Zero Defect Manufacturing of Microsemiconductors – An Application of Machine Learning and Artificial Intelligence

Conference Paper · November 2018

DOI: 10.1109/ICSAI.2018.8599292

5 authors, including:

Zhengwen Huang
Brunel University London

Veerendra C Angadi
Brunel University London

A. Mousavi
Brunel University London

Zero Defect Manufacturing of Microsemiconductors – An Application of Machine Learning and Artificial Intelligence

Zhengwen Huang
College of Engineering, Design and Physical Sciences
Brunel University
London, UK
Zhengwen.Huang@brunel.ac.uk

Veerendra C Angadi
College of Engineering, Design and Physical Sciences
Brunel University
London, UK
Veer.Angadi@brunel.ac.uk

Morad Danishvar
College of Engineering, Design and Physical Sciences
Brunel University
London, UK
Morad.Danishvar@brunel.ac.uk

Ali Mousavi
College of Engineering, Design and Physical Sciences
Brunel University
London, UK
Ali.Mousavi@brunel.ac.uk

Maozhen Li
College of Engineering, Design and Physical Sciences
Brunel University
London, UK
Maozhen.Li@brunel.ac.uk

***Abstract*— A real-time quality monitoring of the detection and prediction of a defect in fluid dispensing systems is presented. A case study of adhesive placement and dispensing in a semiconductor production system demonstrates the applicability of a combination of PCA to explain the variations in the amount of dispensed fluid syringe needle placement and event-based learning to express the causal relationship between machine and production state with defect types. The resulting definitions of system state and interrelationship of control parameters build the building blocks of Gene Expression Program (GEP) that predicts the formation of droplets and fail or pass product. The results show 99.93 % of accuracy in prediction of defect which is based on the obtained data from glue dispensing model. This integrated solution provides the genetic signature of the glue dispensing process helping to eliminate defects and the adjustment of system state prior to defect formation.**

Keywords—Fluid dispensing systems, Data-driven methods, Event-Modeler technique, GEP technique.

I. INTRODUCTION

The time-pressure fluid dispensing systems which is widely used in electronic manufacturing require miniaturized pressure control techniques to release a specified amount of fluid (glue) on a targeted board. To increase quality and efficiency and reduce cost and time, automated fluid dispensing machine usage is growing rapidly. However, controlling such machines to acquire an acceptable quality of the dispensed fluid is the main technical challenge [1]. In these processes, the flow rate of the fluid dispensed, the fluid amount transferred onto the board and/or uniformity of fluid shape are three important performance indexes [2]. These performances may vary with gradual variation in pressure-volume behind the fluid, or temperature of fluid which affecting the dynamics of flow rate.

In the dispensing process also noticed that the fluid amount and shape dispensed on the board are highly correlated to the age of the dispenser needle and the machine deprecation (normally manifesting itself in variation in calibration and settings). It was observed in the experiments that when the dispensed fluid is very small (dimensions) all these impact factors are more important.

In this paper, a dispensing fluid flow rate model is delivered and presented based on an industrial experiment. Principle Component Analysis (PCA) [3] is used to detect defects in the volume of dispensed fluid and dispensed fluid needle release point [4]. Event-Modeler technique [5] as a middleware provides highly coupled input variables (defects sources) relationship with the type of defects and then facilitate the monitored system variables into Gene Expression Programming (GEP) [6] technique to build a prediction model. Based on the simulation and modeling of this effective and robust model of fluid dispensing performance, defects are detected, and their occurrence is predicted in real-time.

This paper is organized as follows: Section II introduces system definition and an industrial case investigated in this paper. Section III presents a PCA based solution for defect detections. Section IV illustrates the defect prediction mechanism. Section V concludes the defect detection and prediction work.

II. SYSTEM DEFINITION FOR AN INDUSTRIAL CASE STUDY

Silicon die and other wire bondable components are attached to a board by means of a conductive glue dispensed via the dispensing system. Once the glue dots have been placed on the board the bond head then picks up the silicon die from a waffle tray and places it into the glue to a controlled height. Two different types of malfunction happen during dispense of the conductive glue. The glue can either be insufficient, or there can be an excess of the dispensed glue. Both malfunctions cause damage to the part and result to the discard such parts. Failing to control the shape and volume of the glue dot dispensed results is part/component failure.

