Classification of Sodar Data Using Fractal Features

Abhik Mukherjee
C. S. T. Department
B. E. College (DU)
Howrah, India
abhik@becs.ac.in

Pinakpani Pal and J. Das
E. C. S. U.
Indian Statistical Institute
Calcutta, India
{pinak,jdas}@isical.ac.in

Abstract

Geometric shape features extracted from sodar data have been used to classify different atmospheric patterns. But such features are scale dependent. In this paper, an attempt is made to classify sodar data using fractal features. Results show that scale invariance of fractal features make them quite useful for classification of sodar patterns.

1. Introduction

Sodar (sound radar) is a remote sensing instrument for continuous monitoring of the lower atmosphere [12, 3, 15] by generating information about the Atmospheric Boundary Layer (*ABL*). This instrument finds applications in the areas of fog monitoring, air pollution monitoring [2, 20] and air quality measurement [19].

Chaudhuri *et al.* [7, 6] developed some basic concepts for automatic classification of sodar-pattern. They observed that decision regarding the shape of the ABL patterns being represented in the sodar-images is to be considered primarily with the boundary between the sodar-pattern and the background in the sodar-image. Tripathi *et al.* [22] presented computer algorithm for plume structures analysis. De *et al.* [8] have identified dot-echo structures and studied their nature.

Different soft computing approaches have been suggested for automatic recognition and analysis of sodar-patterns. Perceptron, a variant of artificial neural network (*ANN*), has been employed for classification [14] of elementary sodar patterns. Similar results have been obtained [18] using knowledge based network approach as well. In both cases geometric shape features like number of maxima and minima, gradient magnitude in a predefined time frame, along with diurnal and seasonal information, have been employed for classification of the patterns. These features are based on sodar data accumulated for half-hour. But with the advent of schemes of ABL extraction based on Kalman filter techniquesthe demand for classification based on dynamic features has been generated.

The seasonal as well as diurnal variability of sodar data and the patterns therein poses a serious problem in identifying appropriate features for pattern classification. Different options of classifier design can be explored following [13]. Exemplar classifier design becomes infeasible due to problems of choosing representative examples and keeping their number small. Simple perceptron-like classifier design has also earned limited success because of the wide variability of chosen feature values. Kernel classifiers, which subdivide the feature space into regions characterized by their mean and standard deviation, can be attempted in such situations. But the feature values are very much scale dependent, i.e. depends on the time frame considered for their measurement. In rule based classifier design, dynamic features like instantaneous magnitude and rate of change of ABL heights can be used. But in that case, due to wind shear and localized convective flows, micro-level changes dominate and disturb the classifier algorithm.

In this context, it is worthwhile to consider fractal features, in particular fractal dimension (*FD*) of the data, which is scale-invariant and also able to reflect the changes at micro level. Fractal features are easily updated and also have physical significance. FD indicates the extent of space filling by the data. In presence of wind, the inversion structure reports slightly higher FD, but evidently this does not quite overlap with the FD range of the plume structures dominated by convective flow. Thus the FD value can be indicative of the kind of pattern as well as the prevailing meteorological conditions. The sodar data recast on the domain of FD can effectively classify the data by employing predefined (through proper training or expertise) receptor fields for inclusion in different classes.

The present work explores the possibilities of using fractal features for interpreting sodar signals. The concept of fractal dimension is introduced in Section 2. The relevance of fractal features in case of sodar data is elaborated in Section 3. The algorithm used for computing fractal dimension of sodar data in the present scheme is explained in Section 4. Section 5 provides the results.

2. Fractal Features

Formal definition of fractal dimension is presented by Barnsley [1] in terms of the number of n dimensional balls of infinitesimally small size required to cover a set defined in R^n. This definition was later extended to the box counting dimension.

