Journal of Mathematical Modelling and Applied Computing

Volume No. 12

Issue No. 2

May - August 2024

ENRICHED PUBLICATIONS PVT.LTD

JE - 18,Gupta Colony, Khirki Extn,
Malviya Nagar, New Delhi - 110017.
E- Mail: info@enrichedpublication.com
Phone :- +91-8877340707

Journal of Mathematical Modelling and Applied Computing

Aims and Scope

The Journal of Mathematical Modelling and Applied Computing is an Indian research journal, which publishes top-level original and review papers, short communications and proceedings on Interdisciplinary Integrative Forum on Modelling, Simulation and Scientific Computing in Engineering, Physical, Chemical Biological, Medical, Environmental, Social, Economic and Other Systems using Applied Mathematics and Computational Sciences and Technology.

Journal of Mathematical Modelling and Applied Computing

Journal of Mathematical Modelling and Applied Computing

(Volume No. 12, Issue No. 2, May - August 2024)

Contents

Comparing Linear Mixed Model and Generalized Estimating Equations Techniques for Longitudinal Data

Kamal Kishore[1], Dr. D. K. Subbakrishna[2], Dr. P. Marimuthu[3]

[1]Assistant Professor, Institute of Management, Christ University, Bengaluru.

[2]Professor, Department of Biostatistics, NIMHANS, Bengaluru.

[3]Additional Prof., Department of Biostatistics, NIMHANS, Bengaluru.

ABSTRACT

The defining feature of longitudinal studies is correlated nature of data as response variable from same subjects are collected repeatedly on subsequent occasions. Moreover, a peculiar characteristic of longitudinal data is missing pattern. Mixed effect models (MEM) and generalized estimating equations (GEE) are two advanced statistical techniques which are popular among researchers for analysis of longitudinal data. These techniques were used to analyze longitudinal dataset on 95 subjects measured on six occasions. Although, these techniques are not comparable in general due to the different assumptions they make about the data. MEM and GEE techniques are comparable when fitted to same longitudinal data. The main focus of the study lies in the group growth trajectory. These approaches were compared on the basis of initial status and rate of growth of parameters, which are representative of a growth profile. Applied researchers are often confused as to which method performs better under different conditions so as to use them for the analysis of data. An attempt is made to demonstrate the differences and similarities in both the techniques for analyzing longitudinal data with and without missing pattern in the data.

Keywords: Longitudinal data, MEM, GEE, Traditional methods

INTRODUCTION

There has been tremendous growth in interest and statistical methodology for the longitudinal data analysis from last few years. One important reason for wide use of longitudinal studies can be attributed to its use to address the substantive research question about change over a period of time. Moreover, popularity is on the rise due to tremendous growth in literature and computational power at hands of end users. This growth leads to integration of advanced statistical techniques in popular software, thereby increasing the application of these methods.

Mixed effect models (MEM) and generalized estimating equations (GEE) approaches provide researchers with powerful and flexible analytic tools for the characteristics displayed by longitudinal data. MEM is an extension of the regression model in which dependent nature of subsequent observations from same subject are accomplished by introducing random effects. Moreover, random effects partitions total variability into within subject variability and between subject variability. MEM are very popular among the applied researchers and these are nicely described in literature by several authors (Fitzmaurice et al. 2003; Singer and Willett 2003; Twisk 2004; Fitzmaurice et al., 2009). GEE is a quasi-likelihood approach (Liang and Zeger, 1986; Zeger and Liang, 1986) which can pursue statistical models by making assumptions about the link function and the relationship between the first two moments, but without specifying the complete distribution of the response. GEE models are quite different than MEM, and these are described in literature by several authors (Diggle et al., 2002; Hardin and Hilbe, 2003; Liang and Zeger, 1993; Zeger et al., 1988). In most of applied research, the interest lies in the shape of population growth trajectory rather than individual growth trajectory over time and correlation among repeated measures. In this paper, an attempt is made to compare the performances of MEM and GEE, when research interest lie in group growth trajectory.

Before development of these advanced statistical techniques, most of the analysis for longitudinal data was carried out with traditional methods such as paired t-test and RM-ANOVA. An important characteristic of using advanced techniques are the way they handle dependence and missingness as compared to traditional methods in longitudinal data. When the subjects are repeatedly measured some of them miss schedules and some may permanently drop out from study due to one or the other reason. Unlike the traditional methods such as paired t-test, RM- ANOVA technique, these statistical techniques use all the subjects for analysis even if responseis available for the single time point. Moreover, these techniques can be applied efficiently when outcome variable is categorical in nature.

MEM and GEE approach was developed in the beginning for different type of longitudinal outcome variable. MEM was initially developed for the analyses of normally continuous outcome longitudinal or clustered variable (Laird and Ware 1982; Bryk and Raudenbush 1992; Goldstein 1999) but were less familiar for analyses of non-normal correlated data. During same time, a quasi-likelihood estimation procedure known as generalized estimating equations (GEE), first introduced by Liang and Zeger(1986; Zeger and Liang 1986) have become very popular to estimate regression coefficients for analysis of longitudinal categorical data. These boundaries are now blurring due to growth in literature as both procedures are applicable for continuous as well as categorical data.

The idea of change and measurement of change is an intriguing concept which fascinate researchers from generations. This article compare performance of MEM and GEE techniques to measure change. Linear growth trajectory and quadratic growth trajectory are two commonly used assumptions to measure change. The linear growth trajectory assumes monotonic increase whereas for quadratic growth trajectory change increases upto a certain time and then levels off gradually.

METHODOLOGY

A MEM is appropriate statistical model for longitudinal data as it takes dependence into accounts through random effects. These models can be represented in mathematical form as:

$$Y_i = X_i\beta + Z_i b_i + \varepsilon_i,\ \ i = 1,2,\dots\dots N \qquad (1)$$

where $Y_i = (y_{i1}, y_{i2}, \dots\dots, y_{in_i})'$, $X_i = (x_{i1}, x_{i2}, \dots\dots, x_{in_i})'$ is a n_ix p design matrix of known covariates, β is a p x 1 vector of fixed regression parameters, $Z_i$is a $n_i$xq design matrix for random effects, b_i is a q x 1 vector of random regression coefficients distributed as N(0, ∑) and b_is are mutually independent, and ε_i is n_i x 1 vector of random errors distributed as $N(0, \sigma^2 In_i)$ and is independent of b_is. The variance-covariance matrix ∑ captures the degree of heterogeneity of subjects. It is important to note that $E(Y_{ij}) = E[E(Y_{ij}|b_i)] = x'_{ij}\beta$ and therefore marginal and conditional parameter are equal.

The MEM for binary outcome variable is an extension of MEM and generalized linear models for correlated non-Gaussian outcome variable. MEM for binary outcome variable can be specified with exponential family distribution for outcome variable, a link function and random effect structure. A MEM for binary outcome variable can be expressed with logit link as follows:

$$E(Y_{ij}) = E[E(Y_{ij}|b_i)] = E\left[\frac{\exp(x'_{ij}\alpha - z'_{ij}b_i)}{1+\exp(x'_{ij}\alpha - z'_{ij}b_i)}\right] = \frac{\exp(x'_{ij}\alpha)}{1+\exp(x'_{ij}\alpha)}$$

GEE is a quasi-likelihood approach (Liang and Zeger, 1986; Zeger and Liang, 1986) which can pursue statistical models by making assumptions about the link function and the relationship between the first two moments, but without specifying the complete distribution of the response. A GEE model for longitudinal data has three part specification:

- Mean of each response is assumed to depend on the covariates through link function.

$$g(\mu)_{ij} = x'_{ij}\beta$$

- The conditional variance of, given the y_{ij} covariates is

$$Var(y_{ij}) = \emptyset\ var(\mu_{ij})$$

where Ø is known or estimated scale parameter and $var(\mu_{ij})$ is known variance function.

- With-in subject association among the vectors of repeated responses over time.

DATA AND ANALYSIS

The data for the present study is utilized from Neurological performance of cohort consist of 95 HIV-1 infected individuals observed over 21/2 years in southern India. In the dataset the subjects were recruited at baseline and followed-up 5 times at six monthly intervals. The objective of research is to investigate the change in neuro-psychological performance over a period of time and the factors that influence change in group. The explanatory variables are either continuous or categorical and time dependent or time independent in nature.

The two advanced statistical techniques MEM and GEE for longitudinal outcome variable with continuous and categorical outcome variable were applied to neuro-psychological performance. The neuro-psychological performance was dichotomized in a way where each measurement of upper tertiles compared to two lower tertile. The dataset for comparing the two techniques were analyzed using "geepack"(Generalized & Equation, 2012; Halekoh, Højsgaard, & Yan, 2006; Højsgaard, 2011) and "lme4" (D Bates, Maechler, Bolker, & Walker, 2013; Douglas Bates, 2011) package of free and flexible R-software..

MEM was developed initially for continuous data, whereas GEE was primarily developed for categorical data. The developments in literature lead to extension of both these techniques for analysis of categorical and continuous data. MEM and GEE generate subject-specific parameter and population specific parameter estimate respectively. Population specific parameter estimates can be obtained from subject specific parameter estimates however, population specific estimates of MEM for categorical data are biased (Agresti 2002).

One of the major challenges in dealing with longitudinal studies is of missing data and the major difference between MEM and GEE is the way they handle missing data. In this paper original dataset analysis is followed by selecting subsets (incomplete datasets) of data from original data. The incomplete datasets were obtained from complete datasets by deleting approximately 21% (N=40) of the observations from fourth occasion. MCAR is the strongest assumption for the data which is rarely met by data.Data following MCAR mechanism can be thought of random sample of the complete data. MCAR data was generated by omitting values completely at random from complete data. Data are said to be missing at random (MAR) when the probability that response are missing depends on the set of observed responses, but is unrelated to the specific missing value that in principle, should have been obtained (G. M. Fitzmaurice, Laird, & Ware, 2004). MAR mechanism occurs more frequently in longitudinal studies. One of the major consequences of MAR is that complete cases are not random samples and it can lead to biased estimates of change in mean response over time. MAR data was generated by arranging the value in ascending order at 3rd occasion and then subsequently removing the lowest 40 values from the fourth occasion onwards. The data is categorized before omitting the values and then same values are omitted which were omitted for continuous outcome variable.

RESULTS

There were total 95 subjects at the beginning of study. Demographic profile of the group suggests comparable gender distribution of male (n=58; 61%) and female (n=37; 39%). The subjects were in the age group of 20-45 years, 58% of the subjects were married and most of them were from urban population (n=40; 42%) followed by rural (n=26; 27%). Further, study group consisted of majority of subjects from high literacy group (n=60; 63%) and most of them were from nuclear families (n=52; 55%). Longitudinal data used for analysis showed significant correlation over a period of time as can be seen below in the table (1).

