Performance Measurement and Management Model of Data Generation and Writing Time in Personal Computer

Dipankar Das, Parichay Chakraborti

Abstract—This paper is aimed at doing a performance measurement of data generation and write time in personal computer by generating relatively large quantity of numeric data and writing them in the text file and along with that developing performance models for them. These models may be used for performance management and at the same time they also provide an insight into the important aspect of relationship between 'time' and 'data size' of a particular platform. In this paper, we have used the curve estimation technique for analyzing the dataset and proposed two different types of generic models which can be viewed as: (i) 'Time' ~ f('Number of Data Elements') and (ii) 'Number of Data Elements' ~ f('Time'). In total six (6) models has been proposed, three from each type. In addition to providing performance measurement models of a particular platform (hardware and software), the proposed generic models also gives us the following information: (i) how much time is required to generate a given amount of numeric data, (ii) how much time is required to generate and write a given amount of numeric data in a text file, (iii) how much time is required to write a given amount of numeric data in a text file, (iv) how much amount of numeric data can be generated in a given time, (v) how much amount of numeric data can be generated and written in a given time and (vi) how much amount of numeric data can be written in a given time.

Index Terms— Performance Measurement, Performance Management, Curve Estimation Technique, Residual Analysis, Data Generation Time, Data Write Time, Data Generation and Write Time.
.

---

1 INTRODUCTION

IN today's business and finance world a very important lesson is "performance" and this concept is of utmost importance for the success of any organizational activity. Quality of any organization is judged by its performance and performance measurement is the building block of Total Quality Management (TQM) [1]. The process of collecting, analyzing and/or reporting information regarding the performance of an individual, group, organization, system or component is known as performance measurement [2]. To improve the quality and for having transparency, performance measurement is done by all the organizations to some extent. Measuring financial performance is relatively easier and most of the times we face steep challenges when try to measure non-financial performances in terms of what to measure, how to measure and/or what to do with the results [3]. The third question is more related to performance management. By performance management we mean the use of performance measures and standards to achieve desired results [4].

It's a well known fact that today we all live in the information society. In this society, all of us – may be an individual or an organization or a business or a process try to gain competitive advantage by using information technology [5]. We clearly understood that wealth can be created by proper cultivation and timely use of information. By information we commonly mean when data is processed, organized, structured or presented in a way to make it useful [6]. Today a lot of data is being generated. Over last two year 90% of all the data in the world has been generated [7]. Nowadays data is generated from everything and from every possible direction. It can be individuals or organizations or processes or any other sources. It is also true that today in most of the cases we store the generated data for future use. For data generation and storing purpose, in case of individuals and/or small enterprises specially in third world countries, mostly personal computers are used rather than servers and/or exploiting the cloud.

In this study, we have tried to measure the performance of data generation and write time in a personal computer and at the same time develop models for management of these functions (data generation and write) in a subtle way. Here, we have only concentrated on numeric data especially integer data generation and writing them in text files. As performance matrix (measuring term) the researchers have chosen data generation and write time, data generation time and data write time. The time taken to write the complete set of data in a text file has been considered as data write time. In this paper, after measuring the above mentioned performances, the authors have introduced two different types of generic models which can be viewed as: (i) 'Time' ~ f('Number of Data Elements') and (ii) 'Number of Data Elements' ~ f('Time'). These models such developed present a simple way for performance measurement and management of data generation and write time in a personal computer which is shown in the following figure (Fig. 1):

• *Dipankar Das is currently working as a Faculty member in The Heritage Academy, Kolkata, India, PH-919831592063. E-mail: dipankr.das@gmail.com*
• *Parichay Chakraborti is currently working as a Faculty member in The Heritage Academy, Kolkata, India, PH-919903897532. E-mail: parichay.c@gmail.com*

have

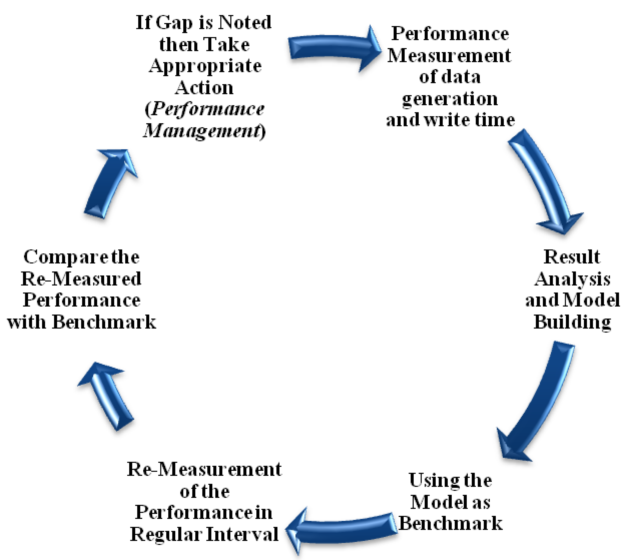

Fig. 1. Performance Measurement and Management of Personal Computer

In the present study we have mainly focused on measuring the performance of data generation and write time in a personal computer and analyzing the result to build models which may be used as benchmarks in future.

