

# Using Optimized Distributional Parameters as Inputs in a Sequential Unsupervised and Supervised Modeling of Sunspots Data

K.Mwitondi<sup>1</sup>, J. Bugrien<sup>2</sup>, K. Wang<sup>3</sup>

<sup>1</sup>Sheffield Hallam University, Department of Computing; <sup>2</sup>Statistics Department, Benghazi University; <sup>3</sup>Radio and Space Services, Australian Bureau of Meteorology.

Email: k.mwitondi@shu.ac.uk, jamal.bugrien@ar.benghazi.edu.ly, K.Wang@bom.gov.au

Received June, 2013

## ABSTRACT

Detecting naturally arising structures in data is central to knowledge extraction from data. In most applications, the main challenge lies in the choice of the appropriate model for exploring the data features. Quite often, the choice is generally poorly understood and any tentative choice may be too restrictive. Growing volumes of data, disparate data sources and modelling techniques entail the need for model optimization via adaptability rather than comparability. We propose a novel two-stage algorithm to modelling continuous data consisting of an unsupervised stage whereby the algorithm searches through the data for optimal parameter values and a supervised stage that adapts the parameters for predictive modelling. The method is implemented on the sunspots data with inherently Gaussian distributional properties and assumed bi-modality. Optimal values separating high from low cycles are obtained via multiple simulations. Early patterns for each recorded cycle reveal that the first 3 years provide a sufficient basis for predicting the peak. Multiple Support Vector Machine runs using repeatedly improved data parameters show that the approach yields greater accuracy and reliability than conventional approaches and provides a good basis for model selection. Model reliability is established via multiple simulations of this type.

**Keywords:** Clustering; Data Mining; Density Estimation; EM Algorithm; Sunspots; Supervised Modelling; Support Vector Machines; Unsupervised Modelling

## 1. Introduction

Many real-life problems are tackled via knowledge extraction from data – a process typically associated with detecting naturally arising structures in the data. A typical example is the sunspots dataset [11] – an average oscillating sequence of the beginning and ending periods of solar cycles with an approximate periodicity of 11 years [7]. Recorded sunspots span across the first cycle (March 1755 to June 1766) to the first few months of the current (24th) cycle. Clustered in non-random positions above and below the equator, the spots are generated by interactions between the sun's surface plasma and its magnetic field [19 and 22]. Solar magnetic activity cycles have attracted the attention of scientists for many years. Solar flares, for instance, affect our planet in different ways - including ejecting plasma and energetic particles and potentially causing geomagnetic storms and damaging satellites [16]. The paper is motivated by the documented effects of sunspots on terrestrial conditions. Correlations between space and terrestrial weather have

been indicated in solar studies dating back many years [13, 18 and 20]. Climatic variations in Lapland via complex variations in the atmosphere, lunar gravitation and solar activity have also been explained [11].

This paper will be subjecting sunspots data to a sequential analysis involving unsupervised and supervised modeling. The two concepts represent the typical data mining problems – data clustering and classification. The primary goal of the former is to partition a given dataset with a known or unknown distribution into subgroups in such a way that data points in each group are as homogeneous as possible while those in different groups are as heterogeneous as possible. The method is typically applied in problems in which there is no clear mathematical formulation for describing the underlying structures. Various approaches to data clustering have been studied and are well-documented in the literature [21, 17 and 6].

However, determining the number of naturally arising structures in data remains a daunting challenge among

the data science community. Many clustering tools in the literature are based on the conventional mechanics of minimization of the distances between data points – a feature which inherently constitutes the same challenge the methods are designed to address – that is, determining the optimal number of clusters. The primary goal of the latter is to allocate new cases in known classes and one of its main challenges is balancing model accuracy and reliability.

