the concept of logical testing and hypothesis checking for detecting potential outliers. Therefore, the outlier detection method presented in this paper is a step toward systematical logical inconsistency analysis to point out the potential bridge data quality problems in NBI based on logics.

3. Methodology

We used geospatial processing and spreadsheet programs to detect the potential logical problems in spatial and temporal domains of the NBI data. The geospatial processing program, ESRI's ArcMap, has spatial statistics resources suitable to find spatial outliers. We used two consecutive years' NBI data of Oklahoma and Texas to identify the temporal outlier bridges.

3.1. NBI data preparation

The procedure for analyzing the NBI data along two dimensions: spatial and temporal dimensions is shown in Figure 1. The NBI data is available to public in American Standard Code for the Information Interchange (ASCII) text files format. In order to perform temporal analysis, we imported the ASCII text files into the spreadsheet program (Microsoft Excel), by following the procedure described on FHWA website's NBI ASCII Files webpage [30]. Details of the temporal analysis are available in the following section. In order to conduct the spatial analysis, we imported the NBI data in shapefile format from the Research and Innovative Technology Administration (RITA) website into ArcGIS 10.3 program [31]. The spatial data contains bridges as point features for the year 2014 and the query function in ArcMap 10.3 extracts the bridges of a selected state/region.

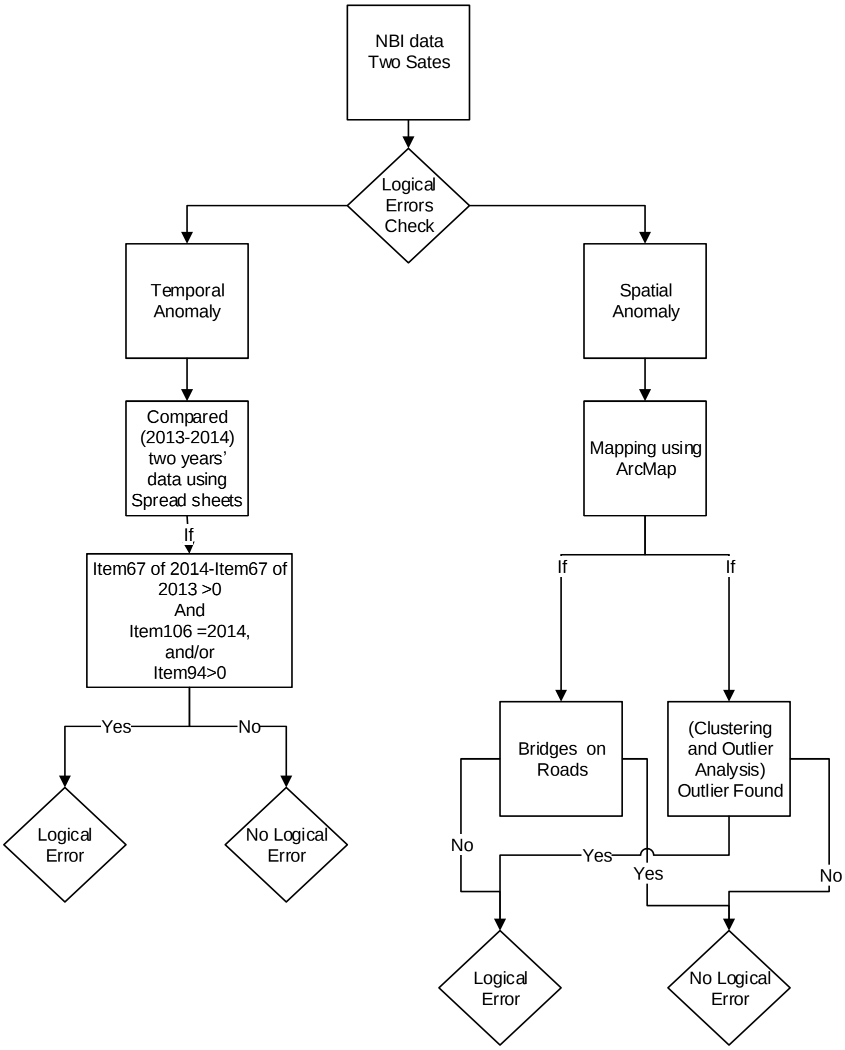


Fig. 1. Flow diagram for logical inconsistency checking

3.2. Determining spatial outliers

First, we plotted the bridges and roads - primary, secondary, and county roads - of Texas and Oklahoma using ArcMap 10.3. The road shapefiles are available as Topologically Integrated Geographic Encoding and Referencing (TIGER) lines on the www.Data.Gov website [32]. The second step was creating a buffer of 100 meters width on both sides of the roads to check if bridges are located within 100-meter distance from the closest road. The third step was using the erasing tool of ArcMap10.3 to remove all the bridges those were in the buffer zone. As a result, the remaining bridges which were not within the buffer zone appeared as a new layer. Only the bridges that were falling outside the buffered polygons were copied to the output feature class [33]. Then off-the-road bridges were exported as a KML file to visualize on Google Earth.

3.3. Determining temporal outliers