2 Related Work

In the following sub-sections (2.1 and 2.2) we have done a brief review on both the performance measurement of information systems and the performance measurement tools.

2.1 Performance Measurement of Information Systems

For evaluating the impact of information system on business performance one approach is contingent evaluation approach (Heo & Han, 2003) [8]. Comprehensive framework to assess the information system linked with the performance of the organization (Prybutok, Kappelman & Myers, 1997) [9] is also there. Measurement scales for four dimensions of the success has also been developed and tested (Saarinen, 1996) [10]. The basic DeLone and McLean information systems success model consists of six categories of information system success (DeLone & McLean, 2003) [11]. End user computing satisfaction (EUCS) measurement instrument (Etezadi-Amoli & Farhoomand, 1996) [12] has been developed long back in 1996. Performance measurement life cycle (PMLC) running parallel with system development life cycle (SDLC) has also been suggested for proper use and implementation of the measurement tools and changes suggested by the tools respectively (Patel & Maheshwari, 2000) [13]. For the strategic growth of SMEs, the information system performance measurement framework work as the foundation (Sharma & Bhagwat, 2006) [14]. The criteria model for information system has five divisions: organization, individual, information, technology and systemics (Palmius, 2007) [15].

2.2 Performance Measurement Tools

In the following we have listed some performance measurement tools –

a. Performance Measurement Tools from Intel®:

Intel® Performance Counter Monitor provides tools for estimating the internal resource utilization of the latest processors (Xeon® and CoreTM) [16]. Measure Intel® IPP Function Performance provides library functions for video, audio, image processing, signal processing, data compression and many more [17]. A cross platform performance analysis tool set is provided by Intel® Performance Tuning Utility which is having powerful data collection, analysis and visualization capabilities [18]. Intel® Graphics Performance Analyzers is a graphics analysis and optimization tool for helping game developers [19]. Intel® VTune™ Amplifier XE is a performance profiler for the programming languages – c, c++, java, Fortran, C# and Assembly [20].

b. Performance Measurement Tools of Ubuntu:

Some of the tools for monitoring the performance of Linux are: top [21], vmstat [22], lsof [23], tcpdump [24], netstat [25], htop [26], iotop [27], iostat [28], mpstat [29] and nmon [30].

3 Objectives of The Study

The main objectives of this study are two folds – (i) performance measurement of data generation and write time in a personal computer and (ii) analyzing the measured data to build models.

As discussed in section 1, the selected performance matrix is given in TABLE1.

TABLE 1
Computer Performance Matrix

Number of Data Elements (DE)	Average Data Generation and Write Time (ADGWT)	Average Data Generation Time (ADGT)	Average Data Write Time (ADWT)

The models to be developed in this study is listed below –

1) To identify the model that can be fitted to the data points 'ADGT' versus 'DE'.
2) To identify the model that can be fitted to the data points 'ADGWT' versus 'DE'.
3) To identify the model that can be fitted to the data points 'ADWT' versus 'DE'.
4) To identify the model that can be fitted to the data points 'DE' versus 'ADGT'.
5) To identify the model that can be fitted to the data points 'DE' versus 'ADGWT'.
6) To identify the model that can be fitted to the data points 'DE' versus 'ADWT'.

4 Research Methodology

4.1 Data Generation for Computer Performance Measurement

We have used Linux shell scripts to generate and write nineteen (19) numbers of relatively large number series on a particular machine. We have noted the 'Data Generation Time' and the 'Data Generation and Write Time' for each theses nineteen number series. We have collected one hundred (100) samples for each of these nineteen number series and calculated the 'Average Data Generation Time' (ADGT) and 'Average Data Generation and Write Time' (ADGWT) for each of these cases to avoid any inconsistencies. We have also computed the 'Average Data Write Time' (ADWT) for each these cases (19 numbers of number series) as given below:

'ADWT' = 'ADGWT' – 'ADGT' ------ (1)

The sample data set is given in TABLE2.