Let a dataset of independently identically distributed random vectors  $\{x_1, x_2, \dots, x_{n-1}, x_n\} \in \mathbb{R}^d$  represent features an underlying density function. The main features of interest may include modes (local maxima), anti-modes (local minima) and *bumps* - regions where the second derivative is negative. In an exploratory setting, the number and locations of these features are not known *a priori*. Many real-life data take this form and with large volumes of data generated from different sources and inputted into different models, we are constantly faced with the challenge to determine optimal stationery points. The challenge is to address model complexity via adaptability rather than comparability. In other words, we seek to minimise inherent randomness in training and test data via novel adaptive methods of data analysis [10 and 1].

This paper proposes a novel approach to detecting naturally arising structures in data that searches for generalising parameter levels and adapts them to supervised modeling. Its main research problem is to develop an algorithm for predicting future cycles given historical solar activity data. We try to address this problem via the following objectives.

1) To determine naturally arising structures in the data. For simplicity, we shall be seeking to identify and separate high from low solar activity cycles. This objective constitutes the unsupervised stage of the algorithm.

2) To predict future cycles based on information in previous cycles. This is the supervised stage.

3) To search for an optimal solution based on repeated simulations at the unsupervised and supervised stages.

The paper is organised as follows. Section 1 provides the introduction followed by methods in Section 2. Data analyses and discussions are in Section 3 and concluding remarks and potential new directions in Section 4.

## 2. Methods

Choosing a parametric form of the density to explore features is generally poorly understood and any tentative choice may be too restrictive. Often under such circumstances non-parametric density estimation, e.g. Kernel Density Estimation (KDE) technique [21] allows for

practical solutions to the classical problem of choosing the level of smoothing (bandwidth), can be efficiently used. For example, given the data points  $\{x_1, x_2, \dots, x_{n-1}, x_n\} \in \mathbb{R}^d$ , the KDE approach to clustering defines clusters as regions of high density separated by regions of no or low density. Its main idea is to first compute a kernel density estimate,  $\hat{f}_t(x)$ , say, from the data, with a Gaussian kernel and isotropic bandwidth  $t > 0$  controlling the amount of smoothing. In its simplest form, KDE can be thought of as an alternative to the histogram as it typically provides a smoother representation of the data, and unlike the histogram, its appearance does not depend on a choice of starting point. The scenario represents a problem amenable to the multivariate kernel function in Equation 1 where  $T$  is a symmetric positive  $d$  by  $d$  bandwidth matrix defined as the diagonal  $T = \text{diag}[t_1, t_2, \dots, t_{n-1}, t_n]$  with a direct effect on model complexity.

$$\hat{f}_T(x) = n^{-1} \sum_{i=1}^n K_t(x - x_i) \quad (1)$$

Without loss of generality, consider a phenomenon with a binary structure of, say, “highs” and “lows”. Depending on the context, a number of models can be applied. For instance, if we assume a Gaussian kernel, we can define a parametric pattern of “lows” and “highs” in the form of a normal mixture model and use the parameter estimates  $\Theta = \{\mu, \Sigma\}$  to track the dynamics of the cycles. Further, if we assume that the probability of a “high” followed by another “high” structure is  $P_{hh}$  and that of a “low” followed by a “low” structure is  $P_{ll}$ , we can define a Hidden Markov Model as in **Table 1**. In this case, an HMM provides a formal foundation for linear sequence labeling of data. Balancing accuracy and reliability amounts to defining an appropriate way of labeling data using the probabilities and interpreting the results probabilistically. We could also define associations, the corresponding scores and the underlying confidence.

### 2.1. Data Description, Research Problem and Objectives

We adopt the sunspots data [11] – an average oscillating sequence of the beginning and ending periods of solar cycles forming the densities in **Figure 1**.