Figure 1 is a schematic of a conventional time-pressure dispensing machine. The representation of the rheological behavior of the glue dispensed will be discussed since it is crucial in a dispensing process and can significantly affect the performance of the production process [7]. The relationship between the amount of glue dispensed per second (Q), Glue level in the syringe (L_S), and applied pressure (P_g) can be presented by Equation (1), which can predict the amount of glue dispensed per second with the applied pressure and glue level in the glue syringe. Furthermore, Glue volume is also related to glue temperature in a syriange and its thermal expansion.

$$Q = -4.54 + 0.00032*(L_S) + 1.86x10^{-5} *(P_g) \quad (1)$$

The shape of the dispensed glue is monitored for detecting faults. A system of camera sensor with an aperture of radius, r_i and centred at $I(x, y)$ is installed to near the needle of glue dropping to capture the state of the system as shown in Fig. 2a. The aperture acts as a tolerance measurement which is assumed to be 80% of the field of view of the camera. The shape of the dispensed glue changes according to the system parameters such as glue temperature, the pressure applied to dispense the glue, and the depreciation of the machine and the needle.

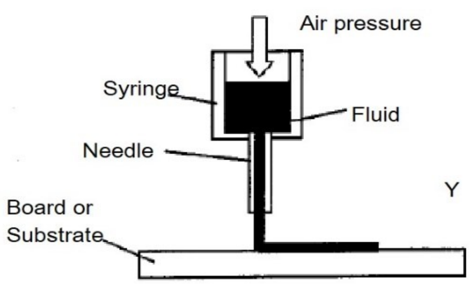


Fig. 1.Time-pressure dispensing schematic (ref. [1])

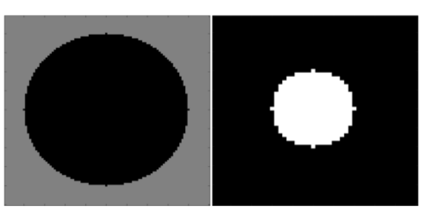


Fig. 2. (a) Sensor image of mask describing the tolerance of 80% of the image. (b) Simulated glue droplet center at (0,0) with a radius of 20% of the image.

The glue dispenses are modeled as a circle in the binary image shown in Fig. 2b with a radius, r_c and centred at $G_c(x, y)$. Depreciation over time is modelled as the change in the position of the centre of the glue drop needle from the centre of the image. As the temperature is increased, shape the of the glue increases based on the liquid thermal expansion. The thermal expansion of the glue is assumed to be linear with respect to the volume.

III. Defect Detection

A. Principle Component Analysis (PCA)

Multivariate analysis such as PCA is a robust analysis in detecting minute changes in the stream of data. The variance between the principal components (PC) and the direction of Eigenvector can be used to detect the number of end-members that are contributing to the data [8]. However, the PCA is highly sensitive to the changes in the data and any tolerance needed are difficult to implement directly.

Fault detection by PCA multivariate analysis on a system of simulated images is carried out by assuming one fault might occur in the series of 10 glue dropping. The faulty frame is simulated by selecting center and radius of the glue drops by the normal random process at a randomly selected location out of 10 frames. This stochastic analysis is useful in understanding the robustness of the process of detecting faults.

B. Measurement of error

The distance (d_c) of furthest part of the glue from the centre of the image, $I(x, y)$, is calculated as shown in equation (2). Hence if $d_c > r_i$ condition is satisfied, then an error has occurred. The error is measured by calculating the distance the glue has crossed the radius of the mask as shown in equation (3). The distance is the Euclidean distance between the centre of the image to the circumference of glue drop.

$$d_c = \sqrt{I(x, y)^2 + G_c(x, y)^2} + r_c \quad (2)$$

$$\text{err.}\% = \left(\frac{d_c}{r_i} - 1\right) \times 100 \quad (3)$$

C. Fault detection

The multivariate analysis detects if there are glue drops that have crossed the 80% aperture (The level of the glue on the board is called glue fillet and there is a threshold of the acceptability of the glue). Fig. 3 shows that a series of 10 images are used to check the faults in the glue drops. Clearly, the first 9 images are the camera sensor data with no glue crossing the aperture. However, in the 10th image, glue has crossed the aperture and the error is 0.40%. The 1st principal component would show the overall shape which in this case would be "No defect" images (from 1 – 9 in Fig. 3) and the "Defect" would be a 2nd principal component.