The NBI contains data from more than 20 years of bridge inspections. To determine temporal outlier, we imported the 'txt' files of the two states from the NBI database into MS Excel. We compared the Structural Evaluation ratings (NBI Item 67) of the year 2014 to the year 2013 and when the difference was greater than zero. Then, we checked two more data columns - Year of Reconstruction (NBI Item 106), and Bridge Improvement Cost (NBI Item 94) - in the NBI database. The assumption of temporal outlier detection is that when reconstruction activity and improvement cost for a bridge is not reported for two consecutive years (the year 2013 and 2014 in this case), then the structural evaluation of the bridge should remain unchanged or downgrade because of bridge aging. Such bridge is an anomalous entry in the database.

3.4. Determining spatial outliers

Bridge condition deteriorates due to average daily traffic and climate conditions (precipitation, frequency of deicing, temperature range, and freeze-thaw cycles). The researchers think that bridge segments made of the same materials, located in the similar environment have more or less similar ratings provided other factors (e.g., age, maintenance period, resource assignment to reconstruction or maintenance) are roughly the same [34]. To find spatial outliers in bridge condition ratings, we downloaded the NBI data files in "shapefile" format from the Bureau of Transportation Statistics website [35] and used ESRI's ArcMap 10.3 to perform spatial analysis using "Cluster and Outlier Analysis (Anselin Local Moran's I)" tool. The tool identifies spatial clusters of bridges with high or low bridge condition ratings, and it identifies spatial outliers, the bridges that have different condition ratings that of neighbors. We performed four types of spatial analysis. The first analysis enables us to find structural evaluation outliers by observing the bridges, which have different condition rating as compared to their neighbors. The second spatial analysis finds outliers when clusters are formed based on bridge deck condition. All of the bridges carrying inconsistent deck condition rating with their neighbors could be perceived as "outliers." The third analysis identified the spatial distribution of bridge clusters based on superstructure condition rating. The bridges that do not follow the clusters and possess inconsistent deck ratings with their neighbors are outliers. Last spatial analysis is performed for scour critical ratings. The bridges showing different scour critical ratings as compared to their geographic neighbors are considered as outliers. The outlier bridges with low condition rating are low-high outliers because neighboring bridge clusters are of higher bridge condition values. Similarly, outlier bridges having high condition rating value are high-low outliers because neighboring bridge clusters are lower bridge condition values.

4. Results and Discussion

4.1. Potential spatial outliers in NBI data of Texas and Oklahoma

Although the NBI database contains bridge location information, its accuracy is unverified. In many cases, spatial coordinates are not provided or incorrect. When we conflated the bridges with a road network on ArcMap, 28 bridges in Texas and 15 bridges in Oklahoma were located at longitude-latitude 0-0. In these two states, we found