**Table 1. A state transition matrix for a binary structure.**

	HIGH	LOW
HIGH	$P_{hh}$	$1 - P_{hh}$
LOW	$1 - P_{ll}$	$P_{ll}$

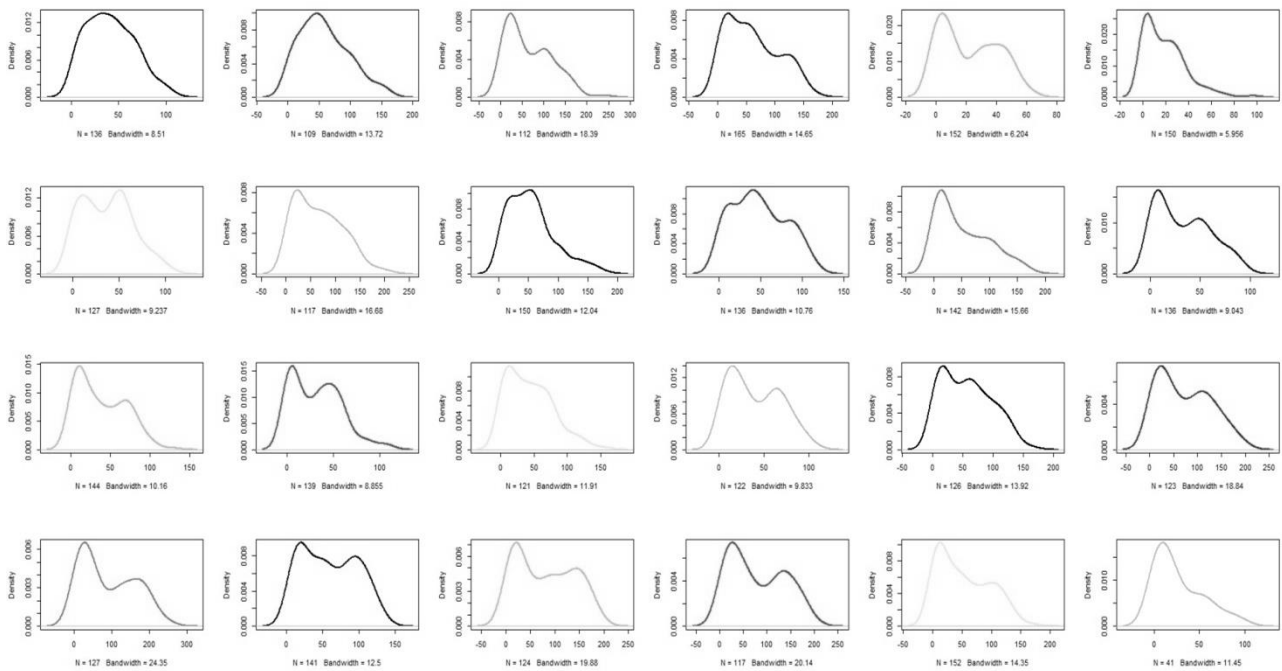


Figure 1. Density representation of the cycles in Figure 1.

The densities in **Figure 1** exhibit different number of modes – a feature typically determined by the adopted level of smoothing. By controlling the level of smoothing via a kernel function of the form in Equation 1 or otherwise we are able to identify different structures in data. **Figure 2** presents a 2-D plot of the sunspots means and standard deviations. The numbers in the plot represent the indices for each of the last 23 cycles and the current cycle (24<sup>th</sup>). Using a rule of thumb, we can identify the high, moderate and low solar activity cycles, say. Following [10 and 1] we can treat each cycle as a separate density and then use their distributional behavior to explore the underlying structures of the cycles.

### 2.2. Modelling Strategy

Conventional approaches to modeling sunspots include data assimilation [8] and rotational solar dynamo-based predictive models for short-term predictions [2 and 14]. The densities in **Figure 1** exhibit typically bivariate patterns and so we shall assume that the cycles form a parametric pattern of “lows” and “highs” and define

$$\prod_{i=1}^N \sum_{k=1}^K \pi_k f_k(S_i | \mu_k, \Sigma_k) \leftrightarrow \prod_{i=1}^N f_{S_i \in k}(S_i | \mu_{S_i \in k}, \Sigma_{S_i \in k}) \quad (2)$$

where  $S_i$  denotes the sunspots numbers,  $K$  is the number of components,  $f_{**}(\cdot)$  is a normal distribution,  $\pi_k$  is the prior probability of class membership and  $S_i \in k$  are class allocations. Statistically, the high-peaked (more than normal) and low-peaked (less than normal) cycles imply high and low solar activities respectively while