Table 1. Observed Inter-occasion Correlation Coefficient For Outcome Variable

	Time1	Time2	Time3	Time4	Time5	Time6
Time1	-	0.807**	0.784**	0.582**	0.467**	0.6**
Time2		-	0.792**	0.697**	0.609**	0.645**
Time3			-	0.731**	0.599**	0.643**
Time4				-	0.685**	0.565**
Time5					-	0.585**
Time6						-

Individual and Group Profile

The facts about mean response trajectory can be visualized from figure (1). It may be noted that mean neuropsychological performance increases smoothly upto certain period and then decreases. This figure indicates that response is not linear with time and polynomial growth curve can be a possibilty to account for curviliear nature of change. However, the final decision to include polynomial trajectory is to be made on the basis of testing procedure. In order to visualize how different persons change over a period of time empirical growth plots of 12 randomly selected subjects are shown here in figure 2.

Figure (1)

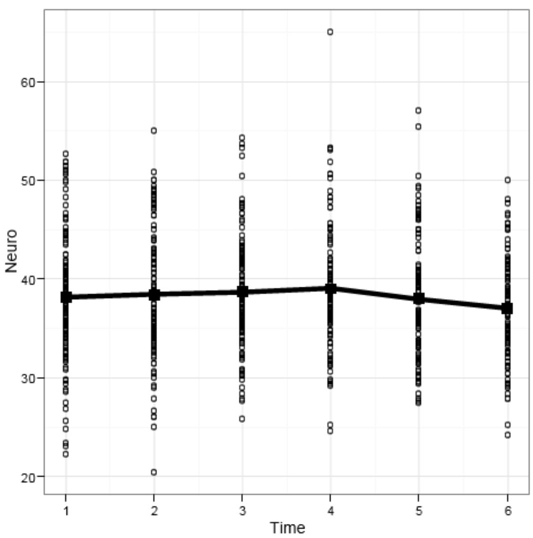
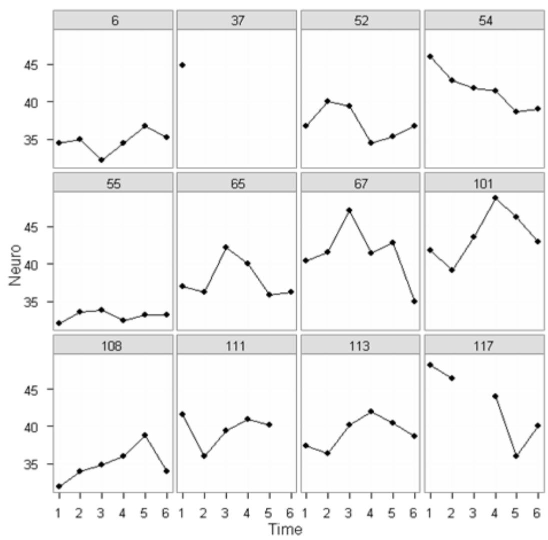


Figure (2)



Comparing Performances of Mem and Gee for Continuous Outcome Variable

The main goal of the study was to infer, how the covariates are related to change in outcome variable. There were not any interest in interactions of covariates with time.The covariates of interest to study are gender, education level, income, locality, family, BDI score, Marital status, MMSE score and CD4 count at baseline. Mathematically a model for continuous outcome for linear growth trajectory can be wriiten as:

$$Y_{ij} = \beta_0 + \beta_1(\text{Time}) + \beta_3(\text{Gender}) + \beta_4(\text{Edu}) + \beta_5(\text{Income}) + \beta_6(\text{Locality}) + \beta_7(\text{Family}) + \beta_8(\text{BDI}) + \beta_9(\text{MMSE}) + \beta_{10}(\text{CD4C}) + \beta_{11}(\text{Mstatus}) + \varepsilon_{ij} \quad (1)$$

A MEM with random intercept is equivalent to exchangeable correlation structure in GEE and MEM with random intercept and random slope is equivalent to unstructured correlation in GEE. The precision of parameter estimates are inversely proportional to standard error, more the standard error lesser the precision and vice–versa. The results obtained after application of MEM and GEE for linear growth trajectory and quadratic growth trajectory for continuous and categorical outcome variable are displayed and discussed subsequently. Missing data is almost always part of longitudinal studies, thus two incomplete data (MCAR and MAR) are generated so as to observe how these techniques handle missing data.

Table: 2

Model		Random Intercept			Random Intercept and Slope		
		Complete	MCAR	MAR	Complete	MCAR	MAR
Intercept	MEM	33.93 (0.87)	33.66 (0.88)	33.53 (0.95)	34.15 (0.92)	33.92 (0.93)	33.74 (1.00)
	GEE	33.93 (0.82)	33.67 (0.82)	33.52 (0.90)	33.51 (0.84)	33.55 (0.78)	33.65 (0.88)
Slope	MEM	-0.15 (0.09)	-0.13 (0.10)	-0.15 (0.11)	-0.15 (0.11)	-0.12 (0.12)	-0.14 (0.13)
	GEE	-0.14 (0.10)	-0.13 (0.12)	-0.15 (0.13)	-0.06 (0.10)	-0.11 (0.12)	-0.2 (0.13)
Gender	MEM	5.08 (0.88)	5.07 (0.87)	5.27 (0.91)	4.93 (0.87)	4.83 (0.86)	5.05 (0.90)
	GEE	5.08 (0.90)	5.07 (0.89)	5.28 (0.92)	4.59 (0.87)	4.64 (0.86)	4.38 (0.89)
Education	MEM	4.25 (0.89)	4.33 (0.88)	4.69 (0.94)	3.99 (0.88)	4.11 (0.87)	4.5 (0.92)
	GEE	4.25 (0.87)	4.33 (0.89)	4.68 (0.90)	4.18 (0.85)	4.7 (0.80)	4.72 (0.87)

In the table (2), parameter estimates and their respective standard errors obtained by MEM and GEE techniques are presented. The exchangeable and unstructured correlation structure models of GEE with equivalent representative models from MEM were compared against each other. The parameter estimates and their standard errors are almost same with GEE and MEM techniques. In the literature it is emphasized that a major difference between MEM and GEE is the way they handle missing data, the former can handle data with MAR mechanism but not the later one. From table (2) it can be seen that GEE estimates and their standard errors are nearby to MEM for data with MAR mechanism. There are some differences in parameter estimates and their standard errors but, these differences can be attributed to maximum likelihood estimation method of MEM and quasi-likelihood estimation method for GEE. The GEE model is more flexible than MEM in making assumptions about the data.

Table (3) is displaying the results for quadratic growth curve models. The model can be obtained by assuming quadratic growth trajectory and adding the same in equation (1). It can be observed from the table that both the techniques are giving almost the same results. Moreover, standard error is more for data with MAR mechanism.

Table (3)

Model		Random Intercept			Random Intercept and Slope		
		Complete	MCAR	MAR	Complete	MCAR	MAR
Intercept	MEM	31.05 (1.03)	30.65 (1.06)	29.99 (1.15)	31.27 (1.05)	30.97 (1.09)	30.14 (1.18)
	GEE	31.05 (1.00)	30.65 (1.00)	29.98 (1.10)	31.43 (0.98)	30.87 (0.95)	30.57 (1.14)
Slope	MEM	2.02 (0.42)	2.16 (0.47)	2.59 (0.52)	2.02 (0.4)	2.1 (0.45)	2.59 (0.5)
	GEE	2.02 (0.43)	2.16 (0.44)	2.59 (0.51)	1.87 (0.39)	1.98 (0.42)	2.3 (0.56)
Quadratic	MEM	-0.31 (0.06)	-0.33 (0.07)	-0.39 (0.07)	-0.31 (0.05)	-0.32 (0.06)	-0.39 (0.07)
	GEE	-0.31 (0.06)	-0.33 (0.06)	-0.39 (0.07)	-0.29 (0.05)	-0.29 (0.06)	-0.34 (0.08)
Gender	MEM	5.08 (0.88)	5.05 (0.87)	5.25 (0.92)	4.93 (0.87)	4.82 (0.86)	5.07 (0.9)
	GEE	5.08 (0.9)	5.05 (0.89)	5.26 (0.92)	4.57 (0.87)	4.58 (0.86)	4.43 (0.89)
Education	MEM	4.25 (0.89)	4.32 (0.88)	4.56 (0.94)	3.99 (0.88)	4.11 (0.87)	4.4 (0.92)
	GEE	4.25 (0.87)	4.32 (0.85)	4.55 (0.9)	4.2 (0.85)	4.64 (0.8)	4.33 (0.86)

The positive parameter estimate for slope component and negative parameter estimate for quadratic component is indicator of growth rate in beginning and then decrease subsequently.It is emphasized in literature that GEE technique is robust to misclassification of correlation structure, but here this trend is also observed for MEM technique. From the table (3) it can be said that GEE and MEM estimates and their standard errors are very stable and almost near for Continuous outcome variable for both complete and incomplete datasets. Thus it was observed that GEE and MEM estimates and their standard errors are not very different from each other for complete and Incomplete (MCAR and MAR) datasets, when the growth trajectories assumed were linear and quadratic.

Comparing Performances of MEM and GEE for binary outcome variable

The model building is done to summarize the important characteristics of data with parsimonious model which should have all the relevant parameters. Up to this point, the repeatedmeasures analysis of the variables was restricted for continuous response variable. However, in practice, categorical variables are also very common in applied research. Because the response variable can take one of two values it cannot be interpreted like continuous response variable. The model for binary response variable was adapted to predict the probability of positive response (D. M. Bates, 2010). A data set was analyzed in which subjects are assesed over a period of time for binary outcome. A marginal logistic regression model for binary response variable with all the covariates of interest is represented by equation 2. The interaction effect among covariates was not of interest and was not assumed for estimation process.

$$\text{logit}[p(Y_{ij}=1)]=\beta_0+\beta_1(\text{Time}_{ij})+\beta_3(\text{Gender})+\beta_4(\text{Edu})+\beta_5(\text{Income})+\beta_6(\text{Locality})+ \beta_7(\text{Family})+ \beta_8(\text{BDI})+ \beta_9(\text{MMSE})+ \beta_{10}(\text{CD4C})+ \beta_{11}(\text{Mstatus})+ \varepsilon_{ij} \quad (2)$$

The parameter estimates for categorical data for linear growth trajectory are displayed in table (4). It can be observed from table that unlike estimates and their standard errors for continuous data, standard error estimates obtained with GEE technique is always better. GEE estimates for data with MAR mechanism holds for categorical data also. The criticism in literature against usage of GEE for data missing at random is not supported by this study. According to this study parameter estimates and their standard errors obtained by GEE are as efficient as estimates obtained by MEM.