TABLE 2
SAMPLE DATA POINTS

Number of Data Elements (DE)	Average Data Generation and Write Time (ADGWT)	Average Data Generation Time (ADGT)	Average Data Write Time (ADWT)
100000	3.988	1.08	2.908
150000	6.049	1.621	4.428
200000	8.215	2.159	6.056
250000	10.285	2.688	7.597
300000	12.374	3.241	9.133
350000	14.403	3.79	10.613
400000	16.516	4.331	12.185
450000	18.604	4.882	13.722
500000	20.704	5.374	15.33
550000	22.648	5.938	16.71
600000	24.745	6.471	18.274
650000	26.892	7.042	19.85
700000	29.007	7.563	21.444
750000	30.983	8.075	22.908
800000	33.031	8.614	24.417
850000	35.047	9.208	25.839
900000	37.301	9.767	27.534
950000	39.33	10.299	29.031
1000000	41.558	11.063	30.495

Units of ADGWT, ADGT and ADWT all are in seconds.

4.2 Model Fitting

a. Expressing the Model in Terms of 'Time' ~ f('Number of Data Elements')

In the first case, we have taken 'Number of Data Elements' as independent variable and 'Average Data Generation Time' as dependent variable and the proposed generic model can be viewed as:

ADGT ~ f(DE) ------ (MODEL1)

In the second case, we have considered 'Number of Data Elements' as independent variable and 'Average Data Generation and Write Time' as dependent variable and the proposed generic model can be viewed as:

ADGWT ~ f(DE) ------ (MODEL2)

In the third case, 'Number of Data Elements' has been considered as independent variable and 'Average Data Write Time' has been considered as dependent variable and the proposed generic model can be viewed as:

ADWT ~ f(DE) ------ (MODEL3)

b. Expressing the Model in Terms of 'Number of Data Elements' ~ f('Time')

In the first case, we have taken 'Average Data Generation Time' as independent variable and 'Number of Data Elements' as dependent variable and the proposed generic model can be viewed as:

DE ~ f(ADGT) ------ (MODEL4)

In the second case, we have considered 'Average Data Generation and Write Time' as independent variable and 'Number of Data Elements' as dependent variable and the proposed generic model can be viewed as:

DE ~ f(ADGWT) ------ (MODEL5)

In the third case, we have taken 'Average Data Write Time' as independent variable and 'Number of Data Elements' as dependent variable and the proposed generic model can be viewed as:

DE ~ f(ADWT) ------ (MODEL6)

c. Goodness of Fit Statistics and Model Diagnostics

We have used curve estimation technique [31] for analyzing the dataset. At first we have explored linear model and if this model is not found suitable then quadratic model is tried and if it is also not found suitable then cubic model has been tested to find an appropriate model. For the purpose of this study we have considered R^2, Adjusted R^2, F – test and Significance of F – test as goodness of fit statistics [32]. For model diagnostics, different residual analysis methods have been used since residual analysis plays an important role in statistical modeling [33][34]. Here, we have used Residual vs predictor plot [35][39], Histogram of the residuals [36][39], Q-Q plot of the residuals [37][39] and Shapiro – Wilk (SW) test [38] for analyzing the residuals. We have performed the entire analysis at 95% confidence interval.

4.3 Hardware Platform

The characteristics of the computer are listed in the following table (TABLE3).

ISSN 2229-5518

TABLE 3
HARDWARE PLATFORM

Processor	Intel(R) Core(TM)2 Duo CPU T5870 @ 2.00GHz
RAM	3.00 GB

4.4 Software Used

Experimental data was generated using Ubuntu 11.04 (Kernel Version 2.6.38). We have used SPSS 17.0 and MS Excel for data analysis.

5 Data Analysis & Findings

5.1 Evaluating Model Fit

We have tried curve estimation technique for all the six (6) cases (MODEL1 – MODEL6) and the model summary is given in TABLE4.

TABLE 4
REGRESSION MODEL SUMMARY

Model Name	Model Type	R	R Square	Adjusted R Square	F	Sig. of F
MODEL1	Cubic	0.999939	0.999878	0.999853	40813.962720	0.000000
MODEL2	Linear	0.999991	0.999982	0.999978	272907.280905	0.000000
MODEL3	Linear	0.999984	0.999969	0.999962	159666.926131	0.000000
MODEL4	Cubic	0.999951	0.999901	0.999882	50672.702617	0.000000
MODEL5	Linear	0.999985	0.999970	0.999968	568285.591858	0.000000
MODEL6	Linear	0.999980	0.999960	0.999957	422360.508030	0.000000

From TABLE 4 we have observed that in all the cases the R^2 value is very high (close to 1) indicating that in all the cases the model fits the data well. The significance of F-test in all the cases is very low (less than 0.05) which indicate that in all the cases the observed R^2 value is reliable and the proposed relationship between the dependent variable and the independent variable is statistically reliable.

