iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** III **Month of publication:** March 2020

DOI:

Generalised Mean and Variance

Dr. J. Suresh Kumar[1], Sarika M. Nair[2],
[1]*Assistant Professor,* [2]*Research Scholar, Post Graduate and Research Department of Mathematics, N.S.S. Hindu College, Changanacherry, Kottayam Dist., Kerala, India-686102*

Abstract: A data consists of several values (called data points) and the aim of Statistical analysis is to explore or describe the data values. In this paper, we investigate the notion of the Generalised Arithmetic mean and the Generalised Geometric mean introduced by Suresh Kumar [1] and will study them in detail for data analysis. We also propose a new notion of the Generalised Covariance to give an approach of Non-linear "Combinatorial" relationship among the data, through the graph models.
Keywords: Data, Graphs, Generalised Arithmetic mean, Generalised Covariance.

I. INTRODUCTION

Statistical data consists of several values (called data points) and the aim of Statistical analysis is to explore or describe the data values and investigates the relationship of the data. The disadvantage of the mean is that it is sensitive to some extreme value, especially when the sample size is small. So, it is not an appropriate measure of central tendency for the skewed distribution. For example, in a company in which a few employees at senior level draw very high salaries plus allowances as compared to the others, the mean will not give any realistic measure of the average income of the employees. In modern new-generation business, there are several layers of employees like Managers, Advisors, Sales Representatives etc. so that there is an inter-relationship (usually Hierarchical) among the employees as well so that the actual income of an employee depends on the amount of business generated by that person, The Generalised Arithmetic mean (GAM) can be used for the data where some data points are more important than some other values so that they shall contribute more to the final "average" and it also gives the inter-relationships among the data values as well. In a University, various courses have to be awarded credit points depending on their relevance and applicability to the core discipline. Professors or teachers handling the curriculum design shall assign credits to the courses, The Generalised Arithmetic mean can be effectively used when calculating a credit for a specific course by seeing the connections among the various courses. The Generalised Arithmetic mean (GAM) is also useful as the atomic mass of a specific element. This refers to the mass and abundance of all the isotopes of an element. An element's atomic mass is the Generalised Arithmetic Mean of the masses of the isotopes of an element. An element's atomic mass can be calculated provided the relative abundances of the element's naturally occurring isotopes and the masses of those isotopes are given. The Generalised Arithmetic Mean is similar to an ordinary arithmetic mean with an exception that instead of each data point contribute equally to the final average, some data points are inter-related to many others and thus contribute more than others. This concept plays a vital role in descriptive statistics and also occurs in a more general form in several other areas of applied mathematics. When the weights of all data points are equal, then the Generalised Arithmetic Mean is the same as the usual arithmetic mean. Covariance is a statistical measure of dispersion used in the Correlation analysis to explore the "Linear" relationship among the data. We propose the notion of the Generalised Covariance to give an approach of Non-linear "Combinatorial" relationship among the data, through graph models.
For statistical terms and notations not explicitly mentioned, reader may refer VK Rohatgi [2]. For graph terms and notations not explicitly mentioned, reader may refer Harary [3].

II. MAIN RESULTS

In this paper, we investigate the notion of the Generalised Arithmetic Mean (GAM) and the Generalised Geometric Mean (GGM) introduced by Suresh Kumar [1] and will investigate them in detail. We extend the investigation in to the concepts such as Variance, Covariance and Correlation coefficient also.

A. Generalised Arithmetic Mean

Arithmetic mean is an average and represents a measure of the central tendency of the data values. Given a set of n elements from a_1 to a_n, the arithmetic mean of these numbers is defined as $(a_1 + a_2 + \cdots + a_n)/n$. and is referred to as AM.
The Generalised Arithmetic mean (GAM) can be used for the data where some data points are more important than some other values so that they shall contribute more to the final "average" and in calculating it, we also consider the inter-relationships among the data values as well.