Table (4)

Model		Random Intercept			Random Intercept and Slope		
		Complete	MCAR	MAR	Complete	MCAR	MAR
Intercept	MEM	-0.80 (0.51)	-0.99 (0.52)	-1.04 (0.61)	-0.41 (0.58)	-0.69 (0.60)	-0.79 (0.72)
	GEE	-0.59 (0.34)	-0.69 (0.35)	-0.67 (0.39)	-0.5 (0.33)	-0.66 (0.35)	-0.76 (0.40)
Slope	MEM	-0.06 (0.07)	-0.04 (0.08)	-0.04 (0.09)	-0.12 (0.08)	-0.09 (0.09)	-0.11 (0.11)
	GEE	-0.04 (0.04)	-0.03 (0.05)	-0.03 (0.06)	-0.03 (0.04)	-0.03 (0.05)	-0.01 (0.06)
Gender	MEM	2.31 (0.52)	2.25 (0.53)	2.49 (0.59)	2.28 (0.53)	2.2 (.54)	2.67 (0.65)
	GEE	1.51 (0.35)	1.47(0.35)	1.54(0.36)	1.48 (0.34)	1.45 (0.34)	1.55 (0.37)
Education	MEM	2.14 (0.05)	2.22 (0.51)	2.43 (0.58)	2.02 (0.51)	2.19 (0.52)	2.61 (0.63)
	GEE	1.45 (0.32)	1.48 (0.33)	1.51 (0.35)	1.33 (0.32)	1.49 (0.33)	1.52 (0.36)

Similarly parameter estimates and standard errors obtained with both the techniques for quadratic growth trajectory are displayed in table (5).

Table (5)

Model		Random Intercept			Random Intercept and Slope		
		Complete	MCAR	MAR	Complete	MCAR	MAR
Intercept	MEM	-2.72(0.71)	-2.52(0.74)	-2.99(0.87)	-2.49(0.77)	-2.32(0.82)	-2.88(1.02)
	GEE	-1.67(0.43)	-1.58(0.46)	-1.64(0.45)	-1.45(0.37)	-1.41(0.43)	-1.68(0.45)
Slope	MEM	1.35(0.35)	1.09(0.37)	1.40(0.45)	1.34(0.36)	1.08(0.39)	1.43(0.49)
	GEE	0.82(0.22)	0.67(0.24)	0.78(0.24)	0.71(0.19)	0.56(0.22)	0.81(0.24)
Quadratic	MEM	-0.20(0.05)	-0.16(0.05)	-0.21(0.06)	-0.20(0.05)	-0.17(0.05)	-0.22(0.07)
	GEE	-0.12(0.03)	-0.10(0.03)	-0.12(0.04)	-0.11(0.03)	-0.08(0.03)	-0.12(0.03)
Gender	MEM	2.43(0.55)	2.33(0.55)	2.64(0.63)	2.45(0.57)	2.27(0.56)	2.80(0.68)
	GEE	1.44(0.35)	1.42(0.35)	1.44(0.36)	1.43(0.34)	1.39(0.34)	1.44(0.36)
Education	MEM	2.25(0.53)	2.29(0.53)	2.52(0.62)	2.17(0.54)	2.25(0.54)	2.68(0.67)
	GEE	1.38(0.32)	1.43(0.33)	1.38(0.35)	1.26(0.32)	1.45(0.32)	1.40(0.35)

The parameter estimates and their standard errors obtained with GEE technique for categorical outcome are always lower than MEM, whereas this trend was not noticed for continuous outcome variables. The standard errors are lower and more stabilized for parameter estimates obtained with GEE technique. The incomplete data had marginal effect on the parameter estimates and their standard errors. The standard errors are higher for data with MAR mechanism for both GEE and MEM. It is surprising to note that, parameter estimates and their standard errors with MEM and GEE techniques are quite different for binary outcome variables compared to continuous outcome variable. More surprising are the results for GEE technique applicable to data with MAR mechanism as these are better than results obtained with MEM. Thus, it was observed that parameter estimates and their standard errors obtained by GEE technique are more stable and efficient than MEM technique for categorical data. Moreover, surprisingly same trend was observed for incomplete data in general and data with MAR mechanism in particular.

DISCUSSION

Both MEM and GEE techniques are highly suitable and preferred among researchers for analyzing longitudinal data. The question arises: which one of these techniques is better? In this paper these two popular statistical techniques for analysis of longitudinal data are compared and discussed. MEM and GEE techniques are very appealing as they include time and their transformations to infer about group growth trajectories. Moreover, both techniques use all the available data for analysis even if data is not available for all but one occasion. The MEM is an extension of linear regression which was extended by adding random effect in regression for longitudinal continuous outcome. Whereas during same time, GEE technique was developed for categorical data where generalized linear models (GLM) was extended by incorporating correlation structure. MEM and GEE approaches were used on empirical longitudinal dataset for comparison of two approaches.

The GEE approach does not make distributional assumptions because estimation of population-averaged model depends on correct specification of few aspects of observed data and it does not depend on the entire data generating distribution. GEE models for mean response depend on the predictors of interest rather than on random effects or previous responses. It requires only a regression model for mean response and does not require distributional assumptions for the data (G. M. Fitzmaurice et al., 2004). On the other hand MEM requires correct specification of distributional assumptions about random effects besides usual data distributional assumptions. The parameter estimates for fixed effect can be misleading and may lead to biased inference as they depends on correct specification of non-identifiable latent random effect for which even large sample size do not help (Hubbard et al., 2010)

Since, interpretation of parameter estimates is different for population-averaged model as compared to subject specific models, their selection is of vital importance to address question of scientific interest. The MEM models are subject-specific whereas GEE are population-averaged model. The researcher should clearly state about the interest of analysis. In case of continuous outcome variable subject-specific estimates can be averaged to get population averaged estimates but same cannot be done for categorical data without introducing bias. Both statistical techniques resulted in comparable parameter estimates and their standard errors for continuous outcome variable. GEE estimates are more efficient for categorical outcome variable but these are population average estimates. Whenever, the research interest lies in individual change MEM are preferred choice.

Missing data is one of the major challenges in longitudinal studies. Both techniques use all the available data for analysis. In literature, it is emphasized that GEE can handle data with MCAR mechanism, whereas MEM can handle data with MAR mechanism. It is observed here that parameter estimates and standard errors with GEE and MEM are slightly different for complete and incomplete continuous datasets. Further, it was observed that there was not much difference in both the techniques for random missing data (MCAR) as compared to dataset with selective missingness (MAR). The same trend was noticed for binary outcome variable for both the techniques but estimates obtained with GEE technique were more efficient as compared to MEM estimates. These results are rather surprising in contrast to existing literature as GEE performed consistently for selective missingness (MAR) in case of both continuous and categorical outcome variable.

CONCLUSIONS

GEE and MEM techniques for continuous outcome variable leads to similar results for complete and incomplete datasets. Any technique can be applied when interest lies in population averaged trajectory. MEM is preferred when research interest is in individual growth trajectory. For categorical growth outcome variable GEE technique is more efficient as compared to MEM technique. Moreover, MEM approach brings bias when averaged to group growth trajectory. It is to be used only when research interest is in individual growth trajectory.

BIBLIOGRAPHY

- *Bates, D. M., 2010, lme4: Mixed-effects modeling with R. URL http://lme4. R-Forge. R-Project.Org/book.*
- *Bates, D., 2011, Mixed models in R using the lme4 package Part 1 : Linear mixed models with simple , scalar random effects. Development, pp 1–27.*
- *Bates, D., Maechler, M., Bolker, B., & Walker, S., 2013. lme4: Linear mixed-effects models using Eigen and S4. R Package Version, pp 0–1.*
- *Bryk, A. S., & Raudenbush, S. W.,(1992),Hierarchical Linear Models. 1st. Ed. Newbury Park, CA: Sage Publications.*
- *Diggle, P., Heagerty, P. J., Liang, K., & Zeger, S. L., (2002). Analysis of longitudinal data, 3rd.*
- *Ed., Oxford University Press, USA.*
- *Fitzmaurice, G., Laird, N., & Ware, J., (2003). Applied longitudinal analysis, 2nd. Ed., John Wiley & Sons.*
- *Goldstein, H.,(1999). Multilevel Statistical Models, 3rd. Ed., London: Edward Arnold.*
- *Halekoh, U., Højsgaard, S., & Yan, J., 2006, The R package geepack for generalized estimating equations. Journal of Statistical Software, 15(2), pp 1–11.*
- *Hardin, J. W., & Hilbe, J. M.,(2003). Generalized Estimating Equations, 2nd. Ed., Chapman and Hall/CRC.*
- *Højsgaard, S. (2011). On the usage of the geepack, (1), pp 1–4.*
- *Hubbard, A. E., Ahern, J., Fleischer, N. L., Van der Laan, M., Lippman, S. A., Jewell, N., Satariano, W. A., 2010, To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology, 21(4), pp 467–474.*
- *Laird, N. M., & Ware, J. H., 1982, Random-effects models for longitudinal data. Biometrics, pp 963–974.*
- *Liang, K., & Zeger, S. L., 1986, Longitudinal Data Aanalysis Using Generalized Linear Models. Biometrika, 73(1), pp 13–22.*
- *Liang, K.-Y., & Zeger, S. L., 1993, Regression analysis for correlated data. Annual Review of Public Health, 14(1), pp 43–68.*
- *Singer, J. D., & Willett, J. B., (2003). Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence, 1st. Ed., Oxford University Press.Twisk, J. W. R., 2004, Longitudinal data analysis: A comparison between generalized estimating equations and random coefficient analysis. European Journal of Epidemiology, 19(8), pp 769–76.*
- *Zeger, S. L., & Liang, K.-Y., 1986, Longitudinal data analysis for discrete and continuous outcomes,Biometrics, pp 121–130.*
- *Zeger, S. L., Liang, K.-Y., & Albert, P. S., 1988, Models for longitudinal data: a generalized estimating equation approach. Biometrics, pp 1049–1060.*

A Review Paper on Mathematical Modelling and Simulation of Fludized Bed Drying System

Prof. Dayanand N. Deomore[1], Prof. R. B. Yarasu[2]
[1]Asst. Prof. D . Y. Patil College of Engineering and Technology , Kolhapur.
[2]Asst. Professor, Government College of Engineering, Nagpur

A B S T R A C T

This is the study on MATHEMATICAL MODELLING AND SIMULATION OF ELECTRON FIRED FLUDIZED BED DRYING SYSTEM, in term of electric arc heating, resistance heating and static electrical energy. FLUDIZED BED DRYING System has been used since long time for drying grains and other types of solid. Drying process is commonly found in agriculture, chemical, mineral, ceramic industries etc. To remove moisture content from materialsfluidized drying system is commonly used. This motivation for FLUDIZED BED was driven by earlier studies that,it consumes a large portion of overall energy consumption and its operating cost is also large.In this study both experimental and theoretical approaches are used in simulation of electron and mathematical modelling. In this study introduced new dimension for mathematical formulation to represent the work of electrons. The inlet air temperature affects on the drying rate of the grains. By this study decrease per unit energy cost as well as drying rate of grains will increase.