5.2 Model Diagnostics

a. Residual vs predictor plot of the models

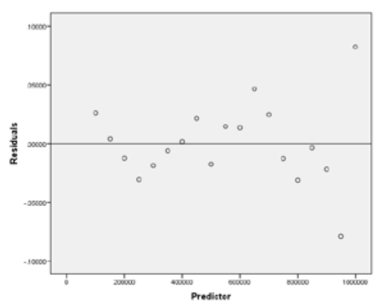

Fig. 2. Residual vs Predictor plot of MODEL1 (cubic)

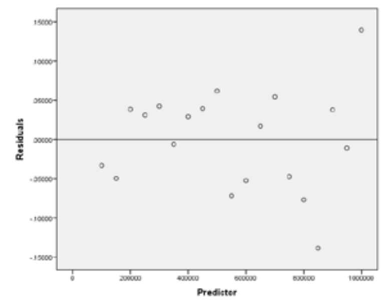

Fig. 3. Residual vs Predictor plot of MODEL2 (linear)

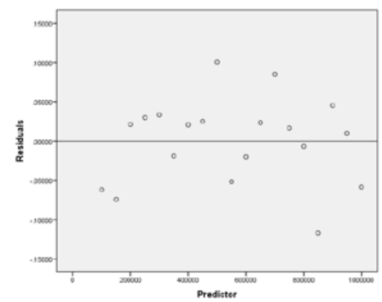

Fig. 4. Residual vs Predictor plot of MODEL3 (linear)

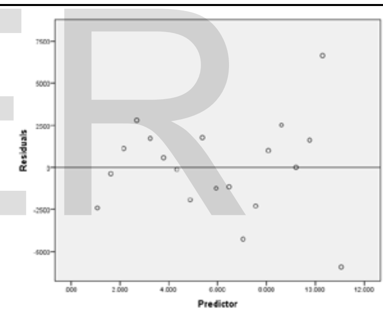

Fig. 5. Residual vs Predictor plot of MODEL4 (cubic)

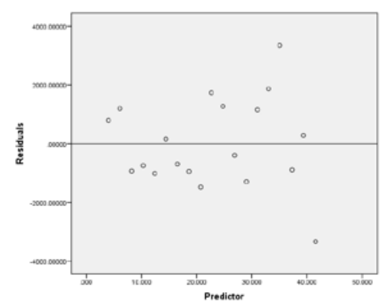

Fig. 6. Residual vs Predictor plot of MODEL5 (linear)

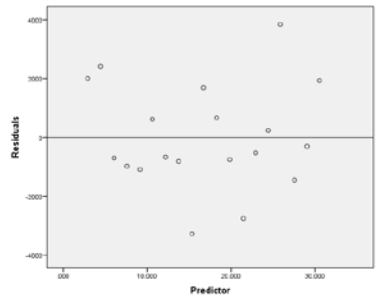

Fig. 7. Residual vs Predictor plot of MODEL6 (linear)

From the above figures (Fig. 2 - 7) it is evident that in all the cases the residuals appear to behave approximately randomly. Therefore, it suggests that in all the cases the model fits the data well.

b. ***Shapiro – Wilk (SW) statistics of the residuals of the models***

The SW statistics of the models discussed above is given in TABLE5.

TABLE 5
SHAPIRO – WILK STATISTICS OF THE MODELS

Model Name	Shapiro – Wilk Statistics	Sig. of Shapiro - Wilk
MODEL1	0.959339	0.559427
MODEL2	0.962542	0.623366
MODEL3	0.963378	0.640456
MODEL4	0.973498	0.843582
MODEL5	0.962837	0.629370
MODEL6	0.963588	0.644760

The significance of the SW Statistic of all the models are higher than 0.05. This suggests that the assumption of normality of the error distribution has been met for all the models.

c. ***Histogram of the residuals***

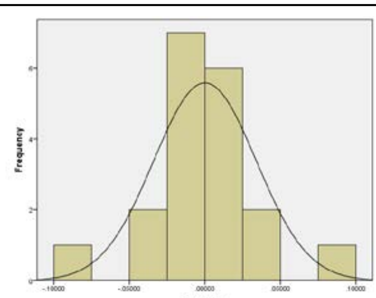

Fig. 8. Histogram of the residuals of MODEL1 (cubic)

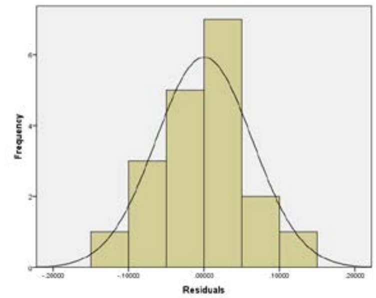

Fig. 9. Histogram of the residuals of MODEL2 (linear)

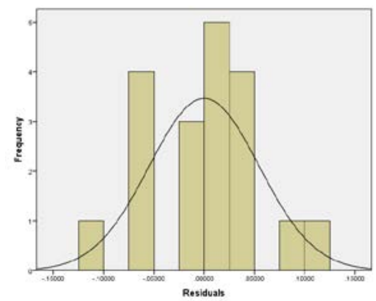

Fig. 10. Histogram of the residuals of MODEL3 (linear)