The robustness of the fault detection of glue drops is simulated at random location, with random radius and center to check if the multivariate method such as PCA can detect the fault and the location of the fault as shown in Fig. 4a and 4b. The variance of the PC indicates the presence of a number of faults and the biplot of 1st and 2nd component show the locations of the faults, i.e. the 2nd image in the batch of 10 drops. In Fig. 4c, the 2nd PC is insignificant, indicating there are no faults present and its biplot in Fig. 4d shows the only 1st image as the defect as a default.

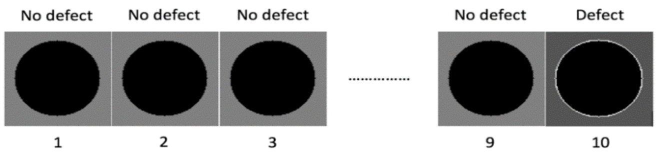
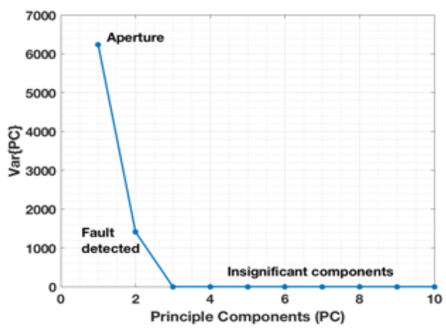
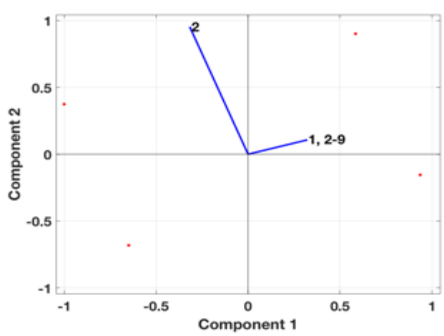


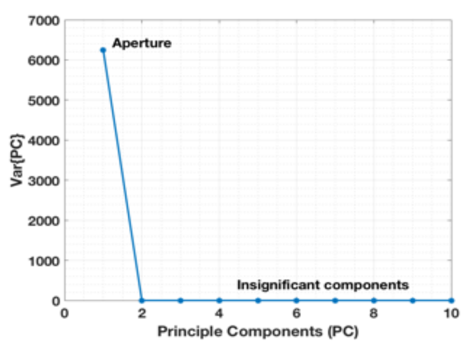
Fig. 3. A batch of glue dropping from the perspective of the camera sensor.



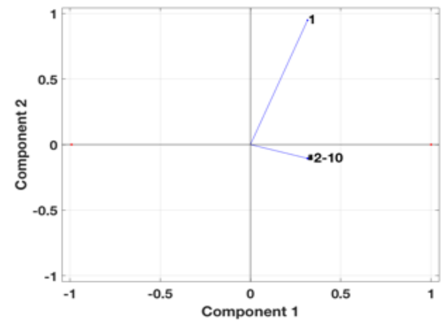
(a)



(b)



(c)



(d)

Fig. 4. PCs sorted according to their variance. (a) 2nd component is indicating a defect and (c) No defect in glue dropping. Biplot of PC #1 vs PC #2 showing (b) 2nd image as the defect and (d) No defect.

IV. Defect Prediction

The events and defects detected from the detection stage and the state of the system need to be engineered into high value and more accurate models that would be able to predict the occurrence of the defect. For each type of defect (insufficient or excess of the dispensed glue) the correlated system parameters and state are identified and tagged to defect, no defect and high-quality events. Event-Modeler information structure is implemented in the data management system, where the data is fed in real-time into the Event Modeler. The Event-modeler output produces more specified results of which group of system parameters such as pressure or temperature is highly coupled with the volume of glue or center of the needle of dispensed glue during the process.

Consequently, Event-Modeler builds a map of the correlation between system input and output parameters and will be used to assisting in building the necessary knowledge for the predictive model.