Keywords: : Fluidised, Electron fire, optimisation, modelling.

1. INTRODUCTION

In many industries like pulp and paper industry, textile industry etc, drying of material and different type of solids has been done. It consumes a large amount of total energy used in different type of processes. For a single dryer the cost of energy is more than half of overall cost. In this process latent heat of vaporization requires to evaporate moisture from materials. If we consider the cost per unit, it is also important to keep eyes on the rate of energy consumption in any drying process. After these earlier studies and investigation drying is still very complex process to be understood. In this study, it is difficult to get accurate mathematical description of the phenomenon. Study with the help of mathematical modelling and numerical simulation is very helpful in solving drying physics. Further drying technology should be cost effective and efficient for this introduced new method, strategies and optimization of system as well.

2. MOTIVATION AND PROBLEM STATEMENT

Research on fluidized bed dryers has focused primarily on the use of pressure fluctuation analysis, namely the S-statistic, as a tool for monitoring the drying process.

However, limited attention has been given to understanding the hydrodynamics associated with fluidized bed dryers, namely the influence of moisture on the fluidization behaviour. To complement basic understanding, as well as to get a better overall understanding of fluidized bed drying hydrodynamics, design features of fluidized beds are also important to consider. Other than operating parameters such as bed loading and superficial gas velocity, distributor design as well as vessel geometry have an impact on the hydrodynamics of drying. With this wide variety of hydrodynamic studies regarding fluidized bed drying, monitoring of drying processes using the S- statistic, or any other techniques, will become more meaningful.

3. SURVEY OF EARLIER WORK

The history of fluidization began in 1922 with Winkler's patent of a fluidized bed process for coal gasification which was used for the production of synthesis gas Scattered study to early observations of what is known today is fluidization can be found in published literature as far back on 1878.

In 1940s the fluidized solid process was commercialized on a massive scale in the petroleum industry to effect intimate contact between the catalyst and hot vapours in the cracking of heavy hydrocarbons to fuel oil.

The concept of fluidized solids actually arose in the field of catalytic cracking process. The standard oil development company has a lot of contribution in the field of fluidization then Prof. W.K Lewis and E. R Gillil and carried out independent research on flow properties of powdered solids suspended in gases and developed the concept of fluid bed. The first commercial plant using the fluid solids technique principle was put into operation in the year 1930 for no pharmaceutical application but the process was first used for pharmaceutical application in the year 1960 by Wurster . The coating of tablets by spraying the coating solution into a bed of tablets suspended in a stream of warm air was invented by Dale Wurster whose first patent for the method was filed in 1953. Granulation of powder in a fluidized bed carried out in 1960 by Wurster. Then the 1980s have seen an explosion in the research, application and commercialization of fluid bed process.

In 1996 'Vanecek'worked in design procedure for fluidized bed dryers:
They proposed a design procedure for fluidised bed dryer based on heat and mass balance of the whole apparatus. Actually the fluidized bed has been treated as a homogenous system, without considering bubbles, and its nitration with bed .

In 1980 'Hoebink and Rietema' development a model for fluidized bed drying: They developed a model for fluidised bed drying and described heat and mass transfer between the dense phase, the cloud, and the bubbles with the assumption of no diffusion limitation inside the solids. They described the moisture transport between particles and gas with single and constant and overall mass transfer coefficient which is useful for constant rate drying. This was first model and second model, Hoebink and Rietema , extended the first model to include a more general case where due to a several diffusion limitation of the particles, the concept of a constant mass transfer coefficient was no longer adequate to describe the moisture transport of solids.

Olazer observed five flow regimes: fixed bed, partially fluidized bed, spouted bed, transition regime from spouted bed to jet spouted bed and jet spouted bed. The hydrodynamic characteristics for these five regimes are presented by empirical correlations.

The difficulty in predicting the hydrodynamics of fluidization from the theory of Pengand Fan possibly lies in the fact that the inlet diameter for the conical spouted bed is less than that of the bed bottom. The same approach by Peng and Fan has beenadopted by Gelperin and Nishi for the incipient fluidization of gas–solidconical beds. The model developed by Shiis based on Ergun's equation and neglects friction between the particles and the wall. Biswal developed theoretical models, for minimum fluidization velocity and pressure dropin a packed bed of spherical particles for gas–solid systems in conical vessels. Due to the angled walls, random and unrestricted particle movement occursin a tapered bed with reduced back mixing.

Olazer compared their experimental results with that calculated using the models developed by Gelperin and Gorshtein and Mukhlenov for maximum pressure drop and found that the predictions were not very accurate. They therefore proposed a modified equation for calculation of maximum pressure drop.

Later, Peng and Fan made an in-depth study of the hydrodynamic characteristics of solid–liquid fluidization in a tapered bed and derived theoretical models for the prediction of minimum fluidization velocity and maximum pressure drop, based on the dynamic balance offorces exerted on the particle. The experiments were however carried out for spherical particles only.

Jing and Shan developed models forgas–solid conical fluidized beds for spherical coarse and fine particles based on the Peng and Fan models but neglected the pressure drop due to the kinetic change in the bed.

Depypere have carried out studies in a tapered fluidized bed reactor and proposed empirical models for determination of expanded bed height by using static pressure and wall surface temperature measurements.

Levey 1960 have successfully used tapered beds in chemical reactions.

Peng and Fan [1997] have mentioned that the beds could be used for biochemical reactions and roasting of sulfide ores.

Kumar [1981] and Yogesh Chandra and Jagannath Rao [1981] have investigated the hydro dynamics of gas solid fluidization in tapered vessels using single size particles. Tapered fluidized beds have many attractive features, among which are their capabilities for handling particles with different sizes and properties

(Scott and Hancher, 1976; Ishii 1977) and for achieving extensive particle mixing (Babu 1973), These beds have been widely applied in various processes including biological treatment of wastewater, immobilized bio film reaction, incineration of waste materials, coating nuclear fuel particles, crystallization, coal gasification and liquefaction, and roasting sulfide ores.

Interestingly, industrial fluidized beds are, more often than not, fabricated with tapered sections at the bottom. Nevertheless, fundamental understanding of the behavior of tapered fluidized beds appears to lag far behind their applications. Some of the previous investigations include studies on pressure drop of fixed and fluidized beds in tapered vessels (Koloini and Farkas, 1973; Biswal 1984), flow regimes, incipient condition of fluidization, voidage distribution and bed expansion (Hsu,1978), and particle mixing (Ridgway, 1965 Maruyama and Sato, (199l).

A fluidized bed is formed when the particles in the bed are in dynamic equilibrium; the fluid drag force and the buoyancy force are exerted in the upward direction against the gravitational force, which pulls the particles downward (Wilhelm and K wauk, 1948; Davidson and Harrison, 1971). This drag force is constant at any position of a columnar bed of uniform particles; however, it decreases in the upward direction in a tapered bed accompanied by the reduction in the superficial velocity of the fluid. Thus, the particles at the lower part of the bed will first be fluidized upon an increase in the flow rate: in contrast, those at the upper part of the bed remainstatic. This phenomenon of partial fluidization is peculiar to the tapered fluidizedbed. Relatively little has appeared in the open literature on hydrodynamic characteristics of tapered fluidized beds; the majority of what has been published deals with flow regimes and incipient condition of fluidization.

Toyohara and Kawamura (1989) have reported the flow regime of partial fluidization in a gas solid tapered bed. Descriptions have been given in Kwauk's monograph (1993) on the change of the flow regime in a gas-solid tapered bed. The incipient condition of fluidization in a tapered bed can be predictedbased on the dynamic balance of forces exerted on the particle bed. This approach was adopted by Gelperin et al. (1960) and Nishi (1979) for gas-solidtapered beds, and Shi et al. (1984) for liquid-solid tapered beds. Nevertheless, noneof these works took into account the phenomenon of partial fluidization in predicting the incipient condition of fluidization and the concomitant, maximum pressure drop.

Fluidization in tapered beds have found wide applicability in many industrial processes such as, waste water treatment (Scott and Hancher,1976), coating of nuclear fuel particles, crystallization, coal gasification and liquification and roasting of sulfide ores (Peng and Fan, 1997), coating of food powder particles (Depypere et al., 2005), etc. Tapered fluidized beds are useful for fluidization of materials with wide particle size distribution and also for exothermic reactions. It can be operated smoothly without any instability i.e. with less pressure fluctuations (Ridgway, 1965) and also for extensive particle mixing(Babu et al. 1973, Maruyama and Sato, 1991). Various techniques including introduction of baffles, operation in multistage unit and imparting vibrations have been advocated from time to time to tackle slugging problem inconventional bed. Introduction of tapered bed instead of a conventional cylindrical one is an alternative technique in gas-solid fluidization to tackle such problem. Better solid fluid mixing and improved quality of fluidization can be achieved in a tapered bed. The gradual decrease in superficial fluid mass velocity due to increase in the cross-sectional area in the upward direction necessitates the use of continuously decreasing size particles for smooth and stable operation of such a fluidizer. Due to angled wall, random and unrestricted particle movement occurs in tapered bed thereby reducing back mixing (Singh et al, 1992).Although some information for liquid- solid system in tapered bed is available, very little work related to gas-solid fluidization in tapered bed is available. In spite ofits advantages and usefulness not much work has been reported in literature for understanding certain important characteristics, especially fluctuation ratio of the bed. Some of the previous investigations include fixed bed pressure drop calculations (Koloini and Farkas, 1973), flow regimes, incipient condition of fluidization, void age distribution and bed expansion calculations (Hsu et al, 1978).

Maruyama and Koyanagi (1993) have proposed analytical methods to predict the bed expansion and pressure drop in tapered fluidizedbed. Depypere et al (2005) have carried out studies in a tapered fluidized bedreactor and proposed a model for expanded bed height by the use of static pressureand wall surface temperature measurement.

'Yashwant Kumar and Seema. A. Belorkar' studied fluidized bed system for drying of Fruits and Vegetables, this paper reviews the various fruits and vegetables dried in the fluidized bed dryer and their retention of sensory attributes in the dried products. This review states that fluidized bed drying indeed has proved better compare to conventional drying (i.e. tray/cabinet drying) in respect to drying time of fruits and vegetables.