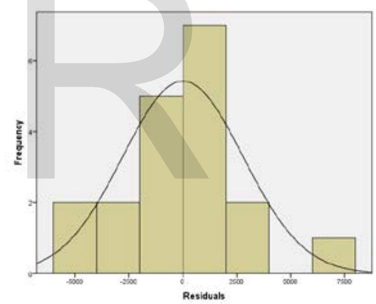

Fig. 11. Histogram of the residuals of MODEL4 (cubic)

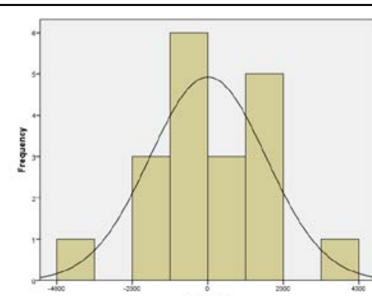

Fig. 12. Histogram of the residuals of MODEL5 (linear)

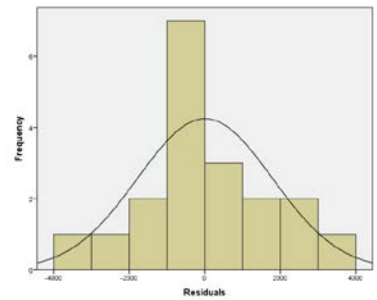

Fig. 13. Histogram of the residuals of MODEL6 (linear)

All the above figures (Fig.8 – 13) exhibit a symmetric bell shaped histogram which is evenly distributed around zero. Therefore it suggests that in all the cases the residuals are normally distributed.

d. *Q-Q plot of the residuals*

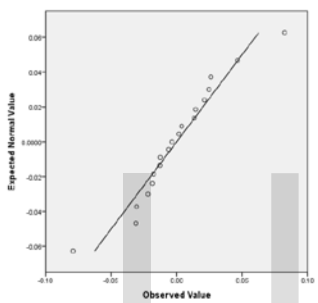

Fig. 14. Normal Q-Q plot of the residuals of MODEL1 (cubic)

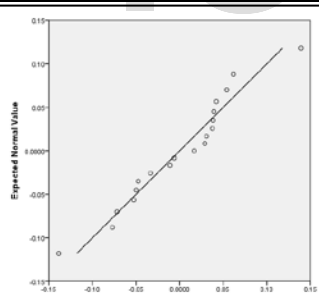

Fig. 15. Normal Q-Q plot of the residuals of MODEL2 (linear)

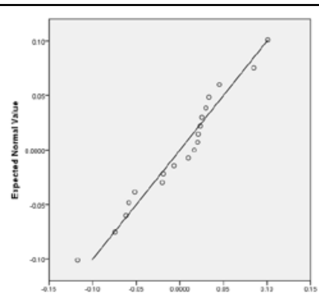

Fig. 16. Normal Q-Q plot of the residuals of MODEL3 (linear)

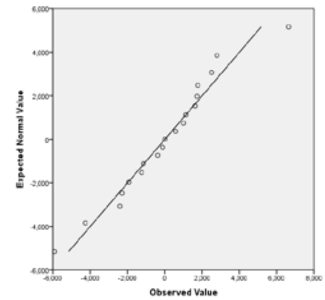

Fig. 17. Normal Q-Q plot of the residuals of MODEL4 (cubic)

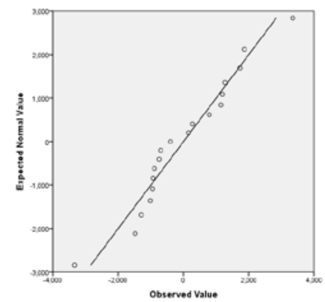

Fig. 18. Normal Q-Q plot of the residuals of MODEL5 (linear)

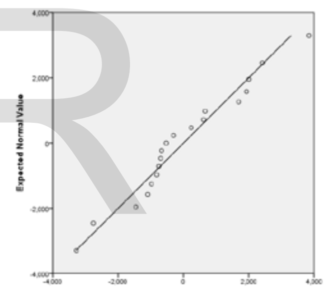

Fig. 19. Normal Q-Q plot of the residuals of MODEL6 (linear)

It may be easily observed from all the above figures (Fig. 14 – 19) that in all the cases the residuals are approximately linear. Therefore it suggests that in all the cases the residuals follow approximately normal distribution.