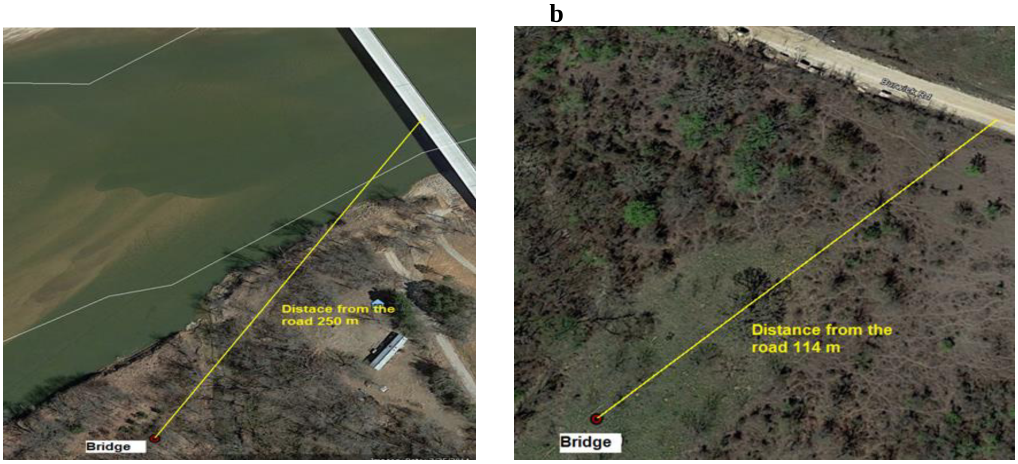
471 bridges with suspicious spatial coordinates that were inconsistent with the location of roads extracted from Google Maps, and consequently none of the bridges were within 100 meters of any roads on Google Maps. These bridges deviated from 101 to 1235 meters from the nearest road. Table 1 highlights information regarding bridges with suspicious spatial coordinates in Texas and Oklahoma.

Table 1. The bridges with suspicious spatial coordinates

Deviation from the nearest road (m)	Number of bridges in Texas State	Number of bridges in Oklahoma State
>1000	3	1
500 - 1000	42	11
250 - 500	80	41
100.1 - 250	190	93

In Figure 2, two bridges on Google Earth show that there are no roads within a 100-meter proximity of the bridges. For example, in Figure 3(a), a bridge on Google Earth is located 250 meters away from the bridge reported in NBI database, and similarly, Figure 3(b) shows that a bridge on Google Earth is located 114 meters away from the bridge reported in the NBI database. The inconsistent spatial coordinates may mislead inspectors when they inspect bridges. The consequences may be serious when funds' allocation for maintenance may be misappropriated due to spatial outliers. The NBI data of both states contains 43 bridges in the year 2014 which are reported with zero spatial coordinates showing that the NBI data quality checkers are not effective as they cannot capture such discernable inconsistencies in the database.

a **b**



(a) A bridge 250 meters off-the-road (b) A bridge114 meters off-the-road

Fig. 2. Examples of the bridges located off-the-road

4.2. Spatial outliers of condition ratings

4.2.1. Spatial outliers - based on structural evaluation of concrete bridges – Item 67

We observed that 3,034 (5.7% of the total bridges) bridges in Texas and 883 (3.8% of the total bridges) bridges in Oklahoma possess different structural evaluation (NBI Item 67) as compared to their adjacent bridges. Further analysis showed that 1,158 and 361 bridges in Texas and Oklahoma, respectively, have a lower structural evaluation than the neighboring bridges. Moreover, 1,876 and 522 bridges have a higher structural evaluation as compared to their neighbors in Texas and Oklahoma. There can be several reasons for these outliers ranging from staff training to the equipment they use for taking observations of different bridge components: the deck, substructure, superstructure, average daily traffic, etc. These components are the basis for determining the structural evaluation of a bridge. In the spatial outlier analysis, we found that outliers which constitute a fraction of total bridges should be checked to get an idea why these bridges carry different structural evaluation ratings as compared to their neighbors.