'Doyce Tesoro-Martinez, Tomas U. Ganiron Jr and Harold S.' Studied Fluidized Bed System in Solid Waste Management, this has stimulated the development of advanced combustion, gasification, drying and or cooling of solid particles utilizing the fluidize bed technology The ability of fluidized bed to burn a wide variety of fuels, while meeting strict emission-control regulations, makes them an ideal choice for burning such fuels as high- sulphur coal, lignite, peat, oil, sludge, petroleum coke, gas and wastes. The fluidized bed technology is presented to inform the public about the important applications which can be utilized in dealing with our solid waste disposal problems under environmental restrictions and possible integration of this topic in the power plant design course.

'Christopher Tremblay and Dongmei Zhou' studied Efficient drying Parameters for Bed Dryers. The main aim for this present work is to prove that higher efficiencies can be achieved when the allotted drying time is considered as a parameter of the drying cycle. The paper investigates how to theoretically calculate the most efficient drying parameters for wheat based on the ambient conditions and allotted drying time. Drying in the constant and falling rates is discussed through mathematical models developed for each drying period. Drying air temperatures between 290 and 370 Kelvin, and drying air velocities between 0.3 and 5.3 meters per second are explored. The wheat is dried from a moisture content of 0.22 kilograms of water per kilograms of dry basis to a moisture ratio of 0.05. Energy and energy efficiencies are utilized as a determining factor for most efficient drying parameters. The results prove that the dryer is most efficient when the dryer runs at 370 Kelvin and 0.3 meters per second when the allotted drying time is less than 21.7 hours. An allotted drying time between 21.7 and 25.4 hours would require a drying air temperature between 290 and 293 Kelvin, and a drying air velocity of 0.3 meters per second. If the allotted drying time is greater than 25.4 hours, the ambient drying air temperature is most efficient due to no energy input. Results from mathematical models are compared to experimental results and it shows a good correlation with an average percent error of 5.9 percent.

'Oluleye, A. E., Ogungbemi, A .A. and Anyaeche' worked on Design and Fabrication of a low cost Fluidised Bed Reactor, this work therefore is an attempt to develop a simple low cost drying bed that can lead to reduced drying time. The fluidized bed designed and fabricated in this work consists of the drying column, fluidized plate, the inlet and outlet unit, the heating unit and the fan. The evaluation considered the drying time andtemperature in achieving quality. The drying efficiency and the amount of moisture reduced per time andwere investigated using rice and wheat with moisture contents of 23% and 33% respectively. The dryer hasshorter drying times and efficiencies of 89% and 90% for rice and wheat respectively.

'Edward K. Levy, Hugo S. Caram, Zheng Yao, Zhang Wei and Nenad Sarunac' studied Kinetics of coal drying in bubbling fluidized beds, this paper describes laboratory experiments, to determine the kinetics of coal drying in abubbling fluidized bed, and a simple theoretical model of coal drying. The experiments were performed with coal which had been crushed to minus 6 mm and fluidized with air with temperatures ranging up to 66°C and with velocities of 0.9 to 1.7 m/s. In-bed electrical heaters, used to simulate an in-bed tube bundle, provided additional thermal energy for drying. The experiments determined the effects of superficial air velocity, drying temperature and inlet air humidity level on rate of drying. A theoretical model of the drying process was developed in which the air and coal particles are assumed to be at the same temperature and the air–water vapour mixture leaving the bed at the free surface is in equilibrium with the local values of particle moisture. This model is in good agreement with laboratory data, showing that for this application, the drying rates do not depend on fluidized bed bubble behaviour or on particle- gas contact, but are controlled by in-bedheat transfer, flow rate, moisture content and temperature of the feed air, and the equilibrium moisture content of the coal.

'Babita Soren' studied on drying kinetics using fluidized bed dryer, Some physical characteristics of different vegetables were determined as a function of moisture content. Parameters like time of drying, temperature of air, flow velocity of air and material to be dried were varied and the drying rates were determined. The effects of different parameters like temperature, time and velocity on drying kinetics were determined. Critical moisture content was also determined. The drying kinetics of vegetables was thus obtained from these data by considering the different experimental conditions. Effect of diffusivity also observed.

'Okoronkwoc.a, nwufo, o.c, nwaigwek.n, ogueken.v, and anyanwue.e' have done Experimental Evaluation of A Fluidized Bed Dryer Performance, the drying experiments were carried out according to the following parameters: initial moisture content of the material (High and Low), drying time and various optimum temperatures. The results obtained show that it took a total of 150minutes with an optimum temperature of 60oC to reduce the moisture content of the cassava and yam from 75.4%(w.b) to the equilibrium moisture content of 11% which is suitable for storage and preservation while maintaining the external conditions. While the optimum temperature that gives faster drying time for maize was 400C. Similarly, a preliminary sun drying experiment was carried out to ascertain the drying time of the cassava and yam chips. It was observed that in the case of sun drying of cassava, it takes a total of 72hrs with an average ambient temperature of 30oC to reduce the moisture content from 75.4 % (w.b) to the equilibrium moisture content of 11 % (w.b). Similarly, the effect of temperature at 300C on the drying curves for the maize shows that the value of the constant rate period of the experiment was smaller than the value obtained at the temperatures of 350C and 400C respectively. The drying rate of products below their optimal temperature was affected by the lower rate of moisture removal and the equilibrium moisture content by mass was high. While drying at temperatures above the optimum temperature, ultimately may cause the thermal degradation of the products been dried. This could be in form of physical defects, such as, decoration, cracking, shrinking and non-uniform drying. From the drying kinetic curves and visual observations during the experiments, it could be concluded that the fluidized bed dryer is an alternative for the processing of cohesive solids that preserve the final quality of the dry solids.

'A. Karbassi, Z. Mehdizadeh' studied Drying Rough Rice in a Fluidized Bed Dryer. The objective of this study was to examine the effects of the fluidized- bed drying method on the final quality of two varieties of Iranian rice, medium- and long grain. The results were compared to that of paddy drying using a traditional method. Rough rice was treated in the fluidized bed drier at 140ºC for 2 minutes. Similar samples were dried for8-10 hrs by the traditional method. Dried samples were dehisced and polished. Quality factors, including trade quality (head rice yield percent and whiteness), cooking quality(amylose content, gelatinization temperature, gel consistency, aroma and flavor) and nutritional quality (thiamine and lysine contents), were then measured for each sample. Finally, the data was analyzed. Results show that paddy drying in a fluidized bed dryer would reduce the quality factors except for rice whiteness for which conventional drying is more acceptable. Therefore modification of fluidized-bed drying technique is recommended.

'Zeljko B. Grbavcic, ZoranaLj. Arsenijevic and Radmila V. Garic-Grulovic' studied Drying of suspension and pastes in fluidized bed of inert particles. A fluid bed dryer with inert particles was used for the drying of suspensions and pastes. The effects of the operating conditions on the dryer throughput and on the product quality were investigated. Experiments were performed in a cylindrical column 215mm in diameter and 1200 mm in height with 0.925 mm diameter glass spheres as the fluidizing media. Cineb fungicide, copper hydroxide and pure water were used as the feed material. With respect to the main efficiency criteria, i.e., specific water evaporation rate, specific heat consumption and specific air consumption, a fluid bed dryer with inert particles represents a very attractive alternative to other drying technologies. A high drying efficiency results from the large contact area and from the large temperature difference between the inlet and outlet air. A rapid mixing of the particles, due toaggregative fluidization and mechanical agitation, leads to nearly isothermal conditionsthroughout the bed. In our experiments, suspensions and very dense pastes were successfullytreated. Suspension and product hold- up in the bed varies between 6 and 8 %by mass and a product with the same particle size as the raw material is obtained.

Jaehyeon Park_ Dowon Shun, Dal-HeeBae, Sihyun Lee, Jeong Hak SeoyJae Hyeok Parkz studied the effect of gas temperature and velocity on coal drying in fluidised bed dryer. The objective of this research work is to develop fluidized bed coal dryer to overcome the disadvantages of low rank coal with high moisture such as low calorific values, costly transportation, high emissions of pollutants, and operational problem. In this paper, laboratory scale bubbling fluidized bed was used to dry high moisture, low-rank Indonesian coal to produce low moisture, high-rank coal. The effects of temperature, gas velocity and bed height to diameter ratio (L/D) on drying rate were studied to obtain information relating to optimum operating conditions.

Coal characterizations (proximate analysis, ultimate analysis, Thermogravi metric Analysis (TGA), BET, Higher Heating Value (HHV), Lower Heating Value (LHV)) were performed to identify the effect of the change of moisture content. This investigation aims to study the drying process under moderated heating conditions.

As a result of the experiments the conclusion is that the thermal fluidized bed process can be successfully applied to reducing moisture in Indonesian coal. Results also indicate that about 80~90% of total moisture could be reduced, including some of the inherent moisture, yielding high heating value product. The drying rate of coal in a fluidized bed is increased by increasing the temperature and velocity of the drying gas. However gas temperature had limitations causing from the spontaneous combustion and gas velocity has to be decided considering energy efficiency.

'Yashwant Kumar1, Mohammad Ali Khan and Krishna Kumar Patel' studied Effect of Microwave on Fluidized Bed Drying of Beetroot. In the present work, an attempt has been made to study the effect of inlet air temperature and velocity on the drying characteristics of beetroot's (Beta vulgaris L.) pieces in microwave assisted fluidized bed drying (MAFBD) system. The results were compared with samples of beetroot dried in a fluidized bed dryer (FBD) at the same combination of temperatures and air velocities. The selected inlet air temperatures and inlet air velocities were 60°C, 67.50°C and 75°C and 9 m/s, 10.50 m/s and 12 m/s, respectively. Moisture content and outlet air humidity was measured at 5 minutes interval. The MAFBD method offered two to three times reduction in drying time as compared to the FBD method. It was also observed that the beetroot samples obtained from the MAFBD system had lower final moisture content than those obtained from the FBD system.

SCOPE OF WORK:

After studying different thesis proposed by different Authors as well as my further studies on Fluidized Bed System, I conclude that we can compare models and experimental results for wheat and corns particles, effect of hydrodynamic modelling on drying characteristic, method for increasing thermal efficiency of fluidized bed, recycling of exhaust gas, intermittency and easily understand the gas flow pattern in fluidized bed, conditioning of grains, moisture content determination of grains, determine the properties of wheat and corns, understand the procedure of drying run of fluidized bed, able to understand the principles of electric arc heating, resistance heating, static electric energy, parting arc, electric fire bed system and drying.