5.3 Proposed Mathematical Model

From the above sub-sections (5.1 and 5.2) we can conclude that the following models have been selected as proposed models:

ISSN 2229-5518

TABLE 6
PROPOSED MODELS

Model Name	Model Type	Model Equation
MODEL1	Cubic	ADGT = 1.182E-005*DE – 2.588E-012*DE**2 + 1.849E-018*DE**3 – 0.014
MODEL2	Linear	ADGWT = 4.155E-005*DE – 0.134
MODEL3	Linear	ADWT = 3.065E-005*DE – 0.095
MODEL4	Cubic	DE = 84038.096*ADGT + 1984.397*ADGT**2 -130.103*ADGT**3 + 9499.439
MODEL5	Linear	DE = 24065.206*ADGWT + 3238.53
MODEL6	Linear	DE = 32626.72*ADWT + 3117.783

ADGT: Average Data Generation Time,
ADGWT: Average Data Generation and Write Time,
ADWT: Average Data Write Time,
DE: Number of Data Elements

The graphical representations of the proposed models are given below (Fig. 20 – 25). The circles indicate the data points and the solid line represents the straight line (i.e. Linear model) in case of MODEL2, MODEL3, MODEL5 and MODEL6 models and cubic curve (i.e. Cubic model) in case of MODEL1 and MODEL4 models.

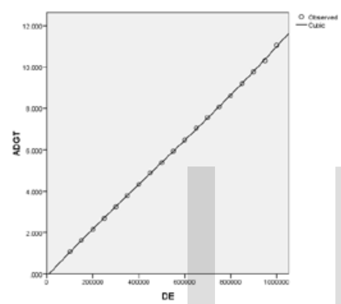

Fig. 20. Graphical representation of MODEL1 (cubic)

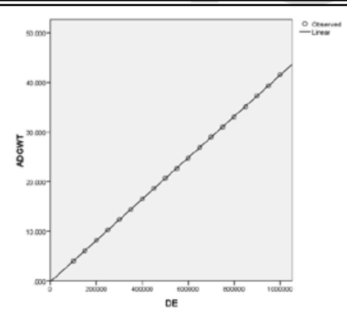

Fig. 21. Graphical representation of MODEL2 (linear)

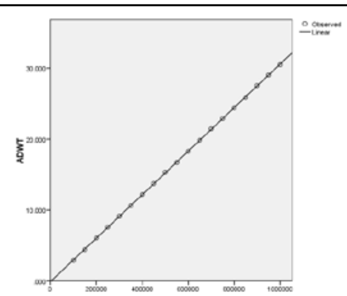

Fig. 22. Graphical representation of MODEL3 (linear)

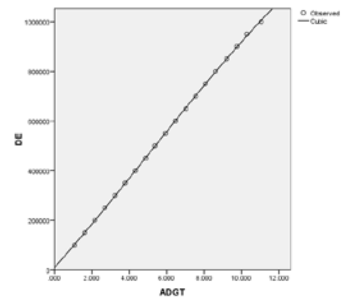

Fig. 23. Graphical representation of MODEL4 (cubic)

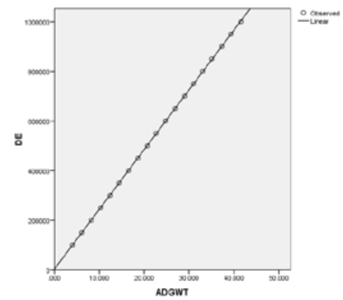

Fig. 24. Graphical representation of MODEL5 (linear)

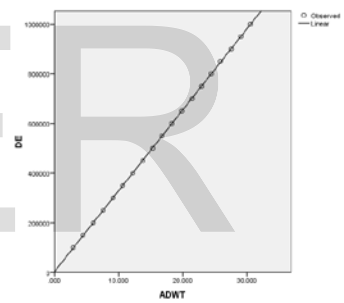

Fig. 25. Graphical representation of MODEL6 (linear)

6 LIMITATIONS & FUTURE SCOPE

We have done our performance measurement on a particular hardware and a particular operating system. Therefore the result obtained in the above corresponds to this particular configuration. In this study, we have only measured time, keeping aside other important performance measurement matrix e.g. percentage CPU utilization, percentage memory utilization etc. The entire performance measurements of data generation and writing those data in a file have been done in GUI mode of the operating system. While doing these measurements, we have not run any application programs. Hence, what happens in case of single or multiple application programs runs in the background while measuring the performance are yet to be explored.

Hence, we are looking forward to do the performance measurement considering these above factors in future because we believe that these factors may contribute to the actual performance.

6 Conclusion

In this paper we have used curve estimation technique for analyzing the dataset. The researchers have proposed two different types of performance measurement and management models for a particular platform (hardware and operating system) in this study. The first type is in the form of 'Time' ~ f('Number of Data Elements') and the second type is in the form of 'Number of Data Elements' ~ f('Time'). In total there are six (6) models, three (3) from each type. These six models give us the following information: how much time is required to generate a given amount of numeric data (MODEL1), how much time is required to generate and write a given amount of numeric data in a text file (MODEL2), how much time is required to write a given amount of numeric data in a text file (MODEL3), how much amount of numeric data can be generated in a given time (MODEL4), how much amount of numeric data can be generated and written in a given time (MODEL5) and how much amount of numeric data can be written in a given time (MODEL6).