4.2.2. Spatial outlier - based on superstructure condition of concrete bridges - Item 59

The superstructure is the portion of the bridge that supports the deck and consists of beams or girders that may be constructed of timber, concrete or steel, and the bearings that connect the superstructure to the substructure. In this analysis, we focused on concrete bridges with reported superstructure ratings to find spatial outliers. In Texas and Oklahoma, the bridges having superstructure rating different from the adjacent bridges are 1,506 (14.7% of the total concrete bridges) and 189 (6% of the total concrete bridges), respectively. Of them, 820 and 113 bridges in Texas and Oklahoma, respectively, have high superstructure rating as compared to the adjacent bridges. However, 686 and 376 bridges in Texas and Oklahoma, respectively, have lower superstructure rating as compared to the neighboring bridges. We considered the bridges with inconsistent superstructure rating as outliers, and outliers may be due to inconsistencies in data collection practice. Figure 3(a) shows bridges 'A' and 'B' constructed in 1930 and 1938, respectively, but the superstructure rating of the relatively newer bridge (B) is 5 while the older bridge (A) has 7. The both bridges are only 3,190 meters apart from each and subject to almost similar environmental conditions. A similar example in Figure 3 (b) shows that bridges A and B constructed in 1943 and 1938, and they have superstructure ratings 7 and 3, respectively. These examples indicate few bridges have inconsistent conditions with their contemporaries and geographical neighbors. Therefore, further investigation is necessary to understand possible reasons for outliers which may be the skill level of inspectors, and equipment used, but we do not have that capability to get such data now.

a b

Fig.3. Typical superstructure condition rating outliers

4.2.3. Spatial outlier - based on substructure condition - Item 60

Substructure (piers and abutments) transfers structural load from the superstructure to the foundation. The structural evaluation of a bridge is obtained from the lowest rating of the superstructure, annual daily traffic, inventory rating, and substructure. Therefore, any amount of uncertainty is the assessment of the substructure, can have a cascading effect on the structural rating calculation and higher order decisions like structurally obsolete and/or functionally obsolete bridge [5]. In Texas and Oklahoma, 873 (2.6%) and 1,253 (7.8%) bridges have substructure rating different from the nearby bridges. Of them, 441 in Texas and 746 in Oklahoma have a higher superstructure rating as compared to the adjacent bridges and 432 in Texas and 507 in Oklahoma has a lower superstructure rating as compared to their adjacent bridges.

4.2.4. Spatial outlier - based on scour critical bridges –Item 113

Scour is the removal of soil from around the bridge abutments or piers due to water flow. Scour is the primary cause of bridge failures in the United States [37][34]. The authors filtered the bridges crossing over the waterways from the database and then checked the internal consistency of the scour critical ratings of bridges in Texas and Oklahoma. There are 10,801 and 1,903 bridges having different scour critical ratings as compared to their neighboring bridges in Texas and Oklahoma, respectively. Moreover, we found that 7,524 bridges are better in scour critical rating as compared to their adjacent bridges, and 3,277 bridges have a lower scour critical rating than their neighboring bridges in Texas. Similarly, in Oklahoma, 1,081bridges are surrounded by the bridges with lower scour critical rating (lower scour critical means more affected by scouring) and 822 bridges are the neighbors of the bridges having higher scour critical rating (higher scour critical means lesser affected by scouring).

4.2.5. Spatial outlier - based on deck condition - Item 58

The surface of the bridge used as a roadway is called bridge deck. The construction material of the deck may be concrete, or wood which is mostly covered with a pavement material like asphalt concrete [38]. The engineers record the condition of bridge deck elements to assess the top surface deterioration [36]. Our analysis shows that in Texas 1,568 and in Oklahoma 768 bridges have an inconsistent rating of the deck from the neighboring bridges. Moreover, in Texas 785 bridges have a higher rating as compared to the adjacent bridges, and 783 bridges have a lower deck condition as compared to the adjacent bridges.

4.3. Temporal outlier of structural evaluation of bridges

Time plays an important role in the deterioration of bridges. However, sometimes the condition rating of bridges does not follow the trend. Such bridges which show improvement in their structural evaluation without recorded reconstruction and improvement activities in the previous or current year are considered as temporal outliers. For the purpose of detection of anomalous rating, we analyzed the structural evaluation of bridges for the year of 2013 and 2014. We found that 37 bridges in Oklahoma and 669 bridges in Texas have improved structural evaluation as compared to the structural evaluation of bridges reported in 2013, but they have no reconstruction and improvement cost allocations on record in 2013 and 2014. This analysis will give the opportunity to find problems with data collection and processing. The structural evaluation value comes from the lowest code obtained from the superstructure and substructure rating, annual daily traffic, inventory and/or culverts rating. Therefore, any outlier bridges due to inconsistent structural evaluation would make all NBI item ratings unreliable from which it is derived.