A. Event-Modeler

The EventTracker [9] and the EventiC [5] (jointly called Event-Modeler) are methods for real-time sensitivity analysis in large-scale complex systems. The real-time ability to group and rank relevant input-output event data in order of its importance and relevancy. The correlation analysis allows the key parameters that explain the system state to be coupled together, thus facilitating the modeling of the system through upper layer modeling. Consequently, the Even-Modeling algorithm will provide a new insight into the correlations between the key parameters of the system. Such input variable selection mechanism will allow for the development of new and improved performance models (e.g. transfer functions, regression and so on).

B. Eventi algorithm implementation

This research is based on a long-term glue dot radius (r_i) and center of glue drop point monitoring campaign performed with the historical data (3000 drops) samples over 8 input variables. There are some other inputs variables such as glue properties, syringe size or threshold of glue fillet which are fixed and unchanged during the experiment and have not been considered in the input variable parameters list. The results are presented in Table I.

The outcome of the EventiC provides four high highly coupled input variables for glue radius control and two high highly input variables for dispensed needle position control. The results, thus facilitating the modeling of the system through an evolutionary algorithm called GEP in the next stage.

TABLE I
Averaged SA weight of glue radius and Needle position over system's eight input variables

Input variable parameters	Sensitivity Level of glue radius, r_i	The sensitivity level of Needle center of glue drop
Glue temperature	**High**	Low
Glue level in the syringe	**High**	Low
Pressure behind the Glue	**High**	Low
The threshold of glue fillet	**High**	Low
Air pressure	Low	Low
Room temperature	Medium	Medium
Machine calibration decay	Medium	**High**
Needle age	Medium	**High**

C. GEP

Gene Expression Programming (GEP) [6] is a variant of the evolutionary algorithm. It has a special phonotype-genotype separated solution presentation structure. With such structure, the search of target space is well organized by locating selection pressure into the corresponding position efficiently. The well-organized search mechanism grants GEP a stable data analyses functionality which ensures a solid performance on correlation mining problem [10]. GEP is particularly good at mining the interactive correlation among multiple factors working in a target system. The generated correlations are represented with mathematical functions which take some causal factors as input and return result factor as outputs. In other words, the change on system components (in terms of value changes on related system components) can be linked to a change of system status with such functions. The mathematical function provides a projection image of the target system. If the accuracy level of such a mathematical function can be well developed after the evolution process, a prediction of system status change is achievable. The next status of the target system (output of such function) can be calculated with given information of major system components (settings of input component of such function). The general evolution process of GEP algorithm is provided in Fig. 5.

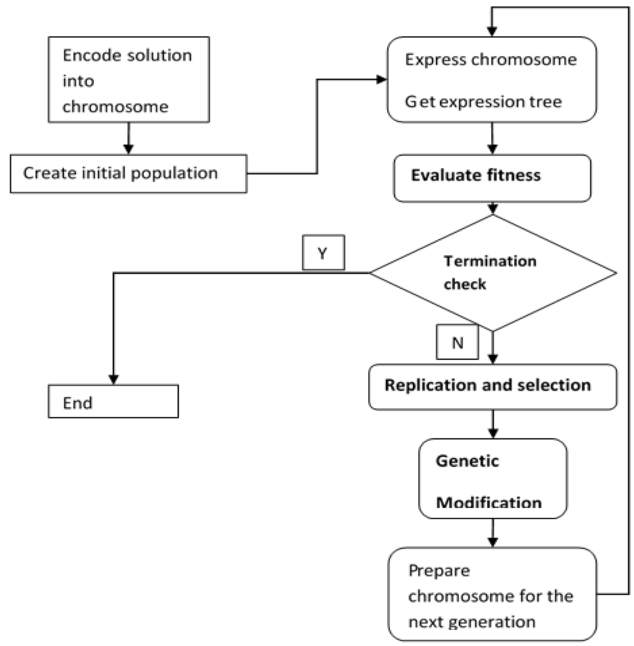


Fig. 5. The evolution process of GEP.

A correlation mining example of GEP is provided in Fig. 6. This example demonstrates the correlation mining process of two parameters (x_0 and x_1). Based on the data types of parameter setting samples, a function candidate is selected from mathematical functions to fit the requirements of x_0 and x_1. The requirements of x_0 and x_1 are target driven object functions. With different targets the requirement can be defined with different concerns. In this example the requirements include a) valid – the selected can be used to describe the correlation of x_0 and x_1; b) accurate – the selected function provides an accurate coverage on the behaviour of x_0 and x_1. At the end of evolution, the process, the best function is eventually selected. In this example, the Plus function is selected.