REFERENCES

- *Yashwant Kumar and Seema. A. Belorkar' "Study of fluidized bed system for drying of Fruits and Vegetables" Volume 01, No. 9, September 2015*
- *Doyce Tesoro-Martinez, Tomas U. GanironJr and Harold S.' "Study Fluidized Bed System in Solid Waste Management", International Journal of u- and e- Service, Science and Technology Vol.7, No.1 (2014)*
- *Christopher Tremblay and Dongmei Zhou' "study Efficient drying Parameters for Bed Dryers"*
- *Oluleye, A. E., Ogungbemi, A .A. and Anyaeche' "Work on Design and Fabrication of a low cost Fluidised Bed Reactor" Vol 3, No 3, 2012*
- *'Edward K. Levy, Hugo S. Caram, Zheng Yao, Zhang Wei and NenadSarunac' "Study Kinetics of coal drying in bubbling fluidized beds" Proceedings Fifth World Congress on Particle Technology, Orlando, Florida, April 2006.*
- *BabitaSoren' "Study on drying kinetics using fluidized bed dryer"*
- *OKORONKWO C.A, NWUFO, O.C, NWAIGWE K.N, OGUEKE N.V, AND ANYANWU E.E' "Done Experimental Evaluation of A Fluidized Bed Dryer Performance" The International Journal Of Engineering And Science (IJES) ||Volume||2 ||Issue|| 6 ||Pages|| 45-53||2013|| ISSN(e): 2319 – 1813 ISSN(p): 2319 – 1805*
- *A. Karbassi, Z. Mehdizadeh' "Study Drying Rough Rice in a Fluidized Bed Dryer" J. Agric. Sci. Technol. (2008) Vol. 10: 233- 241*
- *Zeljko B. Grbavcic, ZoranaLj. Arsenijevic and Radmila V. Garic-Grulovic' "Studied Drying of suspension and pastes in fluidized bed of inert particles"*
- *Jaehyeon Park_ Dowon Shun, Dal-HeeBae, Sihyun Lee, JeongHakSeoy Jae HyeokParkz "Studied the effect of gas temperature and velocity on coal drying in fluidised bed dryer" Published by ECI Digital Archives, 2010*

Data Mining Technique Applied To DNA Sequencing

Dr. R. Jamuna
Associate Professor,
Department of Computer Science, S. R. College, Bharathidasan

ABSTRACT

*CpG islands (CGIs) play a vital role in genome analysis as genomic markers. Identification of the CpG pair has contributed not only to the prediction of promoters but also to the understanding of the epigenetic causes of cancer. In the human genome [1] wherever the dinucleotides CG occurs the C nucleotide (cytosine) undergoes chemical modifications. There is a relatively high probability of this modification that mutates C into a T. For biologically important reasons the mutation modification process is suppressed in short stretches of the genome, such as 'start' regions. In these regions *2+ predominant CpG dinucleotides are found than elsewhere. Such regions are called CpG islands. DNA methylation is an effective means by which gene expression is silenced. In normal cells, DNA methylation functions to prevent the expression of imprinted and inactive X chromosome genes. In cancerous cells, DNA methylation inactivates tumor-suppressor genes, as well as DNA repair genes, can disrupt cell-cycle regulation. The most current methods for identifying CGIs suffered from various limitations and involved a lot of human interventions. This paper gives an easy searching technique with data mining of Markov Chain in genes. Markov chain model has been applied to study the probability of occurrence of C-G pair in the given gene sequence. Maximum Likelihood estimators for the transition probabilities for each model and analgously for the model has been developed and log odds ratio that is calculated estimates the presence or absence of CpG is lands in the given gene which brings in many facts for the cancer detection in human genome.*

1. INTRODUCTION

CpG islands (CGIs) play a vital role in genome analysis as genomic markers. Identification of the CpG pair has contributed not only to the prediction of promoters but also to the understanding of the epigenetic causes of cancer. In the human genome [1] wherever the dinucleotides C-G occurs the C nucleotide (cytosine) undergoes chemical modifications. There is a relatively high probability of this modification that mutates C into a T. For biologically important reasons the mutation modification process is suppressed in short stretches of the genome, such as 'start' regions. *2+In these regions predominant CpG dinucleotides are found than elsewhere. Such regions are called CpG islands. The most current methods for identifying CG islands suffered from various limitations and involved a lot of human intervention which can be easily searched with Markov Chain based Java coding developed in this paper. Java program has been created to identify the presence or absence of CpG Islands in the given Genome sequence. Therefore, given an annotated training data set, our coding has the capability to find other specific nucleotides sequences in DNA. DNA methylation is an effective means by which gene expression is silenced. In normal cells, DNA methylation functions to prevent the expression of imprinted and inactive X chromosome genes. In cancerous cells, DNA methylation inactivates tumor-suppressor genes, as well as DNA repair genes, can disrupt cell-cycle regulation. DNA methylation in instances such as these has enormous impact in the prevention and treatment of human cancers.

1.1 Data Mining with Markov Chains Model

A Markov chain is a model that generates sequences in which the probability of a symbol depends only on the previous symbol. A Markov chain model *3+ is defined by (a) a set of states, Q, which emit symbols and (b) a set of transitions between states. States are represented by circles and transitions are represented by arrows. Each transition has an associated transition probability, a_{ij} , which represents the conditional probability of going to state j in the next step, given that the current state is i. The sum of all transition probabilities from a given state must equal 1.The row values are summed up to unity in Table[1.1].

1.2 Finite Markov Chain

An integer time stochastic process, consisting of a domain D of m>1 states $\{s_1 \ldots s_m\}$ and

1. Consider a m dimensional initial distribution vector
$(p(s_1) \ldots p(s_m))$.
2. It has n×n transition probabilities matrix $M = (a_{sisj})$

For example, in a domain D can be the letters {A, C, T, G}, p (A) the probability of A to be the 1st letter in a sequence, and aAG the probability that G follows A in a sequence. For each integer n, a Markov Chain [4] assigns probability to sequences $(x_1 \ldots x_n)$ over D (i.e, x_i D) as follows:

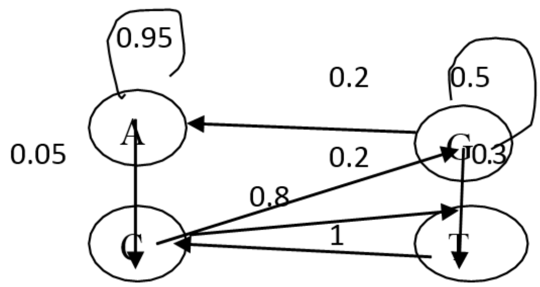
$$p((x_1, x_2, \ldots x_n)) = p(X_1 = x_1)\prod_{i=2}^{n} p(X_i = x_i \mid X_{i-1} = x_{i-1})$$

$$= p(x_1)\prod_{i=2}^{n} a_{x_{i-1} x_i}$$

Similarly, $(X_1, \ldots X_i \ldots)$ is a sequence of probability distributions over D.

There is a rich theory which studies the properties of such “Markov sequences” $(X_1 \ldots X_i \ldots)$. Table[1.1]

	A	G	C	T
A	0.95	0	0.05	0
G	0.2	0.5	0	0.3
C	0	0.2	0	0.8
T	0	0	1	0



Each directed edge say A→G is associated with the positive transition probability from A to G.

1.3 Modeling CpG islands in Genome

In human genomes the pair CG often transforms to (methyl-C) G which often transforms to TG. Hence the pair CG appears less than expected from what is expected from the independent frequencies of C and G alone. [5]Due to biological reasons, this process is sometimes suppressed in short stretches of genomes such as in the start regions of many genes. These areas are called CpG islands (p denotes "pair").

The "-" model: Use transition matrix $A^- = (a^-_{st})$, Where:

a^-_{st} = (the probability that t follows s in a non CpG island)

1.4 Maximum Likelihood estimators for the transition probabilities:-

Maximum Likelihood estimators for the transition probabilities for each model is calculated with sample CpG – islands of a human DNA one estimates the following transition probabilities[6].

The following table shows

- Two Markov chain models: CpG islands (the '+' model)
- Remainder of the sequence (the '-' model).Each row sums to 1.Tables are asymmetric Table[1.2]

Table of Frequencies: Table[1.2]

+	A	C	G	T
A	0.18	0.27	0.43	0.12
C	0.17	0.37	0.27	0.19
G	0.16	0.34	0.38	0.13
T	0.08	0.36	0.38	0.18
-	A	C	G	T
A	0.3	0.21	0.29	0.21
C	0.32	0.3	0.08	0.3
G	0.25	0.25	0.3	0.21
T	0.18	0.24	0.29	0.29

To use the model for discrimination one calculates the log-odds ratio

$$S(x) = \log\left(\frac{P(x|model+)}{P(x|model-)}\right) = \sum_{i=1}^{L} \log \frac{a^+_{x_{i-1}x_i}}{a^-_{x_{i-1}x_i}}$$

$$= \sum_{i=1}^{L} \beta_{x_{i-1}x_i}$$

REFERENCES

1. *Bandyopadhyay, anghamitra. Gene Identification: Classical and Computational Intelligence Approach. 38 vols. IEEE, Jan,2008.*
2. *Durbin, R., S. Eddy, and A. Krogh. Biological Sequence Analysis. Cambridge: Cambridge University, 1998.*
3. *Koski, Timo. Hidden Markov Models for Bioinformatics. Sweden: Kluwer Academic, 2001.*
4. *Birney, E. "Hidden Markov models in biological sequence analysis". July 2001:*
5. *Haussler, David. David Kulp, Martin Reese Frank Eeckman "A Generalized Hidden Markov Model for the Recognition of Human Genes in DNA".*
6. *Boufounos, Petros, Sameh El-Difrawy, Dan Ehrlich. "Hidden Markov Models for DNA Sequencing".*