those skewed to the right imply few increases and frequent decreases in solar activity and vice versa. Our strategy involves two main levels – unsupervised and supervised. At the former level, we examine the initial and subsequent patterns of the cycles in order to separate the “lows” from the “highs”. The maximum likelihood estimates (MLEs) of the random finite mixture densities are estimated and passed on to a predictive model at the supervised level as outlined below.

```

Begin
  Let  $k \in \{1, h\}$ 
  Cluster  $S_{i=1,2,\dots,N}$  into finite groups, say,  $S_{1:l}$  and  $S_{h:l}$ :  $l$  are lows and  $h$  are highs
  Define  $\theta_{1,h} = \{\mu_{1,h} \text{ and } \Sigma_{1,h}\}$ 
   $\omega = \arg \max_S \lambda \mu_h := \{\forall S_i | \mu_i \ll \mu_h : \lambda \text{ a predefined constant}$ 
  Define  $\theta_{1,h} = \{\mu_{S_{1:l,h}(l;n)} \text{ and } \Sigma_{S_{1:l,h}(l;n)}\} : n = \{n_l, n_h\}$  are cases in each cluster
  Set  $\theta = \{\theta_{1,h}, \theta_{1,h}\}$ 
  Initialise  $z_i = (0, 1)^2 = (z_{1k}, z_{hk})$  so that  $z_{1k} = \begin{cases} 1 & \text{if } (S_{i,l,h}(l;n)} \in k) > \omega \\ 0 & \text{Otherwise} \end{cases}$ 
  Obtain MLE of  $f_{S_i \in k}(S_i | \theta_{1,h}, \theta_{1,h})$ 
  For  $m := 1$  to  $M$  (Large positive integer)
     $\hat{\mu}_k := \frac{\sum_{i=1}^N z_{ik} S_i}{\sum_{i=1}^N z_{ik}}$ ;  $\hat{\mu}_k := \frac{\sum_{i=1}^N z_{ik} S_i}{\sum_{i=1}^N z_{ik}}$  and  $\hat{\sigma}_k = \sqrt{\frac{\sum_{i=1}^N z_{ik} (S_i - \hat{\mu}_k)^2}{\sum_{i=1}^N z_{ik}}}$ 
    Update  $\theta := \theta[m]$ 
    Update  $z_{ik} := z_{ik}[m]$ 
    Update MLE of  $f_{S_i \in k}(S_i | \theta) := f_{S_i \in k}(S_i | \theta)[m]$ 
  End For
  Select best model
  Use the model to predict new cycles  $S_{i>N}$ 
  Output best model parameters
  Determine whether  $S_{i>N} \in S_l$  or  $S_{i>N} \in S_h$  based on, say,  $\psi = \frac{\sum_{i=1}^N 1_{\{z_{1k} \leq z_{hk}\}}}{N} \propto \hat{\mu}_k$ 
End
    
```

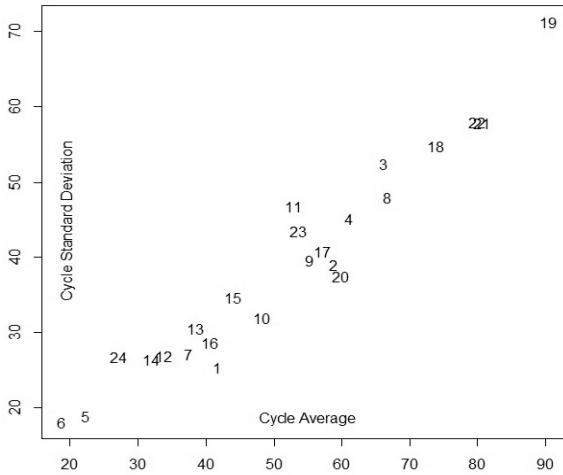


Figure 2. Sunspot means and standard deviations.

The above algorithm adapts the EM converging features described in [5 and 9]. Its form suits any supervised modelling technique. In this paper it is implemented in Support Vector Machines (SVM).