5. Conclusion

In this research, we presented that existing data quality checking programs proposed by the Federal Highway Administration and state department of transportations focus on the code checking and ensuring the conformance to the NBI guidelines and do not capture the logical inconsistencies in the NBI database. Examples of inconsistencies are 1) bridges are reported with inconsistent spatial coordinates, and 2) condition of bridges improving without reported improvement works. We analyzed the NBI data of Texas and Oklahoma and found that 461 bridges are spatial outliers due to inconsistent spatial coordinates of these bridges with Google Earth. The inconsistent coordinates may mislead the bridge inspectors and maintenance operations in future. Furthermore, the results showed that 37 bridges in Oklahoma and 669 bridges in Texas have improved rating without any improvement and reconstruction activity reported in the NBI database. Our analysis of the NBI data also shows that many bridges have different condition ratings as compared to the neighboring bridges in Texas and Oklahoma. For example, the Structural Evaluation of 3,917 bridges, deck rating of 2,336 bridges, superstructure rating of 1,695 bridges, substructure rating of 2,126 bridges, and Scour Critical rating of 12,704 bridges are spatial outliers. By finding such outlier bridges, quality of data will improve, and the Federal Highway Administration may perform rechecking on about 10% of the total bridges only. It will reduce the cost of rechecking and improve the quality of the data. Further analysis of the logical inconstancies could show counties having a larger number of suspicious bridges. The possible reasons of the inconsistent data could be the human factor; the training of staff, technological factor; error produced due to equipment use or workflow related; the collection, transfer, and processing. There is a need to study the overall process of data collection and processing to understand the effect of uncertainty accumulations in the database. Such uncertainties can be caused by human factors and technological limitations; can change the decision, which can be serious for bridge maintenance and operations.

6. Reference

[1] W. Cook, P. Barr, and M. Halling, “Bridge Failure Rate,” *J. Perform. Constr. Facil.*, vol. 29, no. 3, p. 4014080, Dec. 2013.

[2] L. W. Luther and S. F. Sundlof, *National Bridge Inspection Standards Regulations (NBIS)*, vol. 69, no. 239. 2004, pp. 15–35.
[3] "23 CFR 650.311 - Inspection frequency. | US Law | LII / Legal Information Institute." [Online]. Available: https://www.law.cornell.edu/cfr/text/23/650.311. [Accessed: 22-Dec-2015].
[4] "Public Disclosure of National Bridge Inventory (NBI) Data - National Bridge Inventory - Bridge Inspection - Safety - Bridges & Structures - Federal Highway Administration." [Online]. Available: https://www.fhwa.dot.gov/bridge/nbi/20070517.cfm. [Accessed: 28-Jul-2015].
[5] W. Dekelbab, A. Al-Wazeer, and B. Harris, "History lessons from the national bridge inventory," *Public Roads*, vol. 71, no. 6, p. 30, 2008.
[6] "2015-call-for-nbi-update.pdf." .
[7] Federdral Highway Adminstration, "Recording and Coding guide for the Strcutre Inventory and Appraisal of the Nation's Brdiges," Washington, D.C. 20590, 1995.
[8] F. H. Administration, "Bridge & Structures," *NBI Submittal File Checks*. .
[9] T. Maheswaran, J. G. Sanjayan, D. L. Dowe, and P. J. Tan, "MML mixture models of heterogeneous poisson processes with uniform outliers for bridge deterioration," in *AI 2006: Advances in Artificial Intelligence*, Springer, 2006, pp. 322–331.
[10] R. B. Buchheit, "Vacuum: Automated Procedures for Assessing and Cleansing Civil Infrastructure Data," Carnegie Mellon University, Pittsburgh, 2002.
[11] S. Newell, M. Robertson, H. Scarbrough, and J. Swan, *Managing Knowledge Work and Innovation, 2nd Edition*. Palgrave Macmillan, 2009.
[12] T. H. Davenport and L. Prusak, *Working Knowledge: How Organizations Manage What They Know*. Harvard Business Review Press, 2013.
[13] W. E. Winkler, "Methods for evaluating and creating data quality," *Inf. Syst.*, vol. 29, no. 7, pp. 531–550, Oct. 2004.
[14] A. Karr, A. Sanil, and D. Banks, "Data quality: A statistical perspective," *Stat. Methodol.*, no. 151, 2006.
[15] Y. Wand and R. Y. Wang, "Anchoring Data Quality Dimensions in Ontological Foundations," *Commun. ACM*, vol. 39, no. 11, pp. 86–95, Nov. 1996.
[16] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques: concepts and techniques*. Elsevier, 2011.
[17] G. Andrienko, D. Malerba, M. May, and M. Teisseire, "Mining spatio-temporal data," *J. Intell. Inf. Syst.*, vol. 27, no. 3, pp. 187–190, 2006.
[18] M. Das, "Spatio-Temporal Anomaly
Detection," 2009.
[19] R. Buchheit, J. Garrett, S. McNeil, and P. Chen, "Automated Procedure to Assess Civil Infrastructure Data Quality: Method and Validation," *J. Infrastruct. Syst.*, vol. 11, no. 3, pp. 180–189, Sep. 2005.
[20] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Pearson Education, Limited, 2014.
[21] Y. Kou, C.-T. Lu, and D. Chen, "Spatial Weighted Outlier Detection.," in *SDM*, 2006, pp. 614–618.
[22] J. I. Maletic and A. Marcus, "Data Cleansing: Beyond Integrity Analysis.," in *IQ*, 2000, pp. 200–209.
[23] W. K. Härdle and L. Simar, *Applied multivariate statistical analysis*. Springer Science & Business Media, 2012.
[24] M.-S. Chen, J. Han, and P. S. Yu, "Data mining: an overview from a database perspective," *Knowl. data Eng. IEEE Trans.*, vol. 8, no. 6, pp. 866–883, 1996.
[25] A. L. N. Fred and A. K. Jain, "Combining multiple clusterings using evidence accumulation," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 27, no. 6, pp. 835–850, 2005.
[26] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD Record*, 1993, vol. 22, no. 2, pp. 207–216.
[27] "NBI Data Checks - National Bridge Inventory - Bridge Inspection - Safety - Bridges & Structures - Federal Highway Administration." [Online]. Available: http://www.fhwa.dot.gov/bridge/nbi/checks/. [Accessed: 28-Jul-2015].
[28] S. Mcclure and M. Langehennig, "An Evaluation of the Condition of the Nation's Bridge Infrastructure Using Indicators Extracted from National Bridge Inventory Source Files," pp. 1–22, 2013.
[29] Delaware DOT, *NBI Database Quality Assurance Checks*. 2015.
[30] "NBI ASCII Files - National Bridge Inventory - Bridge Inspection - Safety - Bridges & Structures - Federal Highway Administration." [Online]. Available: http://www.fhwa.dot.gov/bridge/nbi/ascii.cfm. [Accessed: 28-Jul-2015].
[31] "National Transportation Atlas Databases 2014 | Bureau of Transportation Statistics." [Online]. Available: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2014/index.html. [Accessed: 28-Jul-2015].
[32] "Data Catalog - Data.gov." [Online]. Available: https://catalog.data.gov/dataset?q=NM&sort=none&collection_package_id=c440235f-8bf9-4001-ac37-a4f4397609cc&ext_location=&ext_bbox=&ext_prev_extent=-142.03125%2C8.754794702435605%2C-59.0625%2C61.77312286453148. [Accessed: 21-Jul-2015].
[33] "ArcGIS Help 10.1 - Erase (Analysis)." [Online]. Available: http://resources.arcgis.com/en/help/main/10.1/index.html#//00080000000m000000. [Accessed: 21-Jul-2015].
[34] H. Ludescher and R. Hajdin, "Distinctive features of the Swiss road structures management system."
[35] "Points | Bureau of Transportation Statistics." [Online]. Available: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2013/points.html. [Accessed: 21-Jul-2015].
[36] G. Scroggins, "Washington State Bridge Inspection Manual M 36-64.05," 2015.
[37] "U.S. Geological Survey Open-File Report 93-480." [Online]. Available: http://ma.water.usgs.gov/publications/ofr/scour.htm. [Accessed: 07-Jul-2015].
[38] B. Chavel, "Steel Bridge Design Handbook: Bridge Deck Design," Washington, D.C. 20590, 2012.