Let $N_\epsilon(S)$ be the number of cubes of size ϵ required to cover a surface S. Now if $N_\epsilon(S) = \frac{1}{\epsilon^d}$ then d is the dimension of S as ϵ tends to 0. Then, $d = \lim_{\epsilon \to 0} \frac{\log N_\epsilon(S)}{\log(1/\epsilon)}$. For a set S defined in R^n, cuboids of size ϵ are considered. From this definition, a square has dimension 2, cube has dimension 3 etc. But fractional dimension value also comes into fray whenever a curved line or curved surface is considered. Though the concept was originally defined in terms of spheroids or balls but box counting is simpler to implement numerically instead of balls.

The concept of FD originated as a scale invariant metric to distinguish between different dynamical systems. In case of the Cantor set, the elements in the n-th iteration are evolved by depleting the middle-one-third of all the subintervals in the $n-1$-th iteration. The reduction of the scale successively by $\frac{1}{3}$-rd results in the increase of number of boxes (one-dimensional) by powers of 2 , i.e. for a scale of $(\frac{1}{3})^n$ the number of boxes is 2^n. This yields a FD of $\frac{\log 2}{\log 3}$. Similarly in case of Sierspinskii triangle, which is obtained by successively removing the middle triangle of all remaining triangles (obtained by joining mid-points of the larger triangles) considering boxes whose size is successively reduced by $\frac{1}{2}$, every four new boxes include 3 triangles. Thereby the box count increases by powers of 3. Hence for box size of $\frac{1}{2^n}$, number of boxes is 3^n, thus resulting in FD of $\frac{\log 3}{\log 2}$. These results, derived for perfect fractal objects, reveal that dynamical systems are indeed characterized by their FD independent of scale.

However FD computed from experimental data of unknown dynamical system is not constant as in case of perfect fractals. This is mainly due to presence of noise in the system. Also, the actual systems are often not exactly as self repetitive as proposed in the model. However the relevance of self repetitiveness in finding a scale-invariant metric like FD is evident from measurements in case of different systems.

Actually in case of a sodar-based experiment some of the finer details of dynamics of ABL evolution in lower atmosphere are missed out. A scale invariant metric like the FD can account for the dynamics in a better way.

3. Fractalness of Sodar data

In order to appreciate the characteristics of sodar data, it is worthwhile to provide some physical insight into the dynamics of ABL which this data captures. During daytime, the boundary layer is unstable due to convection. The elevated inversion height rises with solar heating of earth surface. the height is maximum around midday, when the solar heat flux maximizes and then gradually decays in the afternoon. The height of the convective layer can be expressed in closed form equation [19] which involves C_p, specific heat of air at constant pressure; H, the heat flux; ρ, the air density; T, the temperature near earth in absolute scale. The height h is characterized by convective velocity scale w_*, which has a typical value of 2 m/s.

$$w_* = \left(\frac{gHh}{\rho C_p T} \right)^{(1/3)} \tag{1}$$

The convective element in the unstable boundary layer mainly consists of thermal plumes of up-drafts surrounded by large down-drafts. Typical heights rise upto 1 km. The day time boundary layer is considered as unstable stratified [11] due to predominance of thermal convection.

The stable boundary layer evolves in evening with the beginning of the radiational cooling of earth surface. The night time boundary layer is formed much lower than the day time boundary layer and is mostly stable stratified [16, 17]. It contains shears above of the inversion layer due to occasional sharp increase in temperature layer [21]. There is also shear instability along with some development of a super-geostrophic jet at lower levels [4, 23]. Weak turbulence due to the local shear stress during nights is characterized by the frictional velocity u_* defined [19] as

$$u_*^2 = \frac{\tau}{\rho} = -u'w' \tag{2}$$

where τ is the surface stress, ρ is the air density and u', w' are fluctuations in vertical and horizontal wind velocity respectively. Local stress becomes zero at the top of the boundary layer. u_* is typically 0.1 – 0.3 m/s. The overall height is limited to few tens or 100 m. However due to development of gravity waves, local instabilities as well as circulation of land and sea breezes make the estimation of stable ABL height difficult. Gravity waves with regular periods are sometimes driven by wind shears at higher levels [10, 9].