1) *Definition.* Let $a_1, a_2 \dots\dots a_n$ be a given set of numbers. Consider a graph G with n vertices and assign these numbers as its vertex labels. For any edge $\{a_i, a_j\}$ of the graph G, assign the label $(a_i + a_j)/n$. Then Generalised Arithmetic Mean (GAM) is defined as the sum of all the edge labels of G $and\ is\ d$and is denoted by GAM(G).
2) *Theorem.* For a graph G with vertex degrees, $d_1, d_2 \dots\dots\dots d_n$, Generalised Arithmetic mean is given by $GAM = \sum_{i=1}^{n} a_i d_i/n$
a) *Proof:* For any edge $\{a_i, a_j\}$ assign the label: $(a_i + a_j)/n = a_i/n + a_j/n$. Hence, if we count the sum of all the edge labels of G, then each vertex v_i with degree d_i contribute $d_i a_i/n$ to the sum of the edge labels of all edges of the graph. Hence the Generalised Arithmetic mean is given by $\sum_{i=1}^{n} a_i d_i/n$. The following corollaries are immediate from the above theorem.
3) *Corollary.* For a 1-Regular graph GAM is same as the usual AM of given set of numbers and for a 2-Regular graph GAM will be the twice that of the usual AM of given set of numbers.
4) *Corollary.* For a k-regular graph, the Generalised Arithmetic Mean of a given set of numbers is k times that of the usual Arithmetic Mean. The following corollary gives a useful bound for the Generalised Arithmetic Mean of a given set of numbers.
5) *Corollary.* For a simple connected graph G with n vertices

$$AM \leq GAM \leq (n-1)AM$$

a) *Proof:* Since for a graph G with vertex degrees, $d_1, d_2 \dots\dots\dots d_n$, Generalised Arithmetic mean is given by $\sum_{i=1}^{n} a_i d_i/n$, the first inequality $AM \leq GAM$ follows at once since the usual arithmetic mean is $\sum_{i=1}^{n} a_i/n$. Also the GAM is the maximum when $d_1, d_2 \dots\dots\dots d_n$ have the maximum value. Since the maximum degree of a vertex of a graph with n vertices is (n-1), the second inequality follows. Now we proceed to investigate the Nordhaus-Gaddum type result for the graph parameter, the Generalised Arithmetic Mean. Let G be a graph with vertex degrees, $d_1, d_2 \dots\dots\dots d_n$ so that $\bar{G}$ is the graph with vertex degrees, $n-1-d_1, n-1-d_2 \dots\dots\dots n-1-d_n$. Then the Generalised Arithmetic mean of the numbers $a_1, a_2 \dots\dots\dots a_n$with respect to the graph G satisfies the same lower and upper Nordhaus-Gaddum bounds as below.
6) *Theorem.* For a graph G with n vertices, $GAM(G) + GAM(\bar{G}) = (n-1)\, AM$
a) *Proof:* Let G be a graph with n vertices. $a_1, a_2 \dots\dots a_n$ be its vertex labels. Any edge $\{v_i, v_j\}$ of G gets the label $(a_i + a_j)/n$. Thus when we calculate the Generalised Arithmetic mean with respect to G, each vertex v_i contribute a_i/n to each edge to which it is incident with. Thus each vertex v_i contribute $a_i/n\, , for\ all\ i = 1,2, \dots (n-1)$ to $GAM(G) + GAM(\bar{G})$. That is, each vertex v_i contributes $(n-1)\, a_i/n$ since $E(G) \cup E(\bar{G}) = E(K_n)$. Hence the Theorem follows.

B. Generalised Geometric Mean

In this section, we recollect the definition of the Generalised Geometric mean (GGM) of a graph as in Suresh Kumar [1] and study it in detail.