//IDENTIFICATION OF //CpG IS LANDS IN //SAMPLE GENOME //JAVA CODING

```
import java.io.*;
import java.math.*;
import java.lang.*;
class agct2
{
public static void main(String args[])throws IOException
{
BufferedReader in=new BufferedReader(new InputStreamReader(System.in));
String inp="";
float  cgp[][]={{0.180f,0.274f,0.426f,0.120f},{0.171f,0.368f,0.274f,0.188f},{0.161f,0.339f,0.375f,0.125f},{0.079f,0.355f,0.384f,0.182f}};
floatcgm[][]={{0.300f,0.205f,0.285f,0.210f},{0.322f,0.298f,0.078f,0.302f},{0.248f,0.246f,0.298f,0.208f},{0.177f,0.239f,0.292f,0.292 f}};
float lor=0.0f;
System.out.println("\n\tEnter DNA Sequence:");
inp=in.readLine()+"\0";
int n=inp.length();
int totcnt=n-1;
char[] a=inp.toCharArray();
for(int i=0;i<n;i++)
{
System.out.print("\t"+a[i]);
}
System.out.println();
System.out.println();
System.out.println("\n\t\tCG+ MATRIX");
for(int i=0;i<4;i++)
{
for(int j=0;j<4;j++)
```

```
{
System.out.print("\t"+cgp[i][j]);}
System.out.println();
}
System.out.println();
System.out.println();
System.out.println("\n\t\tCG- MATRIX");
for(int i=0;i<4;i++)
{
for(int j=0;j<4;j++)
{
System.out.print("\t"+cgm[i][j]);}
System.out.println();
}
int c1=0,c2=0,c3=0,c4=0,c5=0,c6=0,c7=0,c8=0,c9=0,c10=0,
c11=0,c12=0,c13=0,c14=0,c15=0,c16=0;
float  p1=0.0f,p2=0.0f,p3=0.0f,p4=0.0f,p5=0.0f,p6=0.0f,p7=0.0f,p8=0.0f,p9=0.0f,p10=0.0f,
p11=0.0f, p12=0.0f,p13=0.0f,p14=0.0f,p15=0.0f,p16=0.0f;
for(int i=0;i<n;i++)
{
if(a[i]=='a' && a[i+1]=='a')
{
c1+=1;
}
}
for(int i=0;i<n;i++)
{
if(a[i]=='a' && a[i+1]=='c')
{ c2+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='a' && a[i+1]=='g')
{
c3+=1;
}}
for(int i=0;i<n;i++)
```

```
{
if(a[i]=='a' && a[i+1]=='t')
{
c4+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='c' && a[i+1]=='a')
{
c5+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='c' && a[i+1]=='c')
{
c6+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='c' && a[i+1]=='g')
{
c7+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='c' && a[i+1]=='t')
{
c8+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='g' && a[i+1]=='a')
{
 c9+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='g' && a[i+1]=='c')
```

```
{
c10+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='g' && a[i+1]=='g')
{
c11+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='g' && a[i+1]=='t')
{
c12+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='t' && a[i+1]=='a')
{
c13+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='t' && a[i+1]=='c')
{
c14+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='t' && a[i+1]=='g')
{
 c15+=1;
}}
for(int i=0;i<n;i++)
{
if(a[i]=='t' && a[i+1]=='t')
{
```

```
c16+=1;
}}
System.out.println("\n\t\tCOUNTS OF INPUT OCCURENCES");
System.out.println("\n\t\t*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*");
System.out.println("\n\t\tAA="+c1);
System.out.println("\n\t\tAC="+c2);
System.out.println("\n\t\tAG="+c3);
System.out.println("\n\t\tAT="+c4);
System.out.println("\n\t\tCA="+c5);
System.out.println("\n\t\tCC="+c6);
System.out.println("\n\t\tCG="+c7);
System.out.println("\n\t\tCT="+c8);
System.out.println("\n\t\tGA="+c9);
System.out.println("\n\t\tGC="+c10);
System.out.println("\n\t\tGG="+c11);
System.out.println("\n\t\tGT="+c12);
System.out.println("\n\t\tTA="+c13);
System.out.println("\n\t\tTC="+c14);
System.out.println("\n\t\tTG="+c15);
System.out.println("\n\t\tTT="+c16);
while(c1>0)
{
p1=p1+(float)Math.log10(cgp[0][0]/cgm[0][0]);
c1--;
}
while(c2>0)
{
p2=p2+(float)Math.log10(cgp[0][1]/cgm[0][1]); c2--;
}
while(c3>0)
{
p3=p3+(float)Math.log10(cgp[0][2]/cgm[0][2]); c3--;
}
while(c4>0)
{
```

```
p4=p4+(float)Math.log10(cgp[0][3]/cgm[0][3]);
c4--;
}
while(c5>0)
{
p5=p5+(float)Math.log10(cgp[1][0]/cgm[1][0]); c5--;
}
while(c6>0)
{
p6=p6+(float)Math.log10(cgp[1][1]/cgm[1][1]); c6--;
}
while(c7>0)
{
p7=p7+(float)Math.log10(cgp[1][2]/cgm[1][2]); c7--;
}
while(c8>0)
{
p8=p8+(float)Math.log10(cgp[1][3]/cgm[1][3]); c8--;
}
while(c9>0)
{

p9=p9+(float)Math.log10(cgp[2][0]/cgm[2][0]); c9--;
}
while(c10>0)
{
p10=p10+(float)Math.log10(cgp[2][1]/cgm[2][1]); c10--;
}
while(c11>0)
{
p11=p11+(float)Math.log10(cgp[2][2]/cgm[2][2]); c11--;
}
while(c12>0)
{
p12=p12+(float)Math.log10(cgp[2][3]/cgm[2][3]); c12--;
}
while(c13>0)
{
p13=p13+(float)Math.log10(cgp[3][0]/cgm[3][0]);
c13--;
```

```
}
while(c14>0)
{
p14=p14+(float)Math.log10(cgp[3][1]/cgm[3][1]); c14--;
}
while(c15>0)
{
p15=p15+(float)Math.log10(cgp[3][2]/cgm[3][2]); c15--;
}
while(c16>0)
{
p16=p16+(float)Math.log10(cgp[3][3]/cgm[3][3]);
 c16--;
}
System.out.println("\n\t\tPROBABILITY OF INPUT OCCURENCES");
System.out.println("\n\t\t*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*");
System.out.format("\n\t\t AA="+"%.2f%n",p1);
System.out.format("\n\t\t AC="+"%.2f%n",p2);
System.out.format("\n\t\t AG="+"%.2f%n",p3);
System.out.format("\n\t\t AT="+"%.2f%n",p4);
System.out.format("\n\t\t CA="+"%.2f%n",p5);
System.out.format("\n\t\t CC="+"%.2f%n",p6);
System.out.format("\n\t\t CG="+"%.2f%n",p7);
System.out.format("\n\t\t CT="+"%.2f%n",p8);
System.out.format("\n\t\t GA="+"%.2f%n",p9);
System.out.format("\n\t\t GC="+"%.2f%n",p10);
System.out.format("\n\t\t GG="+"%.2f%n",p11);
System.out.format("\n\t\t GT="+"%.2f%n",p12);
System.out.format("\n\t\t TA="+"%.2f%n",p13);
System.out.format("\n\t\t TC="+"%.2f%n",p14);
System.out.format("\n\t\t TG="+"%.2f%n",p15);
System.out.format("\n\t\tTT="+"%.2f%n",p16);
lor=p1+p2+p3+p4+p5+p6+p7+p8+p9+p10+p11+p12+p13+p14+p15+p16;
 System.out.format("\n\n\t\t LOG ODDS RATIO:"+"%.2f%n",lor);
if(lor>1)
System.out.println("\n\n\t\tGIVEN SEQUENCE IS A C-G ISLAND");
else
System.out.println("\n\n\t\tGIVEN SEQUENCE IS NOT A C-G ISLAND");
}}
```

Data Logger System: A Survey

Megha Goyal, Parminder Malhotra
Doaba Institute of Engineering and Technology,
Kharar.

A B S T R A C T

This paper describes a brief introduction about data logger system. A data logger is an electronic device that combines analog and digital measurements with programming methodology to sense temperature, relative humidity and other parameters such as voltage and pulse. The data loggers take input from the thermocouple temperature and humidity and other sensors. Knowledge of temperature and relative humidity course during a certain time is needed in scientific, medical and industrial applications. For retrieval of this information from various devices such as manual readings, chart recorders or data loggers can be used.

Keywords: Introduction, History of Data logger system, Operation of Data logger, Characteristics of Data logger and its advantage & applications, future Prospectus, References.

I. INTRODUCTION

A data logger (also data logger or data recorder) is an electronic device that records data over time or in relation to location either with a built in instrument or sensor or via external instruments and sensors. Increasingly, but not entirely, they are based on a digital processor (or computer). They generally are small, battery powered, portable, and equipped with a microprocessor, internal memory for data storage, and sensors. Some data loggers interface with a personal computer and utilize software to activate the data logger and view and analyze the collected data, while others have a local interface device (keypad, LCD) and can be used as a stand-alone device.

One of primary benefits of using data loggers is the ability to automatically collect data on a 24-hour basis.The logging interval can be set for any period between 10 seconds and 24 hours, depending on the data requirements.

A complete data logging application generally requires most of the elements/components illustrated below.

1) Sensors: The inputs from various sources are given to the data logger through various sensors to measure various parameters such as temperature, humidity where electrical signals are converted to temperature and humidity values.

2) User Interface: The interface for interaction with the software and sensors is provided and using implemented algorithm analysis is done for storage of data.

3) Software: It displays the information stored from sensors for and also maintains data for long time storage.

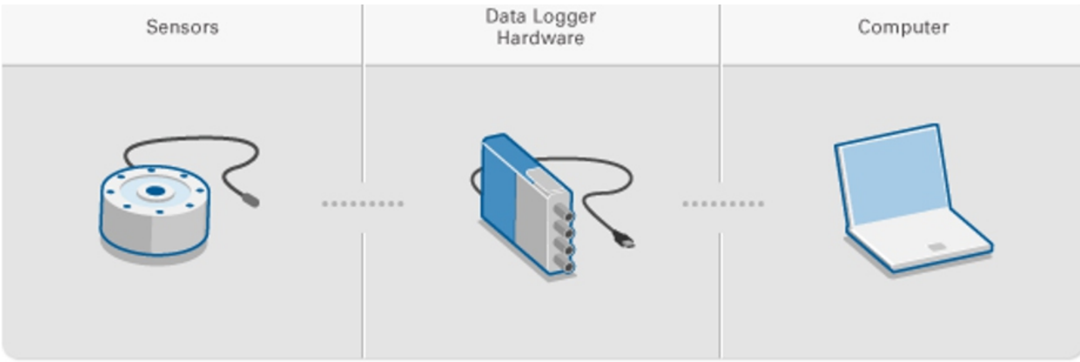


Figure 1. Data Logging System

II. HISTORY OF DATA LOGGER SYSTEM

The earliest form of data logging involved taking manual measurements from analog instruments such as thermometers and manometers. These measurements were recorded into a written log, along with the time of observation. To view trends over time, people manually plotted their measurements on graph paper. In the late 19th century, it became possible to begin automating this process with machines, and strip chart recorders evolved. Strip chart recorders are analog instruments that translate electrical impulses from sensors into mechanical movement of an arm. A pen is attached to the arm, and long rolls of paper are moved at a constant rate under the pen. The result is a paper chart displaying the parameters measured over the course of time. Strip chart recorders were a great leap over manual data logging, but still had drawbacks. For example, translating the traces on the paper into meaningful engineering measurements was tedious at best, and the data recorded took up reams and reams of paper.