### 2.3. Supervised Modeling of Labeled Data

We adopt Support Vector Machines (SVM) - a kernel-based discriminant function the mechanics of which rely on supervised learning of the underlying discriminating rules from the training data [5]. To put it in context, let the “high” and “low” cycles in our modified set  $\{S_i, y_i : i=1, \dots, N\}$ ,  $y_i \in \{-1, 1\}$  and  $S_i \in \mathbb{R}^2$  be separable by the hyper-plane  $H$ . Then the points lying on  $H$  satisfy the equation  $wS+a=0$

where  $w$  is normal to the hyper-plane,  $\frac{|a|}{\|w\|}$  is the perpendicular distance from  $H$  to the origin and  $\|w\|$  is the Euclidean norm of  $w$ . The points on the hyper-planes above and below  $H \{H1, H2\}$  will satisfy the equations  $wS+a=\pm 1$  (both with normal  $w$  and distance to the original  $\frac{|\pm 1 - a|}{\|w\|}$ ) which means that the gap  $\{H1, H2\}$

$= \frac{|2|}{\|w\|}$ . We need to find hyper-planes maximising the gap (minimizing  $\|w\|^2$ ) subject to  $y_i (S_i w + a - 1) \geq 0$ . The numbers in  $\{H1, H2\}$  are the support vectors (supporters) of the optimal location of the decision surface and the hardest to classify. Intuitively, the allocation rule is

$$\begin{cases} S_i w + a \geq +1 \text{ for } y_i = +1 \\ S_i w + a - 1 \leq \text{for } y_i = -1 \end{cases} \leftrightarrow y_i (S_i w + a - 1) \geq 0 \quad \forall_i \quad (3)$$

The SVM kernel [4] is generally defined as

$$F(S) = \sum_{i=1}^V \alpha_i \Phi(S, S_i) + a \quad (4)$$

in which  $\alpha_i$  represents the Lagrange multiplier summed over the values for which  $\alpha_i > 0$ . The upper index  $V$  denotes the number of support vectors as described above. SVM solution relies on the Lagrangian formulation of the problem – an optimisation method requiring  $V \in \mathbb{N}$  positive multipliers ( $\alpha_{i=1,2,\dots,V}$ ) for each of the inequalities on the RHS of Equation 3. The general formulation of the Lagrangian is

$$L = \frac{w^2}{2} - \sum_{i=1}^V \alpha_i y_i (S_i w + a - 1) + \sum_{i=1}^V \alpha_i \quad (5)$$

SVM solution is obtained by minimising Equation 5 with respect to  $w$  and  $a$  and simultaneously requiring that  $\frac{dL}{d\alpha_i} = 0 \forall_i$  or equivalently maximising  $L$  and require that

Both  $w$  and  $a$  disappear. The latter implies that  $w = \sum_i \alpha_i y_i S_i$  and  $\sum_i \alpha_i y_i = 0$  transforming Equation 5

into its dual equivalent  $L_d = \sum_i \alpha_i - \frac{1}{2} \sum_i \alpha_i \alpha_j y_i y_j S_i \cdot S_j$ .

The SVM model weights are calculated as the product of the support vector coefficients and their values and used in forming the allocation rule. Other than the support vectors ( $\alpha_i > 0$ ) the remaining data points have  $\alpha_i = 0$  – these are those lying on the two hyper-planes  $\{H1, H2\} \rightarrow y_i (S_i w + a - 1 = 0)$  or beyond them if  $y_i (S_i w + a - 1 > 0)$ .

## 3. Analyses and Discussions

We now present the two-level analyses described above in order to establish whether sunspots follow identifiable patterns which can be used as inputs in a predictive model.

### 3.1. Unsupervised: Initial Patterns and Maximisation

Figure 3 exhibits the low and high cycles separation based on the cut-off points above alongside their corresponding overall bi-modal densities. It is based on the maximum number of sun spots reached by the full cycles and the number reached in the first 30 and 40 months. The cut-off point in the LHS panel is set to the mean of the averaged maximum early sun spots which, in this case, is 109 - separating the low cycles 1, 5, 6, 7, 9, 10, 12, 13, 14 and 16 from the highs 2, 3, 4, 8, 11, 15, 17, 18, 19, 20, 21, 22 and 23. The densities in the RHS panel exhibit the emerging bi-modality as a function of time.