1) *Definition.* Let $a_1, a_2, \dots\dots\dots a_n$ be a given set of numbers. Consider a graph G with n vertices and assign these numbers as its vertex labels. For any edge $\{a_i, a_j\}$ of G, assign the label $\sqrt[n]{a_i a_j}$. Then Generalised Geometric Mean (GGM) is defined as the product of all edge labels of G.
2) *Theorem.* For a graph G with vertex degrees, $d_1, d_2 \dots\dots\dots d_n$, Generalised Geometric mean is given by $({a_1}^{d_1} {a_2}^{d_2} \dots\dots\dots {a_n}^{d_n})^{1/n}$
a) *Proof:* For any edge $\{a_i, a_j\}$ assign the label: $(a_i a_j)^{1/n} = {a_i}^{1/n} {a_j}^{1/n}$. Hence, if we compute the product of all the edge labels of G, then each vertex v_i with degree d_i contribute $({a_i}^{1/n})^{d_i}$ to the product of the edge labels of all edges of the graph. Hence the Generalised Geometric mean is given by $({a_1}^{d_1} {a_2}^{d_2} \dots\dots\dots {a_n}^{d_n})^{1/n}$

   The following corollaries are immediate from the above theorem.
3) *Corollary.* For a 1-Regular graph GGM is same as the usual GM of given set of numbers and for a 2-Regular graph GGM will be square of the usual GM of given set of numbers.
4) *Corollary.* For a k-regular graph, $GGM = (GM)^k$ The following corollary gives a useful bound for the Generalised Arithmetic Mean of a given set of numbers.
5) *Corollary.* For a simple connected graph G with n vertices, $GM \leq GAM \leq (GM)^{n-1}$
a) *Proof.* Since for a graph G with vertex degrees, $d_1, d_2 \dots\dots\dots d_n$, Generalised Geometric mean is given by $\prod_{i=1}^{n} ({a_i}^{1/n})^{d_i}$, the first inequality $GM \leq GGM$ follows at once since the usual Geometric mean is$\prod_{i=1}^{n} {a_i}^{1/n}$. Also the GGM is the maximum when $d_1, d_2 \dots\dots\dots d_n$ have the maximum value. Since the maximum degree of a vertex of a graph with n vertices is (n-1), the second inequality follows. Now we proceed to investigate the Nordhaus-Gaddum type result for the graph parameter, the Generalised

Geometric Mean. Let G be a graph with vertex degrees, $d_1, d_2 \dots \dots . . d_n$ so that $\bar{G}$ is the graph with vertex degrees, $n-1-d_1, n-1-d_2 \dots \dots . . n-1-d_n$. Then the Generalised Geometric mean of the numbers $a_1, a_2 \dots \dots . a_n$ with respect to the graph G satisfies the the same lower and upper Nordhaus-Gaddum type bounds as below.

6) *Theorem.* For a graph G with n vertices, $GGM(G).GGM(\bar{G}) = (GM)^{n-1}$.

a) *Proof:* Let G be a graph with n vertices. $a_1, a_2 \dots \dots . a_n$ be its vertex labels. Any edge $\{v_i, v_j\}$ of G gets the label $(a_i a_j)^{1/n}$. Thus when we calculate the Generalised Geometric mean with respect to G, each vertex v_i contribute ${a_i}^{1/n}$ to each edge to which it is incident with. Thus each vertex v_i contribute ${a_i}^{1/n}, for\ all\ i = 1,2, \dots (n-1)$ to $GGM(G) + GGM(\bar{G})$. That is, each vertex v_i contributes $\left({a_i}^{1/n}\right)^{n-1}$ since $E(G) \cup E(\bar{G}) = E(K_n)$. Hence the Theorem follows.

C. Generalised Variance

Variance is defined as the arithmetic mean of the squares of the deviations of the observations from their arithmetic mean. The Square root of the variance is called the standard deviation.