In case of sodar, the sound wave sent upwards every 6 s. is backscattered from all heights of lower atmosphere (ranging upto 1 km). The strength of backscatter at the receiver is strongest from heights upto which the temperature inhomogeniety extends on account of convection and wind shear. Sodar receiver therefore captures the ABL information from all heights every 6 s. This information is therefore in the form of an image with pixels in each column representing intensity of backscatter from different heights at a particular time instant. From this information it is possible to extract the ABL heights as a function of time employing some thresholding on the intensity values.

Sodar data therefore resembles the ABL pattern. The ABL formation process is indeed a dynamical system with self-repetitive characteristics. Sodar data is thus a standard candidate for exploring fractal features. In the present case, the ABL height is considered as the data. It is understood that the nocturnal boundary layer having typically no-structure patterns would exhibit a low value of FD.The daytime flat inversion layer should result in medium FD and FD of plume structure should be high. Other inversion structures like bulge, depression have FD in between. Thus by computing and monitoring FD of the data, it is possible to monitor the transitions of sodar data from one pattern to another.

Chanda et al. [5] have worked on characterizing the turbulence from temperature field inside plume structures. Arguing that the pixel intensity of sodar data is proportional to temperature field, they have computed the FD of single column of data. Such an approach is evidently successful in characterizing turbulence at an instant of time. However, this approach is spatial and does not consider any temporal information about the data. FD computed for each column may be plotted against time to get an idea about temporal variation of turbulence. But sodar data also contains information regarding wind shear. Wind shear dominates during times of inversion and is also responsible for the variations in ABL top during nights. Plume structures also undergo distortions due to wind shear. The approach undertaken by Chanda et al. does not take into account the effects of wind shear. This can be done only by adopting a spatio-temporal approach of finding FD.

4. Proposed Algorithm

Since information lies mostly at the top of ABL, it is worthwhile to compute the FD value of ABL top instead of the $3D$ image. It is also time-wise complex to handle image data. However while counting the number of boxes, the count is in terms of the column height and not just the point itself, i.e. space-filling of the area under the ABL top height. The reported FD values would therefore range between 1 and 2.

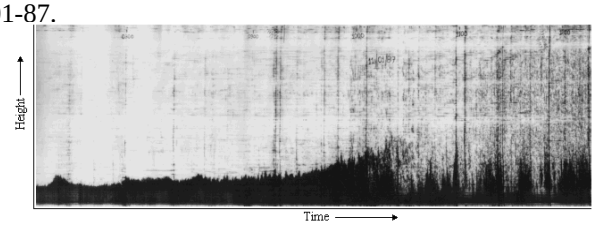
Temporal variation of ABL top height contains information regarding wind shear as well as fluctuations due to thermal convection. It can therefore be argued that FD computed over a certain time of ABL data can effectively characterize sodar data by accounting for both wind and convection. Regions of tranquility are characterized by low FD, the ones with only wind shear have medium values of FD and regions with high convection have characteristically high FD values. It is therefore worthwhile to adopt this approach for pattern classification.

The argument for being able to approximate the FD of the sodar image by computing the FD based on the ABL top height data only is presented here. For detecting the ABL boundary, the image is replaced by the two-tone binary image with the region of active ABL being represented by the highest possible intensity value of 255. Hence considering cubes (instead of squares) of size s, the number of boxes required to cover the sodar image is given by $N_{r,3} = N_{r,2} \times (\frac{256}{s})$. Hence the new dimension of the overall sodar image becomes $d' = \frac{\log N_{r,3}}{\log r} = \frac{\log N_{r,2}}{\log r} + \frac{\log(\frac{256}{s})}{\log r}$. Now since $\frac{\log N_{r,2}}{\log r} = d$, and since $r = \frac{C}{s}$ by definition, for $C = 256$, the above equation reduces to $d' = 1 + d$. Hence it is possible to consider the ABL top heights for classification of sodar patterns instead of the entire image, thus saving both time and space for extracting fractal features.