Figure 3 suggests that the pattern of each solar activity

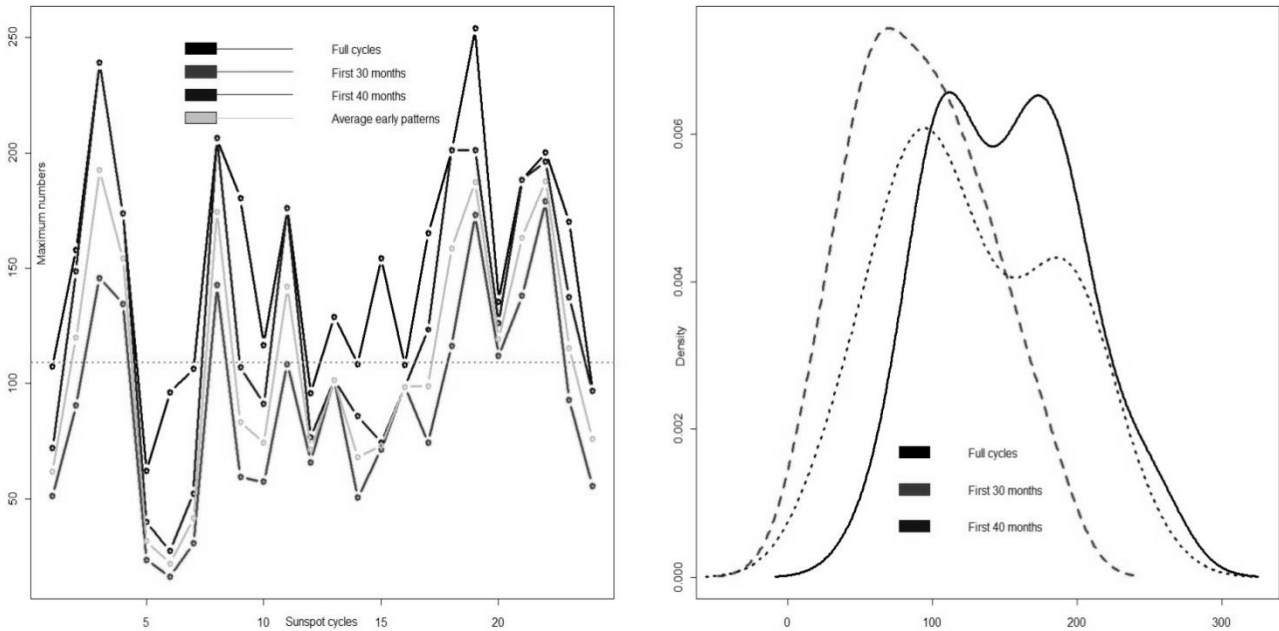


Figure 3. Omega cut-off (LHS) and the corresponding bi-modal density (RHS).

cycle is defined by its early patterns. In particular, the maximum values reached by each cycle appear to provide an insight into the overall activity of the cycle before it starts to subside. The foregoing structural detection of patterns in the sunspots data amounts to unsupervised modelling. Adopting these patterns as a guide to data labelling rule yields the two class priors as

$$\hat{\pi}_l = \frac{\sum_{i=1}^N z_{il}}{N} = 0.46 \leftrightarrow \hat{\pi}_h = 0.54 \text{ with } \theta_{l,h} \text{ computed as}$$

above. As the average early patterns for cycle 24 fall below the cut-off point, it is reasonable to suggest that it will be a low activity cycle. Implementation of SVM modelling follows below.

### 3.2. Supervised Level: SVM Supervised Modelling

Results from SVM modelling based on the initial class patterns with prior probabilities  $\hat{\pi}_l$  and  $\hat{\pi}_h$  gave an averaged accuracy of 58% on a cost range of 0.005 to 5 and a training sample of 500. Posterior class probabilities conditioned on maximised averages of the early low and high group means reached an average accuracy of 98% on the same cost range and training sample size. The support vectors are shown in **Figure 4** with the horizontal and vertical axes corresponding to the support vectors and indices respectively.