1) *Definition.* Let $a_1, a_{2,} \dots \dots \dots a_n$ be a given set of numbers with mean μ. Let G be any graph with n vertices $v_1, v_2 \dots \dots v_n$ and assign the vertex label $f(v_i) = (a_i - \mu)^2$ to each vertex v_i of G. For any edge $\{v_i,v_j\}$ of G assign the label $\left(f(v_i) + f(v_j)\right)/n$. Then Generalised Variance (GV) is defined as the sum of all edge labels of G. Square root of the Generalised variance is called the Generalised Standard Deviation (GSD).

2) *Theorem.* For a graph G with vertex degrees, $d_1, d_2 \dots \dots . . d_n$, Generalised variance is given by $\frac{1}{n}\sum_{i=1}^{n} d_i f(v_i)$, where $f(v_i) = (a_i - \mu)^2$

a) *Proof:* For any edge $\{a_i, a_j\}$ assign the label: $\left(f(v_i) + f(v_j)\right)/n$. Hence, if we count the sum of all the edge labels of G, then each vertex v_i with degree d_i contribute $d_i\left(f(v_i) + f(v_j)\right)/n$ to the sum of the edge labels of all edges of the graph. Hence the Generalised Arithmetic mean is given by $\frac{1}{n}\sum_{i=1}^{n} d_i f(v_i)$, where $f(v_i) = (a_i - \mu)^2$.

The following corollaries are immediate from the above theorem.

3) *Corollary.* For a 1-Regular graph Generalised variance is same as the usual Variance of given set of numbers and for a 2-Regular graph Generalised variance will be twice of the usual variance of given set of numbers.

4) *Corollary.* for a 1-Regular graph Generalised Standard deviation (GSD) is same as the usual Standard deviation of given set of numbers. For a 2- regular graph GSD is same as $\sqrt{2}SD$.

5) *Corollary.* For a k-regular graph, $GV = k(Variance)$ and $GSD = \sqrt{k}\, SD$

D. Generalised Covariance

Covariance is a statistical measure of dispersion used in the Correlation analysis to explore the “Linear” relationship among the data. Generalised Covariance (GCOV) gives an approach to the Non-linear “Combinatorial” relationship among the data, through graph models. In this section, we propose the notion of the Generalised Covariance to give an approach of Non-linear “Combinatorial” relationship among the data, through graph models.

1) *Definition.* Let $\{(x_i, y_i)\}_{i=1,2,\dots n}$ be a given data. Let G be a graph with n vertices. Assign (x_i, y_i) to each vertex v_i of G. Let a be the mean of $x_i's$ and b be the mean of y_i's. To each edge $\{v_i, v_j\}$ of G, assign the label $(\bar{v_i} + \bar{v_j})/n$ where $\bar{v_i} = (x_i - a)(y_i - b)$ and $\bar{v_j} = (x_j - a)(y_j - b)$. Then the Generalised Covariance is defined as the sum of all the edge labels of G.

2) *Theorem.* For a graph G with vertex degrees, $d_1, d_2 \dots \dots . . d_n$, Generalised Covariance is given by $\frac{1}{n}\sum_{i=1}^{n} d_i \bar{v_i}$, where $\bar{v_i} = (x_i - a)(y_i - b)$

a) *Proof:* For any edge $\{a_i, a_j\}$ assign the label: $(\bar{v_i} + \bar{v_j})/n$ where $\bar{v_i} = (x_i - a)(y_i - b)$ and $\bar{v_j} = (x_j - a)(y_j - b)$.Hence, if we count the sum of all the edge labels of G, then each vertex v_i with degree d_i contribute $d_i\bar{v_i}/n$ to the sum of the edge labels of all edges of the graph. Hence the Generalised Covariance is given by $\sum_{i=1}^{n} \bar{v_i} d_i/n$.

The following corollaries are immediate from the above theorem.

3) *Corollary.* For a 1-regular graph Generalised Covariance is same as the usual covariance and for a 2- regular graph generalised covariance is twice of usual covariance.