The proposed models may give us comparatively easy way to measure the performance of a particular platform (hardware and software) while generating relatively large number of numeric (integer) data and writing these data in file. At any point of time we may easily measure the performance matrix of the said platform by using the technique discussed in the sub-section 4.1 and compare these measured parameters with the model generated values. If the two results deviate by an acceptable tolerance limit then we may conclude that the said platform needs performance management and appropriate actions may be taken to answer the issues.

References

[1] Performance Measurement. (n.d.). Retrieved May 2, 2014, from http://www.businessballs.com/dtiresources/performance_measurement_management.pdf

[2] Performance measurement. (2014). *Wikipedia*. Retrieved May 2, 2014, from http://en.wikipedia.org/wiki/Performance_measurement

[3] What is Performance Measurement? - BPIR.com. (n.d.). *BPIR.com, Business Performance Improvement Resource*. Retrieved May 3, 2014, from http://www.bpir.com/what-is-performance-measurement-bpir.com.html

[4] Performance Management & Measurement. (n.d.). *HRSA, U.S. Department of Health and Human Services, Health Resources and Services Administration*. Retrieved May 3, 2014, from http://www.hrsa.gov/quality/toolbox/methodology/performance management/index.html

[5] Information society. (2014). *Wikipedia*. Retrieved May 3, 2014, from http://en.wikipedia.org/wiki/Information_society

[6] Data vs. Information. (n.d.). *Diffen*. Retrieved May 3, 2014, from http://www.diffen.com/difference/Data_vs_Information

[7] SINTEF. (2013, May 22). Big Data, for better or worse: 90% of world's data generated over last two years. *ScienceDaily*. Retrieved May 12, 2014 from www.sciencedaily.com/releases/2013/05/130522085217.htm

[8] Heo, J., & Han, I. (2003). Performance measure of information systems (IS) in evolving computing environments: an empirical investigation. *Information & Management*, *40*(4), 243-256. Retrieved May 3, 2014, from http://www.sciencedirect.com/science/article/pii/S0378720602000071

[9] Prybutok, V. R., Kappelman, L. A., & Myers, B. L. (1997). A comprehensive model for assessing the quality and productivity of the information systems function: toward a theory for information systems assessment. *Information Resources Management Journal*, *10*(1), 6 - 26. Retrieved May 3, 2014, from http://dl.acm.org/citation.cfm?id=2444747

[10] Saarinen, T. (1996). An expanded instrument for evaluating information system success. *Information & Management*, *31*(2), 103-118. Retrieved May 5, 2014, from http://www.sciencedirect.com/science/article/pii/S0378720696010750

[11] DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, *19*(4), 9 - 30. Retrieved May 5, 2014, from http://www.asiaa.sinica.edu.tw/~ccchiang/GILIS/LIS/p9-Delone.pdf

[12] Etezadi-Amoli, J., & Farhoomand, A. F. (1996). A structural model of end user computing satisfaction and user performance. *Information & Management*, *30*(2), 65-73. Retrieved May 5, 2014, from http://www.sciencedirect.com/science/article/pii/0378720695000526

[13] Patel, S. C., & Maheshwari, S. K. (2000). Information system and performance measurement life cycle. *Delhi Business Review*, *1*(2). Retrieved May 6, 2014, from http://www.dbr.shtr.org/v_1n2/dbrv1n2b.pdf

[14] Sharma, M. K., & Bhagwat, R. (2006). Performance measurements in the implementation of information systems in small and medium-sized enterprises: a framework and empirical analysis. *Measuring Business Excellence*, *10*(4), 8-21. Retrieved May 7, 2014, from http://www.emeraldinsight.com/journals.htm?articleid=1581912

[15] Palmius, J. (2007). Criteria for measuring and comparing information systems. *Proceedings of the 30th Information Systems Research Seminar in Scandinavia IRIS 2007*. Retrieved May 7, 2014, from http://www.palmius.com/joel/text/IRIS-30-final.pdf

[16] Dementiev, R., Willhalm, T., Bruggeman, O., Fay, P., Ungerer, P., Ott, A., et al. (2012, August 16). Intel® Performance Counter Monitor - A better way to measure CPU utilization. *Intel Developer Zone*. Retrieved May 2, 2014, from https://software.intel.com/en-us/articles/intel-performance-counter-monitor-a-better-way-to-measure-cpu-utilization