With the development of the personal computer in the 1970s and 80s, people began to use computers for analysis of data, data storage, and report generation. The need to bring data into the PC brought about data loggers – a new special-purpose device for data logging. Data loggers are standalone, box instruments that measure signals, convert to digital data, and store the data internally. This data must be transferred to the PC for analysis, permanent storage, and report generation. Data is typically transferred either by manually moving a storage device (such as a floppy disk) from the data logger to the computer or by connecting the data logger to the PC through some communications link such as a serial cable or Ethernet.

In the 1990s, a further evolution in data logging took place as people begin to create PC- based data logging systems. These systems combine the acquisition and storage capabilities of stand-**alone data** loggers with the archiving, analysis, reporting, and display capabilities of modern PCs. PC-based logging systems finally brought about full automation of the data logging process. The move to PC-based data logging systems was enabled by three technological enhancements:

1. Increasing reliability of PCs.
2. Steadily decreasing cost of hard drive space on PCs.
3. PC-based measurement hardware that could meet or exceed measurement capabilities of stand-alone data loggers.

Today, PC-based logging systems provide the widest range of measurement types, analysis capabilities, and reporting tools. The remainder of this paper will focus on the functionality necessary to implement a PC-based data logging system 1- Wire Weather.

III. OPERATION OF DATA LOGGER SYSTEM

The ability to take sensor measurements and store the data for future use is, by definition, a characteristic of a data logger. However, a data logging application rarely requires only data acquisition and storage. Inevitably, the ability to analyze and present the data to determine results and make decisions based on the logged data is needed. A complete data-logging application typically requires most of the elements illustrated below

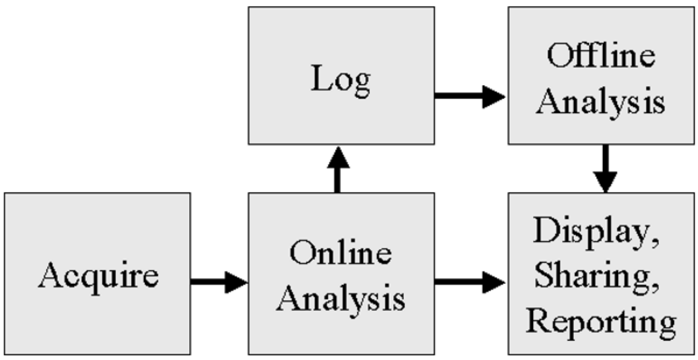


Figure 2. Basic Elements of a Data Logging System

The five major blocks of PC based data logging system are:

1. Data Acquisition
2. Online Analysis
3. Logging & storage
4. Offline Analysis
5. Display, Sharing & Report generation.

1) Acquire – The acquisition is accomplished by the measurement hardware, which can be further broken down into sensors, signal connectivity, signal conditioning, and Analog-to- digital conversion and also as well as conversion the physical phenomenon into digital signals.

2) Online analysis – In PC based systems special software are used to accomplish online analysis. This step includes any analysis that is likely to be done before storing the data. A common example of this is converting the voltage measurement to meaningful scientific units, such as degree celcius. These complex calculations and data compression are completed before logging the data. Every data logging software application should complete this conversion from binary value to voltage and the conversion from voltage to scientific units.

3) Log – The logging or storage block in essential in every data-logging system. Software is of critical importance in PC-based data logging systems, because well-written logging software determines how data is stored, how quickly data can be written to disk, and how efficiently disk space is used. Logging software also gives you data management Capabilities, such as changing data formats, archiving data, and access to databases. There are three general formats commonly used for storage in data logging systems.

- ASCII text files
- Binary files.

4) Offline Analysis - This step includes any analysis that is to be done after storing the data and it is performing mathematical functions on data after it has been acquired in order to extract important information. A common example is looking for trends in historical data or data reduction.

5) Display, share & report generation- This step includes the creation of any reports that are needed to make to present data and displaying the data. However, this can also present data straight from online analysis. The data-logging application requires a display to view the measurements that are being recorded. The display can be of two types,

- Historical Data- Data that has been previously acquired
- Live Data- Data that are currently being acquired

Data viewing utility should provide a user interface with general customization features. Also the data logging application requires some capability for reporting the data. Report generation can be integrated into PC-based data logging applications for increased efficiency. The logging application can be set up to periodically generate specified reports and distribute them. For data that has been logged to be useful, it must be available to theright people. Logging application should have the capability to publish data over a network in order to propagate the information to the concerned people.

In widely distributed data logging applications, each logging node can publish its measurements to the network, and a main computer can serve as the central collection facility. The central computer retrieves the measurements from each node, combines them for further analysis, logs the results for permanent archiving, and periodically generates reports analyzing the data.

IV. CHARACTERISTICS OF DATA LOGGER

1) Modularity: Data loggers can be expanded simply and efficiently whenever required, without any interruption to the working system.

2) Reliability and Ruggedness: They are designed to operate continuously without interruption even in the worst industrial environments.

3) Accuracy: The specified accuracy is maintained throughout the period of use.

4) Management Tool: They provide simple data acquisition, and present the results in handy form.

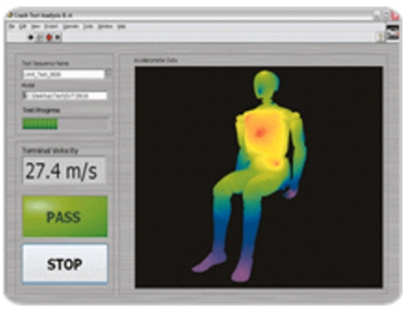
5) Easy to use: These communicate with operators in a logical manner, are simple in concept, and therefore easy to understand, operate and expand.

V. ADVANTAGE AND APPLICATIONS

Advantages:

1. Inline Analysis

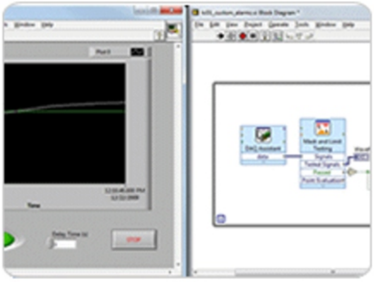
Data analysis with a traditional stand-alone data logger is typically performed offline only after the data has been transferred to the PC. Using a PCbased data logger, you can take advantage of multicore processors and increasingly available RAM in the PC to perform signal processing and analysis on your data as you acquire it. LabVIEW includes many common math and signal processing functions that use configuration wizards and make it easy to add analysis to your measurements.



2. User-Defined Functionality

With a traditional stand-alone data logger, you are generally limited to hardware and software functionality defined by the vendor. These functions are good for accomplishing generalpurpose tasks, but they may not help you meet your unique application requirements. For example, you may want to log data only under certain conditions or generate custom alarms that aren't built into the data logger.

PC-based data loggers are softwaredefined instruments. This means the functionality of the device is defined by the software, and you can customize the software to meet your specific application needs. Using LabVIEW, you can easily build functionality for custom alarms, logging conditions, report generation, and signal analysis. You can log data to virtually any file format for importing into other tools and sharing data with others.

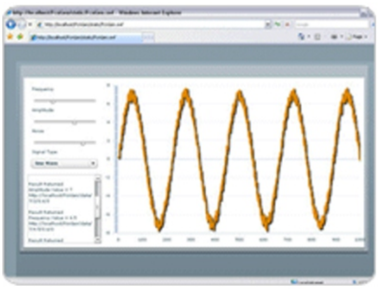


3. Terabytes of Data Storage

Data storage is an important component of a data logger. You can log only as much data as you can store in your data logger. Traditional stand-alone data loggers are limited by the amount of memory built into the device. Because the PC is actually a part of a PC-based data logger, you are limited only by the amount of hard drive space on the PC. Today, it's not uncommon to find a PC hard drive with terabyte capacity that provides ample space for your current measurements as well as permanent storage space.

4. Network Connectivity

For applications that require long-term monitoring over days or weeks, you may have difficulty continually checking results. Remote monitoring is useful because you can see results from a remote location. Using a PC-based data logger, you can take advantage of the PC's network connectivity to transmit results over a network for remote viewing. With LabVIEW, you can create custom alarm conditions that send e-mails or even design a Web service that you can visualize over a Web-based application.



Applications:

1) In unattended recording at weather stations to record parameters like temperature, wind speed / direction, solar radiation and relative humidity.
2) For hydrographic recording of water flow, water pH, water conductivity, water level and water depth.
3) In the recording of soil moisture levels.

4) To record gas pressure and to monitor tank levels.

5) Transportation monitoring, troubleshooting, educational science, quality studies, field studies and general research.

6) Remote collection of recorded data and alarming or unusual parameters are possible with the help of data loggers where these are connected to modems and cellular phones.

VI. FUTURE PROSPECTIVE :

Using data logging, scientists and engineers can evaluate a variety of phenomenon, from weather patterns to factory performance. PC-based data logging systems provide most flexibility, customization, and integration. To define a data logging system, we must evaluate all the requirements for acquisition, analysis, logging, display, and report generation. Based on these requirements, we can customize data logging software and hardware to meet any needs.

VII. REFERENCES:

1. *Suman Saha, Md. Tofiqul Islam and M Zakir Hossain, "Design Of a Low Cost Multi Channel Data Logger," Bangladesh University of Engineering and Technology, ARPN. Journal of Engineering and Applied Science, Vol.1, June 2006.*
2. *Dogan Ibrahim, "Data Logger for Teaching Data Capturing and Analysis to Engineering Students," Department of Computer Engineering, Near East University, Nicosia, Cyprus, Wiley Inter Service DOI 10.1002/cae.20281*
3. *Peter Roberson, "Using data loggers", science teachers' workshop, 2004, north Sydney.*
4. *http/www.loggingtutorial/dr.saulgreenburg/htm.*
5. *DNA technology, "Data logger controller", http://www.dnatechindia.com*
6. *http://www.hanwell.com/index.html*
7. *http ://en.wikipedia.org/wiki/chart-recorder*
8. *www.ni.com/datalogger*
9. *S. J. Perez, M. A. Calva, R. Castaneda, "A microcontroller based data logging System" Instrumentation and Development Vol. 3 Nr. 8, copyright 1997, Journal of the Mexican Society of Instrumentation.*
10. *The university of texas at Arlington,"datalogging & supervisory control in wireless sensor networks",http://dspace.uta.edu/handle/10106/366/u mi-uta-1170.pdf*

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

-

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.