4) *Corollary.* For a k-regular graph $GCOV = k(COV)$

The following corollary gives a useful bound for the Generalised Arithmetic Mean of a given set of numbers.

5) *Corollary*. For a simple connected graph G with n vertices, $COV \leq GCOV \leq (n-1)COV$. Now we proceed to investigate the Nordhaus-Gaddum type result for the graph parameter, the Generalised Covariance. Let G be a graph with vertex degrees, $d_1, d_2 \dots \dots . . d_n$ so that $\bar{G}$ is the graph with vertex degrees, $n-1-d_1, n-1-d_2 \dots \dots . . n-1-d_n$. Then the Generalised Covariance of the numbers $a_1, a_2 \dots \dots . a_n$ with respect to the graph G satisfies the the same lower and upper Nordhaus-Gaddum bounds as below.
6) *Theorem*. For a graph G with n vertices, $GCOV(G) + GCOV(\bar{G}) = (n-1)COV$

a) *Proof:* Let G be a graph with n vertices. $a_1, a_2 \dots \dots . a_n$ be its vertex labels. Any edge $\{v_i, v_j\}$ of G gets the label $(\bar{v_i} + \bar{v_j})/n$ where $\bar{v_i} = (x_i - a)(y_i - b)$ and $\bar{v_j} = (x_j - a)(y_j - b)$. Thus when we calculate the Generalised Covariance with respect to G, each vertex v_i contribute $\bar{v_i}/n$ to each edge to which it is incident with. Thus each vertex v_i contribute $\bar{v_i}/n, for\ all\ i = 1,2, \dots (n-1)$ to $GAM(G) + GAM(\bar{G})$. That is, each vertex v_i contributes $(n-1)\bar{v_i}/n$ since $E(G) \cup E(\bar{G}) = E(K_n)$. Hence the Theorem follows.

E. Generalised Correlation Coefficient

The Correlation coefficient is the specific measure that quantifies the strength of the linear relationship between two variables in a correlation analysis. The correlation coefficient is defined by