s columns of sodar data return ABL top heights h_1 through h_s. It is assumed that the intensity is uniformly distributed. Hence $h = max(h_1, \cdots, h_s)$ represents the height upto which the sodar data extends. Dividing h by the size of box s gives the number of boxes b needed to cover the data represented in these s columns. The number of columns considered is C where C is a multiple of s. Then there would be $r = \frac{C}{s}$ such box counts to be accumulated. The total number of boxes of size s counted over C columns of sodar data is given by $N(r) = \Sigma_{j=0}^{r-1} b_j$ where $b_j = \left\lceil \left(\frac{\max(h_i)}{s} \right) \right\rceil$, $i \in js, \cdots (j+1)s$. Number of boxes are counted for $s = C, \frac{C}{2}, \frac{C}{4}$ successively, giving $r = \frac{C}{s} = 1, 2, 4, \cdots$ etc. When $C = 256$ and $s = 2$; $r = 128$, $j = 0, \cdots, 127$, $i = 0, 1; 2, 3; 4, 5; \cdots; 254, 255$. Now fractal dimension due to box counting is given by the limiting value of $\frac{\log N(r)}{\log r}$ as $\frac{1}{r}$ tends to zero. In present case, $r = 128$ is considered the limiting value of ri due to the practical limitations of working with image data.

5. Results

Figure 1: Portion of continuous sodar data captured on 11-01-87.



This technique is suitable for continuous monitoring. Hence the strength of the technique is brought out using the data for a whole day. Figure 1 shows typical continuous sodar data for which the classifier is tested. Figure 2 evidently shows that the FD variations are smooth when fractal dimension is computed in a window of size 256. But the monitoring covers more details and shows more variations

Figure 2: Variation of fractal dimension with time when size of window is 256.

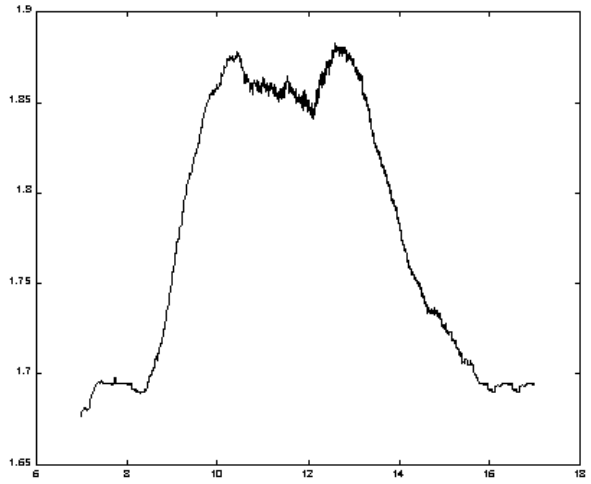
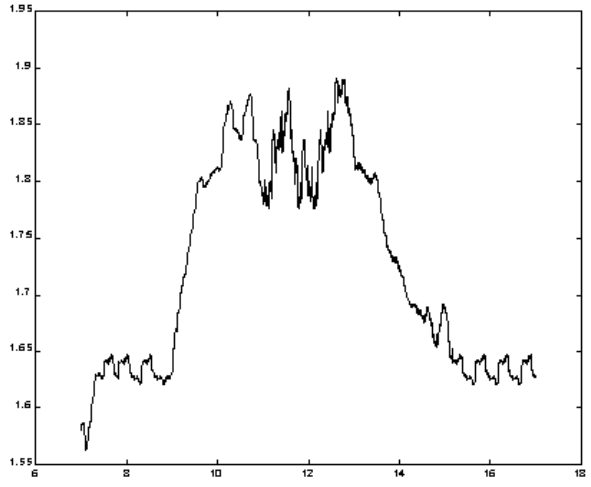


Figure 3: Variation of fractal dimension with time when size of window is 64.



when the window size is chosen to be 64, as is evident from Figure 3.

In order to demonstrate the performance of the classifier, the possibility of plume and inversion are shown graphically.

The variation of possibility of the two patterns shown in Figure 4 and 5 clearly indicate that the scheme is able to classify the patterns. Proper choice of parameters can ensure smooth classification in both cases. The chosen range of FD values are reported in Table 1. The basis of such choice of parameters is rigorous experimentation with the data. Similar continuous monitoring in case of the ANN based classifier would mean use of moving window technique of minimum overlap. That is obviously more time

Figure 4: Variation of possibilities of inversion and plume with time when size is 256.

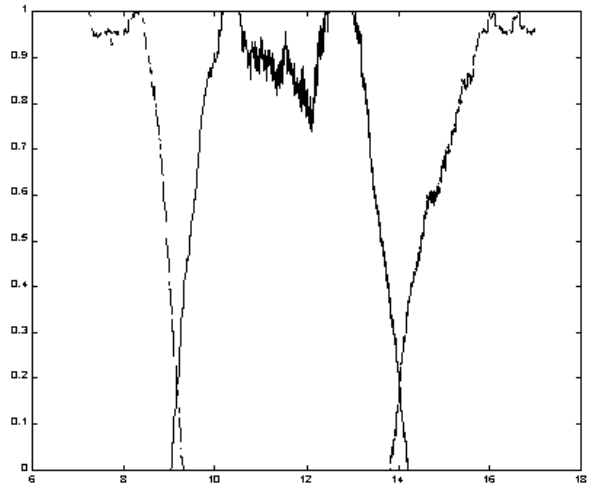
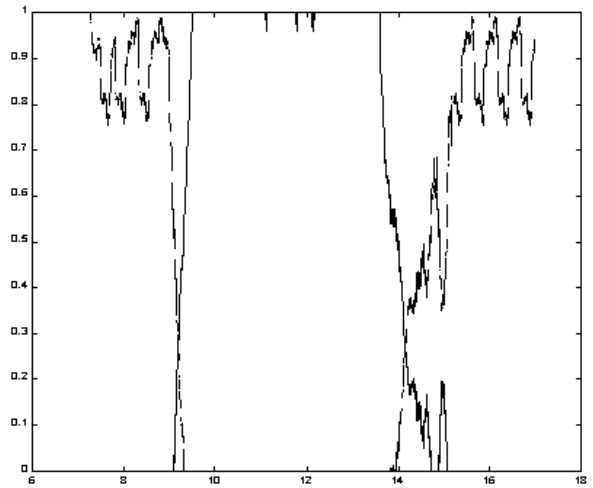


Figure 5: Variation of possibilities of inversion and plume with time when size is 64.



consuming and performs repetitive computations due to scale dependence of the relevant features.

6. Conclusion

The present work has successfully classified the sodar data using fractal features. One extension of the work is to consider multi-fractals. Another significant work is to include the spatial approach reported in [5]. Their approach remains significant for turbulence studies on lower atmosphere. A combination of the two approaches may therefore become significant for estimating parameters of pollutant dispersion model by accounting for effects of wind shear and thermal convection separately.

Table 1: FD ranges for different sodar patterns

Pattern	L (64)	U (64)	L (256)	U (256)
NS	1.20	1.45	1.25	1.50
FI	1.35	1.55	1.40	1.60
OI	1.50	1.70	1.55	1.75
PS	1.65	1.85	1.70	1.90

NS = No Structure, FI = Flat Inversion, OI = Other Inversion, PS = Plume Structure

U = Upper range, L = Lower range

References

[1] M. F. Barnsley. *Fractals Everywhere*. Morgan Kaufmann, second edition, 1993.

[2] D. W. Beran and F. F. Hall. Remote sensing for air pollution meteorology. *Bulletin of the American Meteorological Society*, 55:1097–1105, 1974.

[3] F. Beyrich. On the use of sodar data to estimate mixing height. *Journal of Applied Physics*, B57:27–35, 1993.

[4] A. K. Blackadar. Boundary layer wind maxima and their significance for the growth of nocturnal inversion. *Bulletin of the American Meteorological Society*, 38:283–290, 1957.