In this paper, GEP is employed to investigate the internal correlation among components of the glue dispensing system. The searching space of glue dispensing problem is coordinated with the components of such a system. To search the correlations

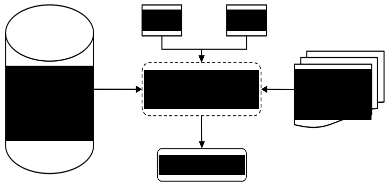


Fig. 6. The correlation mining example of GEP.

of system components, GEP`s element stream chromosome is constructed with the components of the glue dispensing system and potential functions which can describe the behavior of components. The best correlation represented with the best chromosome generated from the end of evolution is further extracted to predict the occurrence of defects.

D. GEP implementation

In this paper, a GEP prediction solution was implemented with data generated with the same dataset as events. The prediction is achieved with the following steps.

1) *GEP is employed to investigate the correlation of involved system components (listed in TABLE II) of the glue dispensing system by learning from a historical dataset.*

In step 1, following our previous method used in work [11] and [12], we select some potential system components which may contribute to the defect as our input factors of the prediction function. By learning from a historical dataset (containing profiles of 3000 drops), the prediction function is refined during the evolution process. The accuracy of the prediction function is also updated generation by generation during the evolution.

TABLE II. Considered system components (Variables)

Id	V0	V1	V2	V3
Component Name	Initial pressure	Pressure behind the syringe	Glue expansion Coefficient	Glue Temperature
Id	V4	V5	V6	V7
Component Name	Glue weight (µg)	Needle age (Drop)	x zero points (µm)	y zero points (µm)
Id	V8	V9	V10	V11
Component Name	X Array (µm)	Y Array (µm)	Z Array (µm)	Glue Volume (*10E-3mm^3)
Id	V12	V13	V14	V15
Component Name	x & y tolerance (mm)	Passed or Failed	Setting error	Glue level in the syringe
Id	V16	V17	V18	V19
Component Name	Drop rate per min	Total Production	Total GHG emission	Unit consumption (kW/part)
Id	V20	V21	V22	V23
Component Name	Total Energy consumption (kw)	Idle Machine consumption (kwh)	Busy Machine consumption (kwh)	Machine Utilization (%)

In this step, a classical GEP evolution process is implemented. The settings of GEP implementation are listed in TABLE III. The parameters were set using the classical values used for traditional GEP applications.

2) *A prediction function which takes listed components as inputs and the indication of defects as output is then generated from the best chromosome containing the best correlation learned from the glue dispensing system.*

TABLE III: GEP parameter settings.

Parameters	Values	
Population size	1000	
No. of genes in a chromosome	1	
Genetic modifications of GEP	one-point recombination rate	30%
	two-point recombination rate	30%
	insertion sequence transposition rate	10%
	inversion rate	10%
	mutation rate	0.44%

In step 2) the final (near best) version of prediction function is extracted from element stream like a chromosome. It is an applicable mathematical function. With the setting of system components, it returns a value which can be used to decide if a defect will be generated.

3) *With prediction function and the new settings of system components, the occurrence of a defect in next drop can be predicted.*

In step 3 the prediction of the future defect is achieved. Given the value of each involved system component, the prediction function is further connected with a glue dispensing system, the success/failure of every new drop can be predicted.

After the above 3 steps, we generate a prediction mechanism for the glue dispensing system. In order to further verify the performance of the proposed prediction mechanism, a comparison experiment was conducted. The prediction mechanism was applied with the historical data (3000 drops). The prediction function of such mechanism works like a classifier which returns an indication value. If the value generated by such function is less than a threshold, the corresponding case is classified as a defect (the value of the indicator is set to 0). Based on input values obtained from historical data, 3000 estimated indicators of defects were generated with the prediction function. The distribution of estimated indicators was collected and compared with the actual distribution of defect indicator recorded in historical data. The result is shown in Fig. 7.

As shown in Fig. 7, the results generated with the prediction mechanism provided by GEP and the actual distribution are identical. Those considered system components are contributing to a defect. Their contribution also can be represented with prediction function. It worth noting that from the prediction mechanism side and the predicted results we can observe similar contribution distribution that is provided by the input variables mentioned in TABLE I.

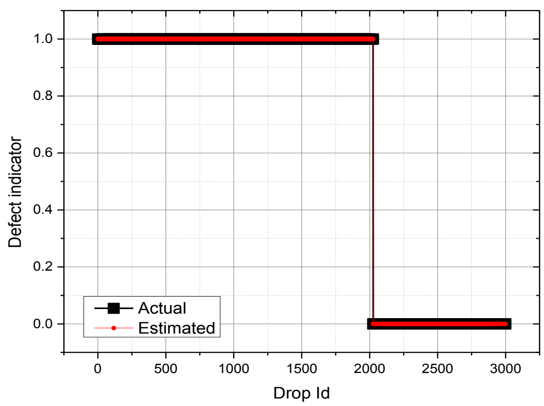


Fig. 7. Distribution of Defects

V. Conclusion

This research presents a solution strategy to minimize generation of a defect in a quality control problem on the basis of near real-time data acquisition, analytics, and decision. Based on historical data (3000 drops) of an industrial experiment, PCA multivariate technique is used to detect defects in the amount of dispensed glue and dispensed fluid needlepoint. Event-Modeler technique distinguishes the defects sources and their relationship with the type of defects and engineered the system variables into GEP technique. At the top layer, based on the glue dispensing model GEP builds a prediction model of fluid dispensing systems which can achieve an accuracy rate at 99.93% on the dataset obtained from our glue dispensing model.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 723906.

References

[1] Ting, Y., Chen, C., Feng, H. and Chen, S.," Glue dispenser route inspection by using computer vision and neural network". The International Journal of Advanced Manufacturing Technology, 39(9-10), 2008, pp.905-918.

[2] Chen, X.B., Schoenau, G., and Zhang, W.J. "On the Flow Rate Dynamics in time pressure Dispensing Processes", ASME Journal of Dynamic Systems, Measurement, and Control (in press), 2002.

[3] Villegas, T., Fuente, M. J., & Rodríguez, M. "Principal component analysis for fault detection and diagnosis. experience with a pilot plant". Proceedings of the 9th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics, 147–152, 2010.

[4] Ming-Da Ma, Wong, D. S.-H., Shi-Shang Jang, & Sheng-Tsaing Tseng. (2010). Fault Detection Based on Statistical Multivariate Analysis and Microarray Visualization. IEEE Transactions on Industrial Informatics, 6(1), 18–24.

[5] Danishvar, M., Mousavi, A., Broomhead, P. "EventiC: A Real-Time Unbiased Event-Based Learning Technique for Complex Systems ", IEEE Transaction on Systems, Man, and Cybernetics: Systems, 99, 1-14, 2018.

[6] C. Ferreira, "Combinatorial Optimization by Gene Expression Programming: Inversion Revisited," in Proc. Argentine Symp. Artif. Intell., pp. 160–174, 2002.

[7] Chen, X., Shoenau, G. and Zhang, W. "Modeling of time-pressure fluid dispensing processes". IEEE Transactions on Electronics Packaging, 2002.

[8] Jin, N., & Zhou, S., "Data-driven variation source identification for the manufacturing process using the eigenspace comparison method". Naval Research Logistics, 53(5), 383–396. https://doi.org/10.1002/nav.20150, 2006

[9] S. Tavakoli, A. Mousavi, P.Broomhead, "Event Tracking for Real-Time Unaware Sensitivity Analysis (EventTracker)," IEEE Trans. Knowledge and Data Engineering, vol.25, no.2, 2013, pp.348-359.

[10] Z. Huang, "Schema Theory for Gene Expression Programming, "Ph.D. Thesis, Brunel University London, 2014.

[11] L. Teodorescu, and Z. Huang, "Enhanced Gene Expression Programming for Signal-Background Discrimination in Particle Physics," in Proc. of the XII Advanced Computing and Analysis Techniques in Physics Research, 2008.

[12] Z. Huang, M. Li, C. Chousidis, A. Mousavi, & C. Jiang, "Schema Theory-Based Data Engineering in Gene Expression Programming for Big Data Analytics," IEEE Transactions on Evolutionary Computation. DOI:10.1109/TEVC.2017.2771445, 2017.