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

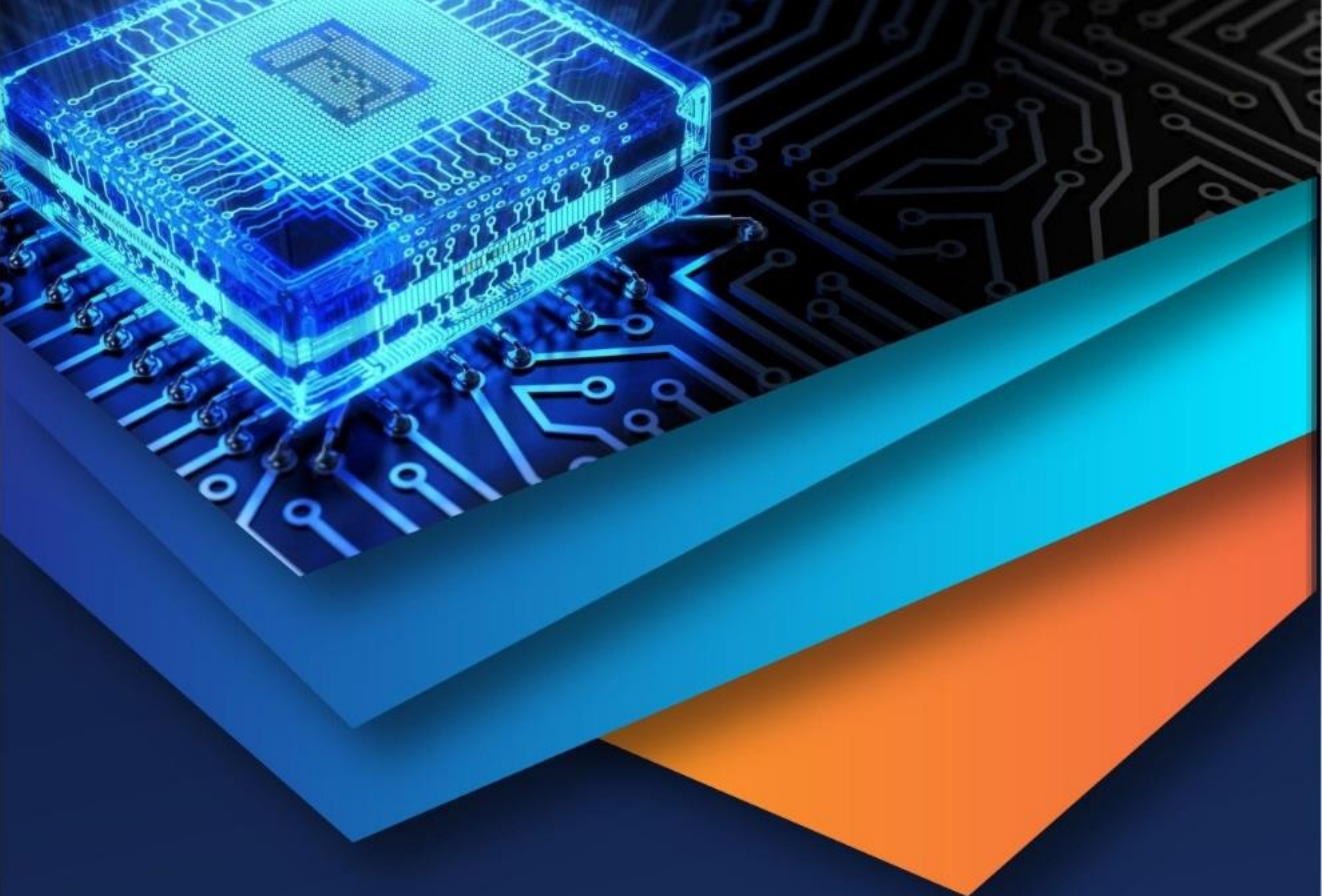
For a graph G with vertex labels (x_i, y_i) we can define Generalised correlation coefficient in terms of Generalised covariance and Generalised standard deviation.

1) *Definition*. Let (x_i, y_i) be a given data. Let G be a graph with n vertices. Assign (x_i, y_i) to each vertex v_i of G. Let $X = \{x_i\}_{i=1,2,\dots n}$ and $Y = \{y_i\}_{i=1,2,\dots n}$. Then Generalised Correlation coefficient is defined as: $\frac{Generalsed\ Covariance\ of\ G}{(GSD\ of\ X)(GSD\ of\ Y)}$
2) *Theorem*. For all regular graphs with n vertices, the Generalised Correlation coefficient is same as the usual Correlation coefficient.

a) *Proof:* Let G be a k-regular graph with n vertices and assign (x_i, y_i) to each vertex v_i of G for some $k < n$. Since G is k-regular, $GCOV(G) = k(COV)$ by corollary 2.4.5. Also, $GSD(X) = \sqrt{k}SD(X)$ and $GSD(Y) = \sqrt{k}SD(Y)$ by corollary 2.3.5. Hence, the Generalised Correlation coefficient is given by $k(\ Covariance\ of\ G)/(\sqrt{k}\ SD\ of\ X)(\sqrt{k}\ SD\ of\ Y) = Covariance\ of\ G/(SD\ of\ X)(SD\ of\ Y)$.

REFERENCES

[1] J. Suresh Kumar, Graph Algorithms and Abstraction for Number Theoretic Concepts, International Journal of Mathematics and its Applications (IJMAA), 7(2) (2019), 39-41.
[2] V. K. Rohatgi, A. K. Md. Ehsanes Saleh, Introduction to Probability and Statistics, John Wiley and Sons, 2011
[3] Harary Frank, Graph Theory, Reading Mass, 1969.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY