Council for Innovative Research
www.cirworld.com

International Journal of Computers & Technology
Volume 4 No. 2, March-April, 2013, ISSN 2277-3061

# A Statistical Analysis of Bhairav-The first morning Raga

Swarima Tewari and Soubhik Chakraborty*
Department of Applied Mathematics, Birla Institute of Technology Mesra, Ranchi-835215, India
Email swarimatewari@gmail.com : soubhikc@yahoo.co.in

## ABSTRACT

A raga, in Indian classical music, is a melodic structure with fixed notes and a set of rules characterizing a particular mood conveyed by performance. *Bhairav* is the first raga of the morning. The present paper gives a statistical analysis of this raga structure.

**Key words:** Raga, melody, entropy, Simple Exponential Smoothing, statistical analysis

## 1. INTRODUCTION

Music, according to Swami Vivekanada, is the highest form of art and also the highest form of worship (provided you understand it!). Understanding music, both aesthetically and scientifically, becomes important. This is especially true for classical music, be it Indian or Western, since each is a discipline in its own right. While the former stresses on the emotional richness of the raga as expressed through melody and rhythm, the latter is technically stronger as, in addition to melody and rhythm, the focus is also on harmony and counterpoint. A raga is a melodic structure with fixed notes and a set of rules characterizing a particular mood conveyed by performance. *Bhairav* is the first raga of the morning. The present paper gives a statistical structure analysis of this raga. Readers interested in performance analysis are referred to [1] where another morning raga, namely, *Ahir Bhairav*, has been analyzed. Those who are interested in some general musical features of *Bhairav* are referred to the appendix.

Our analysis attempts to answer the following questions:-

(a) Can we find a working statistical model that can capture the essence of the *Bhairav* raga structure?

This question is important because the true model is both complex and unknown and that one of the strengths of statistics lies in modeling. Although statistical models are subjective and biased, we can at least make the data objective as far as possible. Also, behind these models stand some beautiful mathematical theorems and they are unbiased [2]. Moreover, the true model may contain multiple parameters related to music theory, the training and background of the artist, the genre of music and even the place of performance and the audience etc. and we do not have an explicit idea as to how exactly (in what functional form) these parameters enter the model. Statistical models are, in contrast, approximate models that use fewer parameters to capture the phenomenon generated by these complex unknown true models. Although approximate, it is possible to verify the goodness of fit of these models as well as control the errors in them.

In the light of these arguments, modeling a musical structure or a musical performance has been a coveted research area in computational musicology.

There are three fundamental steps in statistical modeling: deciding which model to fit, estimate the parameters of the chosen model and verify the goodness of fit of this model. We all know that statistics can be broadly divided into two categories: descriptive and inferential. In statistical modeling, both are involved--as we first describe a pattern (through modeling) and then infer about its validity. Two types of models are used in statistics: probability models and stochastic models. Through a probability model, we can tell the probability of a note or a note combination but cannot predict the next note. Through a stochastic model we can predict (make an intelligent guess of) the next note, given the previous.

In the present paper, a Simple Exponential Smoothing is used to capture the note progression depicting the structure of the raga *Bhairav*. This is a stochastic model used in time series analysis

(b) What are the probabilities of the notes used in the raga? Which note has the highest probability and which the lowest? Interestingly, in music analysis, even a note with low probability cannot be neglected as its surprise element will be more! Entropy is the measure of surprise element in any message. We can treat the realization of a note as message that can be subjected to entropy analysis.

It should be emphasized here that entropy is measuring surprise which should not be confused with meaning. If I write stuoqqzf it is meaningless but since q is coming after q which never happens in English, there is definitely the surprise element. For further literature on entropy, see [3] and the references cited therein. The use of entropy in music analysis has been successfully tried in Western music. We are motivated by the work of Snyder [4].

(c) What are the melody groups and which melody groups are how much significant?

A melody may be mathematically defined as a complete sequence of musical notes that can be regarded as a single entity [5]. Length of a melody refers to the number of notes in it. Significance of a melody is computed by multiplying its length with the number of occurrences in the entire note sequence (this formula should be used for monophonic music such as Indian classical music. Monophonic music means there is a single melody line). For polyphonic music, one may use the formula given in [6].

As a final comment, there is no unique way of analyzing music, be it structure or performance, and hence statistics and probability are likely to play important roles [7]. For a good bibliography of statistical applications in musicology see [8].

## 2. METHODOLOGY

**Simple Exponential smoothing** is used to statistically model time series data for smoothing purpose or for prediction. Although it was Holt [9] who proposed it first, it is Brown's simple exponential smoothing that is commonly used nowadays (Brown [10]). Simple exponential smoothing is achieved by the model $F_{t+1} = \alpha Y_t + (1-\alpha)Ft$, $0<\alpha<1$, to the data $(t, Y_t)$ where $F_t$ is the predicted against $Y_t$ and initially $F_o = Y_o$. Here $\alpha$ is the smoothing factor. This is the only parameter in the model that needs to be determined from the data. The smoothed statistic $F_{t+1}$ is a simple weighted average of the previous observation $Y_t$ and the previous smoothed statistic Ft. The term *smoothing factor* applied to $\alpha$ here is something of a misnomer, as larger values of $\alpha$ actually reduce the level of smoothing, and in the limiting case with $\alpha = 1$ the output series is just the same as the original series (with lag of one time unit). Simple exponential smoothing is easily applied and it produces a smoothed statistic as soon as two observations are available. Values of $\alpha$ close to one have less of a smoothing effect and give greater weight to recent changes in the data, while values of $\alpha$ closer to zero have a greater smoothing effect and are less responsive to recent changes. There is no formally correct procedure for choosing $\alpha$. Sometimes the statistician's judgment is used to choose an appropriate factor. Alternatively, a statistical technique may be used to *optimize* the value of $\alpha$. For example, the method of least squares might be used to determine the value of $\alpha$ for which the sum of the quantities $(F_t - Y_t)^2$ is minimized. (see [11] for further literature).

### Entropy Analysis

*Definition 1*: If P(E) is the probability of an event, the *information content* of the event E is defined as $I(E) = -\log_2(P(E))$. Events with lower probability will signal higher information content when they occur.

*Definition* 2: Let X be a discrete random variable which takes values $x_1\ x_2\ x_3 \ldots\ldots x_n$ with corresponding probabilities $p_1\ p_2\ p_3 \ldots\ldots p_n$. Since X is a random variable, the information content of X is also random which we denote by I(X) (what value I(X) will take depends on what value X takes). When $X = x_j$ which is an event with probability $p_j$ then $I(X)=-\log_2(p_j)$. Accordingly, it makes sense to talk about the mean value of I(X) called its *entropy*, denoted by H(X), so that we have

$H(X) = - \Sigma\, p_j \log_2(p_j)$, where the summation is over j = 1 to n.

Although only raga structure is analyzed in the present work, the ideas are applicable to performance as well. Note that for an impossible event E, $P(E)=0$, $I(E)= -\infty$. As negative information is ruled out, it indicates the non-feasibility of ever obtaining information about an impossible event.

Remark: We shall define $p\log(p) = 0$ when $p = 0$. The range of $p\log(p)$ is thus $[0, \infty)$

### Getting the musical data for structure analysis

Time series is a series of observations in chronological order. Musical data can also be taken as a time series in which a musical note characterized by pitch $\mathbf{Y_t}$ is the *entry* corresponding to the *argument* time **t** which may mean time of clock in actual performance or just the instance at which the note is realized. In our case, since we are modeling only the structure of the raga, so the arguments will be simply the instances 1, 2, 3....The desired note sequence is given in table 1. Western Art Music readers should refer to table 2 where corresponding western notations are provided. The tonic (Sa in Indian music) is taken at natural C. Analyzing the structure of a musical piece helps in giving an approximate model that captures the note progression in general without bringing the style of a particular artist into play. On the other hand, performance analysis gives additional features like note duration and the pitch movements between notes etc [1].

In table 2, pitches of notes in three octaves are represented by corresponding integers, C of the middle octave being assigned the value 0.We are motivated by the works of Adiloglu, Noll and Obermayer [6]. The letters S, R, G, M, P, D and N stand for Sa, *Sudh* Re, *Sudh* Ga, *Sudh* Ma, Pa, *Sudh* Dha and *Sudh* Ni respectively. The letters r, g, m, d, n represent *Komal* Re, *Komal* Ga, *Tibra* Ma, *Komal* Dha and *Komal* Ni respectively. Normal type indicates the note belongs to middle octave; italics implies that the note belongs to the octave just lower than the middle octave while a bold type indicates it belongs to the octave just higher than the middle octave. The terms "*Sudh*", "*Komal*" and "*Tibra*" imply, respectively, natural, flat and sharp.

**Table 1: Raga *Bhairav* Note Sequence**

| t | Note | Y(t) | t | Note | Y(t) | t | Note | Y(t) | T | Note | Y(t) | t | Note | Y(t) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | S | 0 | 51 | M | 5 | 101 | **S** | 12 | 151 | d | 8 | 201 | r | 1 |
| 2 | r | 1 | 52 | r | 1 | 102 | D | 8 | 152 | M | 5 | 202 | r | 1 |
| 3 | S | 0 | 53 | S | 0 | 103 | **R** | 13 | 153 | P | 7 | 203 | S | 0 |
| 4 | *d* | -4 | 54 | S | 0 | 104 | **S** | 12 | 154 | G | 4 | | | |
| 5 | *N* | -1 | 55 | G | 4 | 105 | N | 11 | 155 | M | 5 | | | |
| 6 | S | 0 | 56 | M | 5 | 106 | D | 8 | 156 | r | 1 | | | |
| 7 | r | 1 | 57 | P | 7 | 107 | D | 8 | 157 | r | 1 | | | |
| 8 | r | 1 | 58 | d | 8 | 108 | **S** | 12 | 158 | r | 1 | | | |
| 9 | S | 0 | 59 | M | 5 | 109 | **R** | 13 | 159 | S | 0 | | | |
| 10 | *d* | -4 | 60 | P | 7 | 110 | **S** | 12 | 160 | S | 0 | | | |
| 11 | r | 1 | 61 | d | 8 | 111 | **S** | 12 | 161 | G | 4 | | | |
| 12 | S | 0 | 62 | d | 8 | 112 | **R** | 13 | 162 | M | 5 | | | |
| 13 | *N* | -1 | 63 | N | 11 | 113 | **G** | 16 | 163 | P | 7 | | | |
| 14 | *d* | -4 | 64 | d | 8 | 114 | **r** | 13 | 164 | G | 4 | | | |
| 15 | *P* | -5 | 65 | N | 11 | 115 | **G** | 16 | 165 | M | 5 | | | |
| 16 | *M* | -7 | 66 | **S** | 12 | 116 | **G** | 16 | 166 | N | 11 | | | |
| 17 | *P* | -5 | 67 | d | 8 | 117 | **M** | 17 | 167 | d | 8 | | | |
| 18 | *N* | -1 | 68 | N | 11 | 118 | **r** | 13 | 168 | N | 11 | | | |
| 19 | *d* | -4 | 69 | d | 8 | 119 | **S** | 12 | 169 | d | 8 | | | |
| 20 | *N* | -1 | 70 | M | 5 | 120 | N | 11 | 170 | P | 7 | | | |
| 21 | *S* | 0 | 71 | P | 7 | 121 | **S** | 12 | 171 | G | 4 | | | |
| 22 | S | 0 | 72 | G | 4 | 122 | **G** | 16 | 172 | M | 5 | | | |
| 23 | r | 1 | 73 | M | 5 | 123 | **M** | 17 | 173 | P | 7 | | | |
| 24 | G | 4 | 74 | r | 1 | 124 | **G** | 16 | 174 | d | 8 | | | |
| 25 | G | 4 | 75 | G | 4 | 125 | **M** | 17 | 175 | N | 11 | | | |
| 26 | M | 5 | 76 | M | 5 | 126 | **r** | 13 | 176 | **S** | 12 | | | |
| 27 | r | 1 | 77 | P | 7 | 127 | **M** | 17 | 177 | **r** | 13 | | | |
| 28 | S | 0 | 78 | M | 5 | 128 | **r** | 13 | 178 | **S** | 12 | | | |
| 29 | *d* | -4 | 79 | d | 8 | 129 | **r** | 13 | 179 | d | 8 | | | |
| 30 | *N* | -1 | 80 | M | 5 | 130 | **S** | 12 | 180 | N | 11 | | | |
| 31 | S | 0 | 81 | P | 7 | 131 | d | 8 | 181 | **S** | 12 | | | |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 32 | G | 4 | 82 | G | 4 | 132 | N | 11 | 182 | **r** | 13 | | | |
| 33 | M | 5 | 83 | M | 5 | 133 | **S** | 12 | 183 | **S** | 12 | | | |
| 34 | P | 7 | 84 | P | 7 | 134 | d | 8 | 184 | **G** | 16 | | | |
| 35 | N | 11 | 85 | d | 8 | 135 | **r** | 13 | 185 | **M** | 17 | | | |
| 36 | d | 8 | 86 | N | 11 | 136 | **S** | 12 | 186 | **r** | 13 | | | |
| 37 | N | 11 | 87 | d | 8 | 137 | **S** | 12 | 187 | **S** | 12 | | | |
| 38 | d | 8 | 88 | P | 7 | 138 | N | 11 | 188 | **S** | 12 | | | |
| 39 | P | 7 | 89 | G | 4 | 139 | d | 8 | 189 | N | 11 | | | |
| 40 | G | 4 | 90 | M | 5 | 140 | P | 7 | 190 | d | 8 | | | |
| 41 | M | 5 | 91 | d | 8 | 141 | G | 4 | 191 | P | 7 | | | |
| 42 | P | 7 | 92 | P | 7 | 142 | M | 5 | 192 | d | 8 | | | |
| 43 | G | 4 | 93 | d | 8 | 143 | d | 8 | 193 | M | 5 | | | |
| 44 | M | 5 | 94 | N | 11 | 144 | P | 7 | 194 | P | 7 | | | |
| 45 | d | 8 | 95 | **S** | 12 | 145 | d | 8 | 195 | G | 4 | | | |
| 46 | P | 7 | 96 | **r** | 13 | 146 | N | 11 | 196 | M | 5 | | | |
| 47 | d | 8 | 97 | **r** | 13 | 147 | **S** | 12 | 197 | d | 8 | | | |
| 48 | M | 5 | 98 | **S** | 12 | 148 | N | 11 | 198 | P | 7 | | | |
| 49 | P | 7 | 99 | N | 11 | 149 | d | 8 | 199 | G | 4 | | | |
| 50 | G | 4 | 100 | d | 8 | 150 | P | 7 | 200 | M | 5 | | | |

**Table 2: Numbers representing pitches of musical notes in three octaves [6]**

| C | Db | D | Eb | E | F | F# | G | Ab | A | Bb | B | Western notation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| S | r | R | g | G | M | m | P | d | D | n | N | Notes (lower octave) |
| -12 | -11 | -10 | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | Numbers for Pitch |
| S | r | R | g | G | M | m | P | d | D | n | N | Notes (middle octave) |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | Numbers for Pitch |
| S | r | R | g | G | M | m | P | d | D | n | N | Notes (higher octave) |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | Numbers for Pitch |

## 3. STATISTICAL ANALYSIS

### Simple Exponential Smoothing fitted to *Bhairav* note sequence

The fit is found to be explaining the note progression well enough with smoothing factor 0.762839. Should such a model work well for the other ragas also , it is of interest to see how the smoothing factor varies. Here is a summary of our results obtained using Minitab Statistical package version 16:-

### Single Exponential Smoothing for C1

* NOTE * Zero values of Yt exist; MAPE calculated only for non-zero Y

Data C1

Length 20

Smoothing Constant

Alpha 0.762839

Accuracy Measures

MAPE 48.1267

MAD 1.9867

MSD 5.5180

```
Time   C1   Smooth  Predict     Error
1       0   0.0327   0.1380  -0.13796
2       1   0.7706   0.0327   0.96728
3       0   0.1828   0.7706  -0.77060
4      -4  -3.0080   0.1828  -4.18276
5      -1  -1.4762  -3.0080   2.00802
6       0  -0.3501  -1.4762   1.47622
7       1   0.6798  -0.3501   1.35010
8       1   0.9241   0.6798   0.32019
9       0   0.2192   0.9241  -0.92406
10     -4  -2.9994   0.2192  -4.21915
11      1   0.0515  -2.9994   3.99938
12      0   0.0122   0.0515  -0.05150
13     -1  -0.7599   0.0122  -1.01221
14     -4  -3.2316  -0.7599  -3.24006
15     -5  -4.5806  -3.2316  -1.76841
16     -7  -6.4262  -4.5806  -2.41940
17     -5  -5.3382  -6.4262   1.42621
18     -1  -2.0289  -5.3382   4.33824
```

| | | | | |
|---|---|---|---|---|
| 19 | -4 | -3.5325 | -2.0289 | -1.97114 |
| 20 | -1 | -1.6006 | -3.5325 | 2.53252 |
| 21 | 0 | -0.3796 | -1.6006 | 1.60061 |
| 22 | 0 | -0.0900 | -0.3796 | 0.37960 |
| 23 | 1 | 0.7415 | -0.0900 | 1.09003 |
| 24 | 4 | 3.2272 | 0.7415 | 3.25851 |
| 25 | 4 | 3.8167 | 3.2272 | 0.77279 |
| 26 | 5 | 4.7194 | 3.8167 | 1.18328 |
| 27 | 1 | 1.8821 | 4.7194 | -3.71937 |
| 28 | 0 | 0.4464 | 1.8821 | -1.88209 |
| 29 | -4 | -2.9455 | 0.4464 | -4.44636 |
| 30 | -1 | -1.4614 | -2.9455 | 1.94550 |
| 31 | 0 | -0.3466 | -1.4614 | 1.46140 |
| 32 | 4 | 2.9692 | -0.3466 | 4.34659 |
| 33 | 5 | 4.5184 | 2.9692 | 2.03084 |
| 34 | 7 | 6.4115 | 4.5184 | 2.48163 |
| 35 | 11 | 9.9118 | 6.4115 | 4.58855 |
| 36 | 8 | 8.4534 | 9.9118 | -1.91178 |
| 37 | 11 | 10.3960 | 8.4534 | 2.54660 |
| 38 | 8 | 8.5682 | 10.3960 | -2.39605 |
| 39 | 7 | 7.3719 | 8.5682 | -1.56825 |
| 40 | 4 | 4.7997 | 7.3719 | -3.37193 |
| 41 | 5 | 4.9525 | 4.7997 | 0.20031 |
| 42 | 7 | 6.5144 | 4.9525 | 2.04751 |
| 43 | 4 | 4.5963 | 6.5144 | -2.51441 |
| 44 | 5 | 4.9043 | 4.5963 | 0.40368 |
| 45 | 8 | 7.2658 | 4.9043 | 3.09574 |
| 46 | 7 | 7.0630 | 7.2658 | -0.26581 |
| 47 | 8 | 7.7778 | 7.0630 | 0.93696 |
| 48 | 5 | 5.6588 | 7.7778 | -2.77779 |
| 49 | 7 | 6.6819 | 5.6588 | 1.34122 |
| 50 | 4 | 4.6360 | 6.6819 | -2.68192 |
| 51 | 5 | 4.9137 | 4.6360 | 0.36396 |
| 52 | 1 | 1.9282 | 4.9137 | -3.91368 |
| 53 | 0 | 0.4573 | 1.9282 | -1.92817 |
| 54 | 0 | 0.1085 | 0.4573 | -0.45729 |
| 55 | 4 | 3.0771 | 0.1085 | 3.89155 |
| 56 | 5 | 4.5440 | 3.0771 | 1.92292 |
| 57 | 7 | 6.4175 | 4.5440 | 2.45604 |
| 58 | 8 | 7.6247 | 6.4175 | 1.58248 |
| 59 | 5 | 5.6225 | 7.6247 | -2.62470 |
| 60 | 7 | 6.6733 | 5.6225 | 1.37753 |
| 61 | 8 | 7.6854 | 6.6733 | 1.32669 |

| | | | | |
|---|---|---|---|---|
| 62 | 8 | 7.9254 | 7.6854 | 0.31464 |
| 63 | 11 | 10.2708 | 7.9254 | 3.07462 |
| 64 | 8 | 8.5385 | 10.2708 | -2.27082 |
| 65 | 11 | 10.4162 | 8.5385 | 2.46145 |
| 66 | 12 | 11.6244 | 10.4162 | 1.58376 |
| 67 | 8 | 8.8596 | 11.6244 | -3.62439 |
| 68 | 11 | 10.4924 | 8.8596 | 2.14044 |
| 69 | 8 | 8.5911 | 10.4924 | -2.49237 |
| 70 | 5 | 5.8517 | 8.5911 | -3.59109 |
| 71 | 7 | 6.7277 | 5.8517 | 1.14833 |
| 72 | 4 | 4.6469 | 6.7277 | -2.72766 |
| 73 | 5 | 4.9163 | 4.6469 | 0.35311 |
| 74 | 1 | 1.9288 | 4.9163 | -3.91626 |
| 75 | 4 | 3.5088 | 1.9288 | 2.07122 |
| 76 | 5 | 4.6463 | 3.5088 | 1.49121 |
| 77 | 7 | 6.4418 | 4.6463 | 2.35366 |
| 78 | 5 | 5.3419 | 6.4418 | -1.44181 |
| 79 | 8 | 7.3696 | 5.3419 | 2.65806 |
| 80 | 5 | 5.5620 | 7.3696 | -2.36961 |
| 81 | 7 | 6.6590 | 5.5620 | 1.43802 |
| 82 | 4 | 4.6306 | 6.6590 | -2.65896 |
| 83 | 5 | 4.9124 | 4.6306 | 0.36940 |
| 84 | 7 | 6.5049 | 4.9124 | 2.08761 |
| 85 | 8 | 7.6454 | 6.5049 | 1.49510 |
| 86 | 11 | 10.2044 | 7.6454 | 3.35458 |
| 87 | 8 | 8.5228 | 10.2044 | -2.20443 |
| 88 | 7 | 7.3611 | 8.5228 | -1.52280 |
| 89 | 4 | 4.7971 | 7.3611 | -3.36115 |
| 90 | 5 | 4.9519 | 4.7971 | 0.20287 |
| 91 | 8 | 7.2771 | 4.9519 | 3.04811 |
| 92 | 7 | 7.0657 | 7.2771 | -0.27711 |
| 93 | 8 | 7.7784 | 7.0657 | 0.93428 |
| 94 | 11 | 10.2360 | 7.7784 | 3.22157 |
| 95 | 12 | 11.5816 | 10.2360 | 1.76403 |
| 96 | 13 | 12.6636 | 11.5816 | 1.41836 |
| 97 | 13 | 12.9202 | 12.6636 | 0.33638 |
| 98 | 12 | 12.2182 | 12.9202 | -0.92022 |
| 99 | 11 | 11.2889 | 12.2182 | -1.21824 |
| 100 | 8 | 8.7800 | 11.2889 | -3.28892 |
| 101 | 12 | 11.2363 | 8.7800 | 3.22000 |
| 102 | 8 | 8.7675 | 11.2363 | -3.23634 |
| 103 | 13 | 11.9962 | 8.7675 | 4.23247 |
| 104 | 12 | 11.9991 | 11.9962 | 0.00377 |

| | | | | |
|---|---|---|---|---|
| 105 | 11 | 11.2369 | 11.9991 | -0.99910 |
| 106 | 8 | 8.7677 | 11.2369 | -3.23695 |
| 107 | 8 | 8.1821 | 8.7677 | -0.76768 |
| 108 | 12 | 11.0945 | 8.1821 | 3.81794 |
| 109 | 13 | 12.5481 | 11.0945 | 1.90546 |
| 110 | 12 | 12.1300 | 12.5481 | -0.54810 |
| 111 | 12 | 12.0308 | 12.1300 | -0.12999 |
| 112 | 13 | 12.7702 | 12.0308 | 0.96917 |
| 113 | 16 | 15.2340 | 12.7702 | 3.22985 |
| 114 | 13 | 13.5298 | 15.2340 | -2.23401 |
| 115 | 16 | 15.4142 | 13.5298 | 2.47018 |
| 116 | 16 | 15.8611 | 15.4142 | 0.58583 |
| 117 | 17 | 16.7299 | 15.8611 | 1.13894 |
| 118 | 13 | 13.8846 | 16.7299 | -3.72989 |
| 119 | 12 | 12.4469 | 13.8846 | -1.88458 |
| 120 | 11 | 11.3432 | 12.4469 | -1.44695 |
| 121 | 12 | 11.8442 | 11.3432 | 0.65684 |
| 122 | 16 | 15.0144 | 11.8442 | 4.15578 |
| 123 | 17 | 16.5291 | 15.0144 | 1.98559 |
| 124 | 16 | 16.1255 | 16.5291 | -0.52910 |
| 125 | 17 | 16.7926 | 16.1255 | 0.87452 |
| 126 | 13 | 13.8995 | 16.7926 | -3.79260 |
| 127 | 17 | 16.2647 | 13.8995 | 3.10055 |
| 128 | 13 | 13.7743 | 16.2647 | -3.26467 |
| 129 | 13 | 13.1836 | 13.7743 | -0.77425 |
| 130 | 12 | 12.2807 | 13.1836 | -1.18362 |
| 131 | 8 | 9.0152 | 12.2807 | -4.28071 |
| 132 | 11 | 10.5293 | 9.0152 | 1.98479 |
| 133 | 12 | 11.6512 | 10.5293 | 1.47071 |
| 134 | 8 | 8.8659 | 11.6512 | -3.65121 |
| 135 | 13 | 12.0196 | 8.8659 | 4.13408 |
| 136 | 12 | 12.0046 | 12.0196 | -0.01956 |
| 137 | 12 | 12.0011 | 12.0046 | -0.00464 |
| 138 | 11 | 11.2374 | 12.0011 | -1.00110 |
| 139 | 8 | 8.7678 | 11.2374 | -3.23742 |
| 140 | 7 | 7.4192 | 8.7678 | -1.76779 |
| 141 | 4 | 4.8109 | 7.4192 | -3.41925 |
| 142 | 5 | 4.9552 | 4.8109 | 0.18909 |
| 143 | 8 | 7.2779 | 4.9552 | 3.04484 |
| 144 | 7 | 7.0659 | 7.2779 | -0.27788 |
| 145 | 8 | 7.7785 | 7.0659 | 0.93410 |
| 146 | 11 | 10.2360 | 7.7785 | 3.22153 |
| 147 | 12 | 11.5816 | 10.2360 | 1.76402 |

| | | | | |
|---|---|---|---|---|
| 148 | 11 | 11.1379 | 11.5816 | -0.58164 |
| 149 | 8 | 8.7442 | 11.1379 | -3.13794 |
| 150 | 7 | 7.4137 | 8.7442 | -1.74420 |
| 151 | 8 | 7.8609 | 7.4137 | 0.58635 |
| 152 | 5 | 5.6785 | 7.8609 | -2.86094 |
| 153 | 7 | 6.6866 | 5.6785 | 1.32150 |
| 154 | 4 | 4.6372 | 6.6866 | -2.68659 |
| 155 | 5 | 4.9139 | 4.6372 | 0.36285 |
| 156 | 1 | 1.9282 | 4.9139 | -3.91395 |
| 157 | 1 | 1.2201 | 1.9282 | -0.92823 |
| 158 | 1 | 1.0522 | 1.2201 | -0.22014 |
| 159 | 0 | 0.2495 | 1.0522 | -1.05221 |
| 160 | 0 | 0.0592 | 0.2495 | -0.24954 |
| 161 | 4 | 3.0654 | 0.0592 | 3.94082 |
| 162 | 5 | 4.5412 | 3.0654 | 1.93461 |
| 163 | 7 | 6.4169 | 4.5412 | 2.45881 |
| 164 | 4 | 4.5732 | 6.4169 | -2.41687 |
| 165 | 5 | 4.8988 | 4.5732 | 0.42681 |
| 166 | 11 | 9.5530 | 4.8988 | 6.10122 |
| 167 | 8 | 8.3683 | 9.5530 | -1.55303 |
| 168 | 11 | 10.3759 | 8.3683 | 2.63168 |
| 169 | 8 | 8.5635 | 10.3759 | -2.37587 |
| 170 | 7 | 7.3708 | 8.5635 | -1.56346 |
| 171 | 4 | 4.7994 | 7.3708 | -3.37079 |
| 172 | 5 | 4.9524 | 4.7994 | 0.20058 |
| 173 | 7 | 6.5144 | 4.9524 | 2.04757 |
| 174 | 8 | 7.6477 | 6.5144 | 1.48560 |
| 175 | 11 | 10.2050 | 7.6477 | 3.35233 |
| 176 | 12 | 11.5743 | 10.2050 | 1.79504 |
| 177 | 13 | 12.6619 | 11.5743 | 1.42571 |
| 178 | 12 | 12.1570 | 12.6619 | -0.66188 |
| 179 | 8 | 8.9859 | 12.1570 | -4.15697 |
| 180 | 11 | 10.5223 | 8.9859 | 2.01413 |
| 181 | 12 | 11.6496 | 10.5223 | 1.47767 |
| 182 | 13 | 12.6797 | 11.6496 | 1.35045 |
| 183 | 12 | 12.1612 | 12.6797 | -0.67973 |
| 184 | 16 | 15.0896 | 12.1612 | 3.83880 |
| 185 | 17 | 16.5469 | 15.0896 | 1.91041 |
| 186 | 13 | 13.8412 | 16.5469 | -3.54693 |
| 187 | 12 | 12.4367 | 13.8412 | -1.84119 |
| 188 | 12 | 12.1036 | 12.4367 | -0.43666 |
| 189 | 11 | 11.2617 | 12.1036 | -1.10356 |
| 190 | 8 | 8.7736 | 11.2617 | -3.26172 |

```
191     7   7.4206   8.7736  -1.77355
192     8   7.8626   7.4206   0.57938
193     5   5.6789   7.8626  -2.86259
194     7   6.6867   5.6789   1.32111
195     4   4.6372   6.6867  -2.68669
196     5   4.9140   4.6372   0.36282
197     8   7.2681   4.9140   3.08605
198     7   7.0636   7.2681  -0.26811
199     4   4.7266   7.0636  -3.06359
200     5   4.9352   4.7266   0.27344
201     1   1.9333   4.9352  -3.93515
202     1   1.2213   1.9333  -0.93326
203     0   0.2897   1.2213  -1.22133
```