[17] Song, Y. (2012, August 2). Measure Intel® IPP Function Performance. *Intel Developer Zone*. Retrieved May 3, 2014, from https://software.intel.com/en-us/articles/measure-intel-ipp-function-performance

[18] Intel® Performance Tuning Utility 4.0 Update 5. (n.d.). *Intel Developer Zone*. Retrieved May 12, 2014, from https://software.intel.com/en-us/articles/intel-performance-tuning-utility

[19] Intel® Graphics Performance Analyzers. (n.d.). *Intel Developer Zone*. Retrieved May 12, 2014, from https://software.intel.com/en-us/vcsource/tools/intel-gpa

[20] Intel® VTune™ Amplifier XE 2013. (n.d.). *Intel Developer Zone*. Retrieved May 12, 2014, from https://software.intel.com/en-us/intel-vtune-amplifier-xe

[21] Ubuntu Manpage: top. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/raring/en/man1/top.1.html

[22] Ubuntu Manpage: vmstat. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/hardy/man8/vmstat.8.html

[23] Ubuntu Manpage: lsof. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/hardy/man8/lsof.8.html

[24] Ubuntu Manpage: tcpdump. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/lucid/en/man8/tcpdump.8.html
[25] Ubuntu Manpage: netstat. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/precise/man8/netstat.8.html
[26] Ubuntu Manpage: htop. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/lucid/man1/htop.1.html
[27] iotop – simple top-like I/O monitor. (n.d.). *Ubuntu Geek*. Retrieved May 15, 2014, from http://www.ubuntugeek.com/iotop-simple-top-like-io-monitor.html
[28] Saive, R. (2012, September 5). Linux Performance Monitoring with Vmstat and Iostat Commands. *Techmint.com, Linux Howto's Guide*. Retrieved May 15, 2014, from http://www.tecmint.com/linux-performance-monitoring-with-vmstat-and-iostat-commands/
[29] Ubuntu Manpage: mpstat. (n.d.). *Ubuntu manuals*. Retrieved May 15, 2014, from http://manpages.ubuntu.com/manpages/gutsy/man1/mpstat.1.html
[30] Install and Use nmon Tool To Monitor Linux Systems Performance. (2012, August 5). *nixcraft, Insight Into Linux Admin Work*. Retrieved May 17, 2014, from http://www.cyberciti.biz/faq/nmon-performance-analyzer-linux-server-tool/
[31] Starkweather, J. (n.d.). Model Specification Error...Are you straight, or do you have curves?. Retrieved May 2, 2014, from http://www.unt.edu/rss/class/Jon/Benchmarks/Model%20Specification%20Error_JSD_Apr2010.pdf
[32] Teknomo, K. (n.d.). How can we be sure that the best line is linear?. *Kardi Teknomo's Page*. Retrieved May 17, 2014, from http://people.revoledu.com/kardi/tutorial/Regression/GoodnessOfFit.html
[33] 4.4.4. How can I tell if a model fits my data?. (n.d.). *Engineering Statistics Handbook*. Retrieved May 17, 2014, from http://www.itl.nist.gov/div898/handbook/pmd/section4/pmd44.htm
[34] Williams, T. (n.d.). Residual analysis. *Encyclopedia Britannica Online*. Retrieved May 17, 2014, from http://www.britannica.com/EBchecked/topic/564172/statistics/60718/Residual-analysis
[35] 4.4.4.1. How can I assess the sufficiency of the functional part of the model?. (n.d.). *Engineering Statistics Handbook*. Retrieved May 17, 2014, from http://www.itl.nist.gov/div898/handbook/pmd/section4/pmd441.htm
[36] Normal probability plot of residuals. (n.d.). *PENNSTATE, STAT 501 - Regression Methods*. Retrieved May 17, 2014, from https://onlinecourses.science.psu.edu/stat501/node/40
[37] Bandyopadhyay, G. (2013). Modeling NPA Time Series Data in Selected P ublic S ector B ank s in India with Semi Parametric Approach. *International journal of scientific and engineering research*, *4*(12), 1876-1889. Retrieved May 17, 2014, from http://www.ijser.org/researchpaper%5CModeling-NPA-Time-Series-Data-in-Selected-Public-Sector.pdf
[38] Testing for Normality using SPSS Statistics. (n.d.). *Laerd statistics*. Retrieved May 17, 2014, from https://statistics.laerd.com/spss-tutorials/testing-for-normality-using-spss-statistics.php
[39] Das, D., Chakraborty, A., & Mitra, A. (2014). Sample Based Curve Fitting Computation on the Performance of Quicksort in Personal Computer. *International journal of scientific and engineering research*, *5*(2), 885-891. Retrieved May 17, 2014, from http://www.ijser.org/onlineResearchPaperViewer.aspx?Sample-Based-Curve-Fitting-Computation-on-the-Performance.pdf