[5] A. Chanda, A. K. Dey, and J. Das. Fractal characterization of temperature field in convective boundary layer. *Fractals*, 5:267–274, 1997.

[6] B. B. Chaudhuri, A. K. De, A. Ganguly, and J. Das. Automatic recognition and interpretation of sodar records. *Indian Journal of Radio & Space Physics*, 21:123–128, 1992.

[7] B. B. Chaudhuri, A. Ganguli, A. K. De, and J. Das. Algorithm development for the machine recognition of sodar structure. In S. P. Singal, editor, *Acoustic Remote Sensing*, pages 155–160. TATA McGraw Hill Publication, New Delhi, 1990.

[8] A. K. De, S. Tripathy, and J. Das. On fine structure of dot echoes as observed by acoustic sounder. *International Journal of Remote Sensing*, 15(11):2157–2165, 1994.

[9] F. Einaudi and J. J. Finnigan. The interaction between an internal gravity wave and the planetary boundary layer. Part I. The linear analysis. *Quarterly Journal of the Royal Meteorological Society*, 107:793–806, 1981.

[10] J. J. Finnigan and F. Einaudi. The interaction between an internal gravity wave and the planetary boundary layer. Part II. Effect of the wave on the turbulence structure. *Quarterly Journal of the Royal Meteorological Society*, 107:807–832, 1981.

[11] S. Incecik. A study of daytime planetary boundary layer height with acoustic sounding and radiosonde profiles. *Bulletin of the Technical University of Istambul*, 42(1):99–110, 1989.

[12] M. A. Kallistratova. Physical grounds of acoustical remote sensing of the atmospheric boundary layer. In S. P. Singal, editor, *Acoustic Remote Sensing Application*, pages 3–34. Springer Verlag, 1997.

[13] R. P. Lippmann. Pattern classification using neural networks. *IEEE Communications Magazine*, 27:47–64, November 1989.

[14] A. Mukherjee, P. Pal, and J. Das. Identification of elementary sodar patterns using perceptrons. In *Proceedings of the Fourth International Conference on Advances in Pattern Recognition and Digital Techniques*, pages 83–86, Calcutta, India, December 1999. Narosa.

[15] L. Naiping, C. Jingnan, L. Shiming, Z. Y. S. Lirong, and Z. Mingyu. The detection of boundary layer structure with sodar in western mountain area of Bejing. In S. P. Singal, editor, *Acoustic Remote Sensing*, pages 345–356. TATA McGraw Hill Publication, New Delhi, 1990.

[16] F. T. M. Nieuwstadt. Some aspects of the turbulent stable boundary layer. *Boundary Layer Meteorology*, 30:31–55, 1984.

[17] F. T. M. Nieuwstadt. The turbulent structure of the stable, nocturnal boundary layer. *Journal of the Atmospheric Sciences*, 41:2202–2216, 1984.

[18] P. Pal, A. Mukherjee, S. Acharya, and J. Das. Continuous detection of atmospheric pattern from sodar signals. *Signal Processing*, 74(2):153–168, 1998.

[19] S. P. Singal. Current status of air quality related boundary layer meteorology studies using sodar. In S. P. Singal, editor, *Acoustic Remote Sensing*, pages 453–476. TATA McGraw Hill Publication, New Delhi, 1990.

[20] S. P. Singal. Monitoring air pollution related meteorology using sodar. *Journal of Applied Physics*, B57:65–82, 1993.

[21] A. J. Thrope and T. H. Guymer. The nocturnal jet. *Quarterly Journal of the Royal Meteorological Society*, 103:633–653, 1977.

[22] S. Tripathi, A. K. De, and J. Das. Computer analysis of atmospheric plume structures. *Indian Journal of Radio & Space Physics*, 21:321–328, 1992.

[23] S. Zhong, J. D. Fast, and X. Bian. A case study of strong summertime great plains low-level jet using a high resolution mesoscale model and data from wind profiler network. *Monthly Weather Review*, 124:785–806, 1996.