**Single Exponential Smoothing Plot for C1**

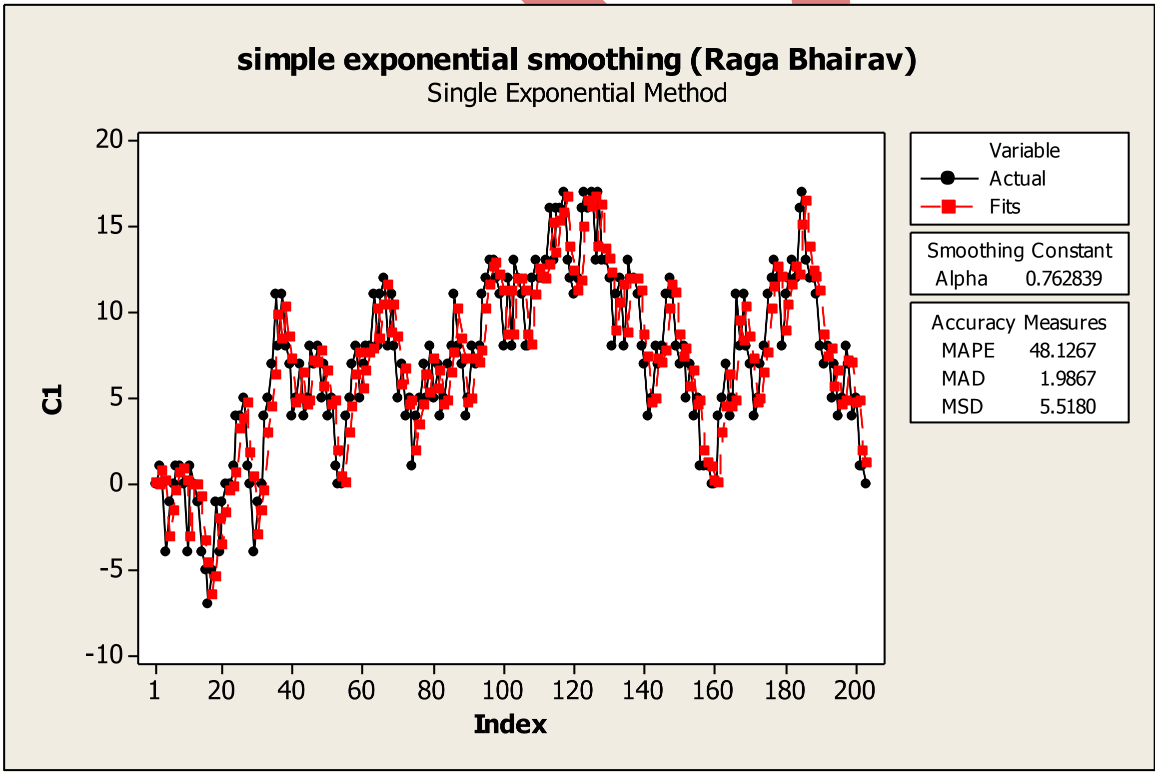

**Fig. 1: Simple Exponential Smoothing captures the Bhairav note sequence**

## Residual Plots for C1

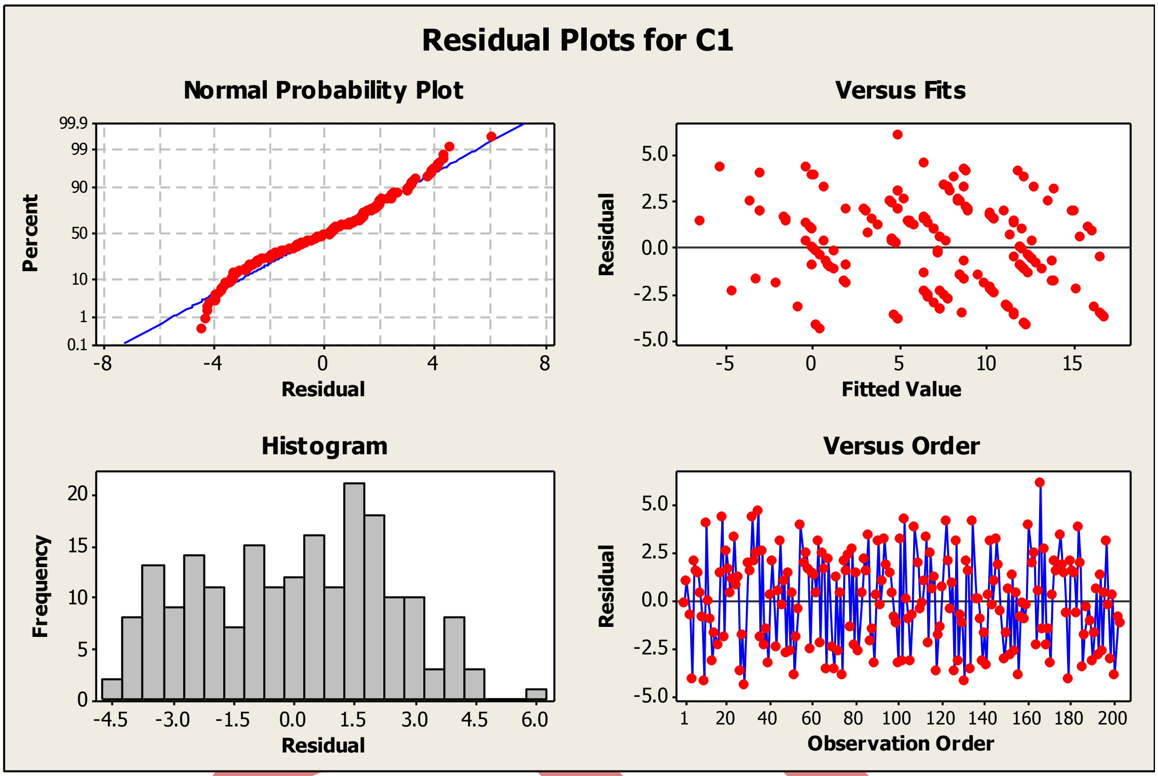

**Fig. 2: Residual Plots**

## Entropy Analysis of *Bhairav*

**Table 3 Information content of notes of *Bhairav* with varying probability**

| **Note Occurrence x** | **Information Content = -log$_2$(x/203)** |
|---|---|
| S=35 | I(S) = 2.5361 |
| r=27 | I(r)= 2.9104 |
| G=24 | I(G)= 3.0804 |
| M=30 | I(M)= 2.7584 |
| P=25 | I(P)= 3.0215 |
| d=38 | I(d)= 2.4174 |
| N=24 | I(N)= 3.0804 |

Also, mean entropy of Bhairav notes ignoring octave = 2. 7851

## Melody analysis

Table 4 with reference to table 1 gives the melody groups

**Table 4 : *Bhairav* Melody Groups**

| Note Sr.No. | Group No. | Length |
|---|---|---|
| 1-3 | G1 | 3 |
| 4-6 | G2 | 3 |
| 7-9 | G3 | 3 |
| 10-15 | G4 | 6 |
| 16-27 | G5 | 12 |
| 28-31 | G6 | 4 |
| 32-34 | G7 | 3 |
| 35-39 | G8 | 5 |
| 40-42 | G9 | 3 |
| 43-46 | G10 | 4 |
| 47-53 | G11 | 7 |
| 54-57 | G12 | 4 |
| 58-60 | G13 | 3 |
| 61-67 | G14 | 7 |
| 68-71 | G15 | 4 |
| 72-74 | G16 | 3 |
| 75-77 | G17 | 3 |
| 78-81 | G18 | 4 |
| 82-85 | G19 | 4 |
| 86-88 | G20 | 3 |
| 89-92 | G21 | 4 |
| 93-98 | G22 | 6 |
| 99-104 | G23 | 6 |
| 105-110 | G24 | 6 |
| 111-115 | G25 | 5 |
| 116-119 | G26 | 4 |
| 120-123 | G27 | 4 |
| 124-126 | G28 | 3 |
| 127-130 | G29 | 4 |
| 131-136 | G30 | 6 |
| 137-140 | G31 | 4 |
| 141-144 | G32 | 4 |
| 145-147 | G33 | 3 |
| 148-159 | G34 | 12 |
| 160-163 | G35 | 4 |
| 164-170 | G36 | 7 |

| | | |
|---|---|---|
| 171-178 | G37 | 8 |
| 179-183 | G38 | 5 |
| 184-187 | G39 | 4 |
| 188-191 | G40 | 4 |
| 192-194 | G41 | 3 |
| 195-198 | G42 | 4 |
| 199-203 | G43 | 5 |

Table 5 gives the significance of the melody groups.

**Table 5: Melody groups and their significance**

| Melody Group | Significance |
|---|---|
| G6,G10,G12,G15,G18,G19,G21,G26,<br>G27,G29,G31,G32,G35,G39,G40,G42 | 64 |
| G1,G2,G3,G7,G9,G13,G16,<br>G17,G20,G28,G33,G41 | 36 |
| G4,G22,G23,G24,G30 | 30 |
| G5,G34 | 24 |
| G11,G14,G36 | 21 |
| G8,G25,G38,G43 | 20 |
| G37 | 8 |

## 4. DISCUSSION

### Interpretations from figures 1 and 2:-

**The random pattern of the residuals (fig. 2) together with the closeness of smoothed data with the observed one (fig. 1) justifies the Simple Exponential Smoothing**. A detailed discussion of the findings is given next.

**MAPE** (Mean Absolute Percent Error) - measures the accuracy of fitted time series values. It expresses accuracy as a percentage.

**MAD** (Mean Absolute Deviation) - measures the accuracy of fitted time series values. It expresses accuracy in the same units as the data, which helps conceptualize the amount of error.

**MSD** (Mean Squared Deviation) - measures the accuracy of fitted time series values. MSD is always computed using the same denominator (the number of forecasts) regardless of the model, so one can compare MSD values across models and hence compare the accuracy of two different models.

For all three measures, smaller values generally indicate a better fitting model. In case we fit other models to the same data, it is of interest to compare the corresponding MAPE, MAD and MSD values. This is reserved as a rewarding future work.

The normal probability graph plots the residuals versus their expected values when the distribution is normal. The residuals from the analysis should be normally distributed. In practice, for data with a large number of observations, moderate departures from normality do not seriously affect the results.

The normal probability plot of the residuals should roughly follow a straight line. One can use this plot to look for the following:

| This pattern... | Indicates... |
|---|---|
| Not a straight line | Nonnormality |
| Curve in the tails | Skewness |
| A point far away from the line | An outlier |

| | |
|---|---|
| Changing slope | An unidentified variable |

**As is clear from fig. 2, the plot roughly follows a straight line.**

The next graph plots the residuals versus the fitted values. The residuals should be scattered randomly about zero. One can use this plot to look for the following:

| This pattern... | Indicates... |
|---|---|
| Fanning or uneven spreading of residuals across fitted values | Non-constant variance |
| Curvilinear | A missing higher-order term |
| A point far away from zero | An outlier |

**As is clear from fig. 2, there are more residuals on the positive side than the negative side. We may make transformations to stabilize variance by working on some function of the response rather than the response itself although its musical meaningfulness is difficult to justify and hence omitted.**

A histogram of the residuals shows the distribution of the residuals for all observations. One can use the histogram as an exploratory tool to learn about the following characteristics of the data:

- Typical values, spread or variation, and shape
- Unusual values in the data

The histogram of the residuals should be bell-shaped. One can use this plot to look for the following:

| This pattern... | Indicates... |
|---|---|
| Long tails | Skewness |
| A bar far away from the other bars | An outlier |

Because the appearance of the histogram can change depending on the number of intervals used to group the data, one should use the normal probability plot and goodness-of-fit tests to assess whether the residuals are normal. We have already given the normal probability plot for residuals.

The graph “residuals versus order” plots the residuals in the order of the corresponding observations. The plot is useful when the order of the observations may influence the results, which can occur when data are collected in a time sequence (as in our case) or in some other sequence, such as geographic area. This plot can be particularly helpful in a designed experiment in which the runs are not randomized.

The residuals in the plot should fluctuate in a random pattern around the center line as in fig. 2. One can examine the plot to see if any correlation exists between error terms that are near each other. Correlation among residuals may be signified by:

- An ascending or descending trend in the residuals
- Rapid changes in signs of adjacent residuals

Remark:

Simple exponential smoothing is useful when (i) there is no trend (ii) there is no seasonal variation (iii) there is no missing value (iv) we want short term forecast.

From the entropy analysis, we find that I(G) = I(N) > I(P) > I(r) > I(M) > I(S) > I(d)

Finally, from the melody analysis, we find that 64 is numerically the highest melody significance measure and that it is attained by a good number of melody groups.

## 4. CONCLUSION

Music analysis, broadly speaking, can be divided into commonality analysis (“what is common?”) and diversity analysis (“what is special*?”). A statistician is essentially a commonality expert in the sense that the philosophy of statistics is to summarize and average and make inferences which are true on the whole and which describe a process rather than an individual entity.* Fortunately, there are issues in music where this traditional mindset of the statistician finds a support. For example, a collection of recordings of the same artist if analyzed statistically will definitely reflect certain common features having to do with the style of the artist. In our analysis, we found different melody groups sharing the same significance measure (commonality). But the statistician must realize that every single music piece will have something special to offer.

Thus the smoothing factor of 0.762839 is typical of this raga structure that was taken from a standard text (Dutta [12]) and hence is a diversity for that text. Another raga from the same text can have a similar smoothing factor but if it comes close, it is of interest to seek musical commonality in the ragas concerned. On the other hand, the particular modeling we have done in this paper may not in general hold good for all raga structures. Similar comments can be made about entropy analysis. There may other ragas that use the same notes but their information contents would be, in general, quite different. Hence the entropy analysis also reflects a diversity. It appears that diversity is more important than commonality in music analysis. Fortunately, there are issues even in statistics where the statistician does take an individual observation seriously-as in the case of an outlier or influential observation for example. There is a whole literature in statistics to deal with outliers. When an outlier comes, the traditional philosophy of summarizing and averaging is brushed aside. The statistician goes after this individual influential observation exploring how it came and what it signifies. The case of outliers is an exception in statistics. It is the very grammar in music as music is a work of art! *If the statistician can use his experience and mindset of handling outliers (regarding every musical performance or structure as a musical outlier) alongwith his commonality expertise, he can be a very effective music analyst*.** That said, we close the paper.

**Remark: We wholeheartedly appreciate a similar point of view expressed very strongly by Nettheim [8].

## APPENDIX:

**Table 6: Musical features of raga *Bhairav***

| **Musical feature** | ***Raga Bhairav*** |
|---|---|
| ***Thaat (a raga group based on scale)*** | *Bhairav* |
| ***Aroh* (ascent)** | S r G M, P d, N **S** |
| ***Awaroh* (descent)** | **S** N d P M G r S |
| ***Vadi* (most imp note)** | d |
| ***Samvadi* (second most imp note)** | r |
| ***Pakad* (catch)** | G M P d, P d M P G M, r S |
| ***Jati* (no. of distinct notes used in ascent and descent)** | *Sampoorna-Sampoorna*<br>(7 distinct notes in ascent;<br>7 distinct notes in descent) |
| ***Nyas swar* (stay notes)** | r, M, P, d |
| ***Anga*** | *Uttaranga-pradhan meaning that the second half is more important* |
| **Time of rendition** | First raga of the morning<br>(5AM – 8 AM) |
| **Nature** | Restful and serious |

**N.B.**

**Suresh Chandra Chakraborty**, a well known musicologist, has asserted that *Bhairav* although called the *Adi*-raga being the first raga of the morning is not older than the raga *Bhairavi*. The "*Komal Re*" (r) of *Bhairav* is actually *ati komal* (a microtone whose pitch is between 0 and 1 in our notation but we have taken it as 1 for simplicity) and *andolita* (oscillating) in a way that helps create the raga mood. The "*Komal Dha*" (d) is also *andolita* but only *komal*, not *ati komal*. In performance, it is necessary to give a *meend* (glide) from M to r and also from S or N to d[13].

## REFERENCES

[1] Chakraborty S, Ranganayakulu R, Chauhan S, Solanki SS, Mahto K (Winter 2009) A Statistical Analysis of Raga Ahir Bhairav. Journal of Music and Meaning, Vol. 8, sec. 4, http://www.musicandmeaning.net/issues/showArticle.php?artID=8.4

[2] Klemens, B., (2008) *Modeling with data : tools and techniques for scientific computing*, Princeton University Press, Princeton, NJ

[3] Applebaum, D., (1996) Probability and information: An integrated approach, Cambridge University Press, chapter 6

[4] Snyder, J. L., (1990) *Entropy as a measure of musical style: the influence of a priori* Assumptions, Music Theory Spectrum, 12, 121-160

[5] Chakraborty, S., Krishnapryia, K. Loveleen, Chauhan, S., Solanki, S. S. and Mahto, K. (2010) Melody Revisited: Tips from Indian Music Theory, International Journal of Computational Cognition, Vol. 8, No. 3, 2010, 26-32

[6] Adiloglu, K., Noll, T. and Obermayer, K., (2006) *A Paradigmatic Approach to Extract the Melodic Structure of a Musical Piece*, Journal of New Music Research, Vol. 35(3), 221-236

[7] Beran, J. and Mazzola, G., (1999) Analyzing Musical Structure and Performance A Statistical Approach, Statistical Science, vol. 14, no. 1, p. 47-79

[8] Nettheim, N., (1997), A Bibliograhy of Statistical Applications in Musicology, Musicology Australia, Vol. 20, 94-106

[9] Holt, Charles C (1957) "Forecasting Trends and Seasonal by Exponentially Weighted Averages". Office of Naval Research Memorandum **52**. reprinted in Holt, Charles C. (January–March 2004). "Forecasting Trends and Seasonal by Exponentially Weighted Averages". International Journal of Forecasting **20** (1): 5–10. doi:10.1016/j.ijforecast.2003.09.015

[10] Brown, Robert Goodell (1963) Smoothing Forecasting and Prediction of Discrete Time Series. Englewood Cliffs, NJ: Prentice-Hall

[11] http://en.wikipedia.org/wiki/Exponential_smoothing (accessed on 28th Jan,.2013)

[12] Dutta, D., *Sangeet Tattwa* (Pratham Khanda) (2006), Brati Prakashani, 5th ed, (Bengali)

[13] Chakraborty, S. C., (1965) *Raga Rupayan* [Bengali], General Printers & Pub. Pvt. Ltd, Kolkata