Automated Detection of Significant Deviations in a Spatial Position of Oil Pipelines

Alla Yu. Vladova

V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
avladova@ipu.ru

Abstract. Selective comparison of the oil pipeline sections based upon datasets of multiple in-line inspections [1] showed that there is a significant group of sections with 3d position changed again and again after repairs. At the same time, increasing volume of in-line inspections makes it impossible to analyze a spatial position of each pipeline section over time. It provokes adapting methods of multidimensional data analysis for automating detection of significant deviations in a spatial position of the pipeline. First phase of data preparation algorithm includes checking the uniqueness headers of dataset, lack of duplicates and gaps, lack of special characters, unprintable characters and extra spaces. The second phase includes checking misses, as well as significant and rapid changes in trends. Method of detecting significant deviations in a spatial position of the oil pipeline consists of four main steps: evaluating correlation coefficients of datasets, selecting the grouping method [2], analyzing intra-group statistics and assigning compensating activities for each group of pipeline sections.

Keywords: Multidimensional dataset · Pipeline sections · Compensating activities · Monitoring · Repair · R-programming · Inline inspections

1 Introduction

Operation of underground pipelines contributes to bending stresses in its walls. The situation is significantly aggravated at plots with changing geological conditions: freezable swamps, landslide slopes and permafrost. Therefore, in order to ensure trouble-free operation, changes in the pipeline spatial position shall be analyzed. A spatial position of every section of a pipeline is characterized with a bending radius and a turn angle and is set up at a design stage. Monitoring changes in an oil pipeline spatial position bases upon regular in-line inspections, strength calculations and comparative analysis.

In: S. Belim et al. (eds.): OPTA-SCL 2018, Omsk, Russia, published at http://ceur-ws.org

2 Data Source

In-line inspections provide information about high-altitude situation, bending radius and turn angle of every section of an oil pipeline. A fragment of comparative analysis of bending radius and turn angles over 3 years is represented in Table 1.

Table 1. Changes in angles and radii in 2013-2016 years.

Section	A2013, °	R2013, m	A2014,°	R2014, m	A2015,°	R2015, m	A2016,°	R2016, m
12570	0	393	2	382	1	388	0	386
33650	187	444	187	423	190	433	186	460
92050	343	547	337	527	347	568	340	518
92060	342	554	336	525	346	571	340	518
95200	349	519	344	566	350	497	348	527
96600	177	502	175	516	179	510	176	528
100920	350	462	349	482	350	495	349	466
102650	176	548	176	504	174	538	177	517
102660	176	548	176	509	174	540	176	525
102670	355	477	350	514	352	519	349	520
102840	354	538	2	494	0	534	357	509
104400	186	397	184	410	187	409	185	406

The in-line inspection database consists of more than 700 000 records for every survey [3]. Fig. 1 illustrates the ratio of stressed sections with non-normative bending radius, to whole amount of stressed sections at some sites of the oil pipeline.

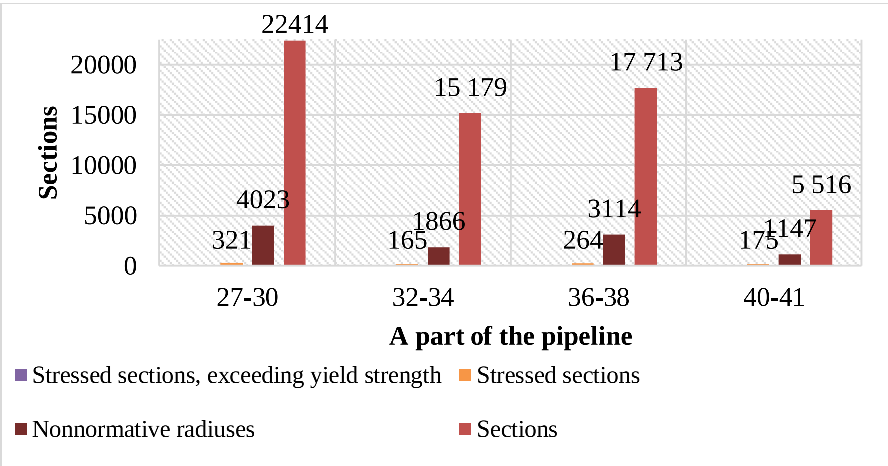
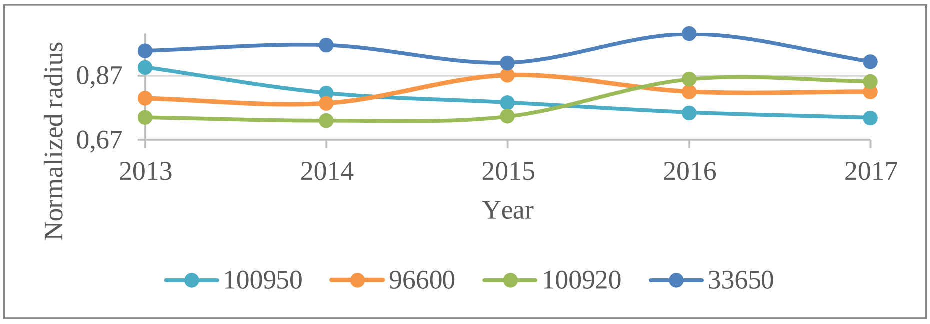
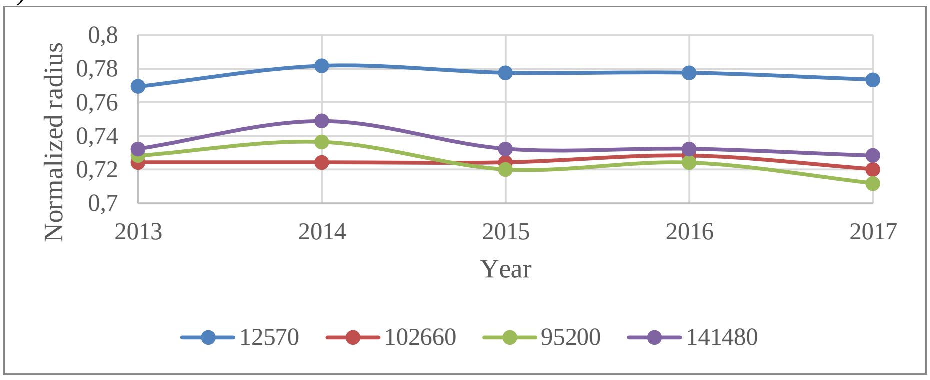


Fig. 1. Comparative analysis of sections within the pipeline sites.

Comparative analysis of bending radii based upon in-line inspections showed that there is a significant group of repaired sections with stable decreasing bend radii (see Fig. 2a, sections No 100950, 96600, reparation works were made in 2015 year and Fig. 2b, sections No 95200, 141480, reparation works were made in 2014 year). Apparently, it depends on the quality of the repairs and soil conditions.



a)



b)

Fig. 2. Changing overtime: a) bend radius; b) turn angles.

Analysis based on the in-line inspection data has shifted from the purpose of finding defects that had to be repaired to monitoring of the pipeline's condition. Thus, the purpose of this work is automated identification of pipeline sections with deteriorating spatial position, despite of compensating activities.

3 Cluster Analysis in Oil and Gas Industry

Previously the in-line inspection results was observed right after delivery and then archived. But today these archives are used in different types of analysis years after the actual inspections have taken place. Cluster analysis allows to categorize and to visualize large amount of data that are specific to the oil and gas industry. Paper [4] suggests diagnosing gas leaks with the sound produced by broken pipeline. Sound analysis is carried out using Fast Fourier transform with subsequent clustering on mind spectrum. Paper [5] uses fuzzy clustering algorithm to classify types of defects of underground pipeline bases upon the in-line inspections data. [6] offers a grouping algorithm of distributed data, analyzes data of independent monitoring systems. The paper [7] shows dimensionality reduction of a pipeline route thermal field-analyzing task based on clustering thermowells.

Patent [8] builds a model of geological environment at drilling process, clustering volumetric and qualitative parameters of the reservoir to optimize trajectory and characteristics of drilling. Patent [9] performs clustering rock formations at the site of well to define their differences, to identify heterogeneity, to offer visual indication of best collectors and to provide best potential for commercial exploitation of specific wells. Patent [10] proposes a method of evolutionary search with clustering of signs of limiting states of constructions of complex objects, their defects and damages leading to pre-emergency situations.

4 Clustering Spatial Position of Pipeline Sections

At the first stage we do focus on dataset formation (see Fig. 3).

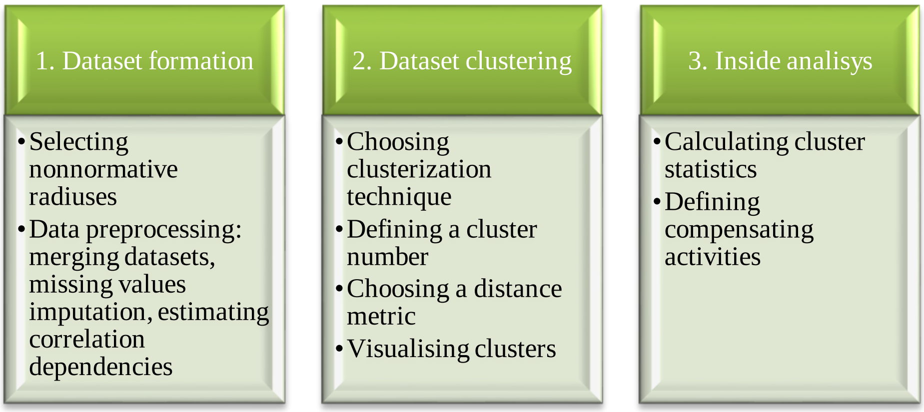


Fig. 3. Stages of clustering analysis of bend radiuses.

The data preprocessing algorithm checks unique headers; absence of duplicates and omissions; presence of special characters, unprintable characters, extra spaces. If missing values are scattered across the entire dataset, record deleting can destroy an appreciable fraction of the data. Therefore, at the first step for each thirty-kilometer site of a pipeline, we delete records if missed measurements exceed 20% [2]. At the second step, we impute missing values with row-means. The data preprocessing algorithm in terms of R language uses functions manyNAs, is.na и na.aggregate from libraries DMwR и zoo.

Distances between cluster objects are calculated according to the following formula:

$$P(x, x^*) = (\sum_{i=1}^{N} |x_i - x_i^*|^v)^{1/p}), \quad (1)$$

i - is a counter, $i = \overline{1, N}$;

N - is the number of TCs;

v and p are parameters of the distance metric. The selection of v and p is based on the following criteria:

- if necessary for lowering the impact of large individual differences, v = p = 1 (the Manhattan distance);
- if necessary, increase or decrease the weight of a dimension for which corresponding objects vary, v = p = 2 (the Euclidean distance) or v = 2, p = 1 (the squared Euclidean distance).

Clustering of a composite set of bending radii with preliminary determination of a number of clusters is realized in the language R using functions kmeans, aggregate and clusplot from the cluster() library [11].

5 Results

Raw datasets show a significant number of missed measurements (Table 2).

Table 2. Fragment of a dataset with multiple missing measurements.

Section	**2012**	**2013**	**2014**	**2015**	**2016**
142980	*NA*	1470	*NA*	1427	1380
144080	*NA*	*NA*	1826	2061	2295
144470	*NA*	1489	1580	1608	1521

Correlation analysis of time-separated measurements showed that the smallest correlation coefficient for datasets bounded by non-normative bending radii is 0.53, and for complete datasets is 0.14. It happened due to different types of in-line inspections equipment, a significant number of repairs, and deterioration of soil bearing capacity.

As a clustering result, we obtained two sets of pipeline sections for each site of the oil pipeline. Visualizing clusters (see Fig. 4) we used principal components and determined the abscissa and ordinate axis as dimensionless values of the first and second principal components [11].

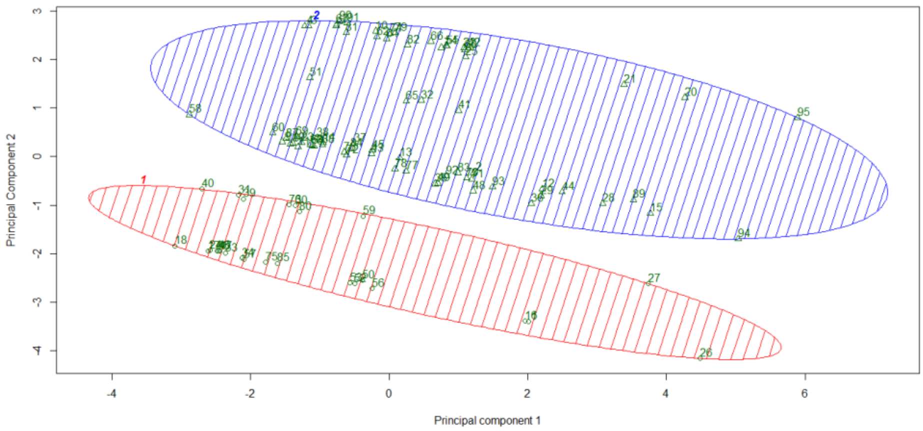
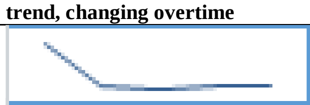
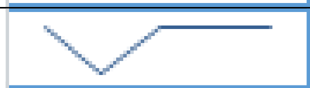
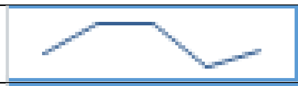
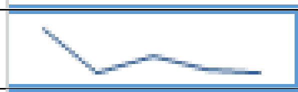


Fig. 4. Sections, distributed in two clusters.

Fragment analysis of appointment of compensatory actions to pipeline sites depending on the cluster is presented in table 3.

Table 3. Summary analysis of selected oil parts

Part	Cluster	Section	A sample of the bend radius trend, changing overtime	Compensatory actions
23-24	1	14130		Repair
	2	15610		Monitoring
27-29	1	119170		Repair
	2	121510		Monitoring
32-34	1	138400		Repair
	2	148580		Monitoring
36-38	1	157510		Repair
	2	173480		Monitoring

Cluster' statistics trends in bending radii over time show that the pipeline sections are predominantly distributed across clusters as follows: a negative trend and a neutral trend.

6 Conclusions

To process in-line inspection's data, we applied cluster analysis. It allowed grouping pipeline sections into two sets: requiring compensating activities and monitoring. It significantly simplified our analysis task and made it possible to identify in relationship between the laying conditions and the spatial position of the pipeline. Novelty of the proposed approach consists of:

- developed method of automated allocation a pipeline sites requiring compensatory activities;
- revealing the trend and detecting significant deviations in the values of controlled parameters, affecting strength, reliability and service life of a pipeline.

References

1. Vanaei, H. R., Eslami, A., Egbewande, A.: A review on pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models. International Journal of Pressure Vessels and Piping, 149, 43-54 (2017)
2. Kabakoff, R.: R in Action., Manning Publications (2011)
3. Surikov, V.I., Mogilner, L.Yu., Vladova, A.Yu., Tambovtsev, A.V., Provorov A.V.: Creation, introduction and support of archive for electronic copies and digitized data of trunk oil pipeline route. Science & Technologies: oil and oil products pipeline transportation 4(20), 52-60 (2015)
4. Shibata, A., Konishi, M., Abe, Y., Hasegawa, R., Watanabe, M., Kamijo, H.: Neuro based classification of gas leakage sounds in pipeline. In: Proceedings of the IEEE International Conference on Networking, Sensing and Control, 298-302 (2009)
5. Ziashahabi, M., Sadjedi, H., Khezripour, H.: Automatic segmentation and classification of pipeline images using mathematic morphology and fuzzy k-means algorithm. In: Machine Vision and Image Processing (MVIP), IEEE, pp. 1-5 (2010)
6. Naldi, M.C., Campello, R.J.G.B.: Evolutionary k-means for distributed datasets. In: Brazilian Symposium on Neural Networks. vol. 127, 30-42 (2014)
7. Vladova, A.Yu.: Algorithmic support of information system for geotechnical monitoring of hydrocarbon transportation in permafrost conditions. Information technologie. 23(3), 205-212 (2017)
8. Gzara, Kais B.M., Dzhain, V.: Determination of characteristics of bed components on site of works performance. RU 2574329 C1, 4 (2016)
9. Suares-Rivera, R., Khandverger, D.A., Soudergren, T.L.: Method and apparatus for multidimensional data analysis to identify rock heterogeneity. RU 2474846 C2, 4 (2013)
10. Bekarevich, A.A., Budadin, O.N., Morozova, T.Yu, Toporov, V.I. Method for adaptive forecasting of residual operating life of complex objects, and device for its implementation. RU 2533321 C1, 32 (2014)
11. Pison, G., Struyf, A., Rousseeuw, P.J.: Displaying a clustering with CLUSPLOT. Computational Statistics & Data Analysis 30(4), 381-392 (1999)