**Figure 5** shows the best discriminating SVM decision values at two different bandwidths. The bandwidths and hence decision values are chosen from multiple simulations as determined by the binary cut-off point demar-

cating low from high cycles. Notice how each of the modes also exhibits sub-modes To avoid spurious modes (over-fitting) or masking effects (under-fitting) it is recommended to use significance test for changes or, for clear patterns, graphical visualization.

Typically, SVM model weights for each of the support vectors are obtained as a cross product of the model coefficients and support vectors [15]. Weights from multiple SVM runs can be recorded and their graphical patterns be used to guide model selection. Other SVM outputs include the individual probabilities and decision values as in **Figure 6**. The difference between the lower accuracy case in the top panel - highlighting the random nature of class allocation – and the higher accuracy model in the bottom panel) showing clear concentrations of  $\hat{\pi}_l$  and  $\hat{\pi}_h$  on either side of the class boundary. The observations corresponding to the vectors in **Figure 4**, the decision values in **Figure 5** and to the corresponding probabilities in **Figure 6** can be identified by indexing.

### 4. Concluding Remarks and Potential Future Directions

Predicting solar activity cycles remains one of the major challenges the scientific community faces with intricacy being compared to predicting, say, the severity of next year’s winter. In this weather analogue, if all that is available is a long vector of temperature readings over many years, the only sensible approach is to search for

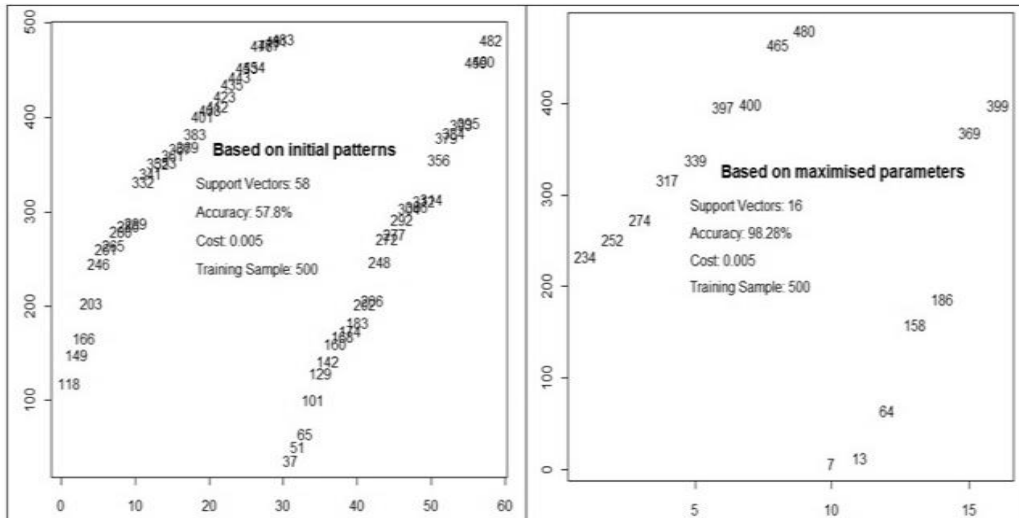


Figure 4. Support vectors for the initial patterns (LHS) and maximised parameters (RHS).

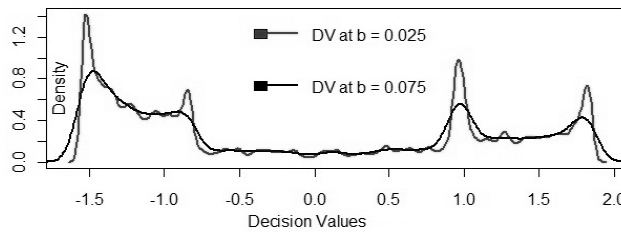


Figure 5. SVM decision values at two different bandwidths showing well-separated structures.

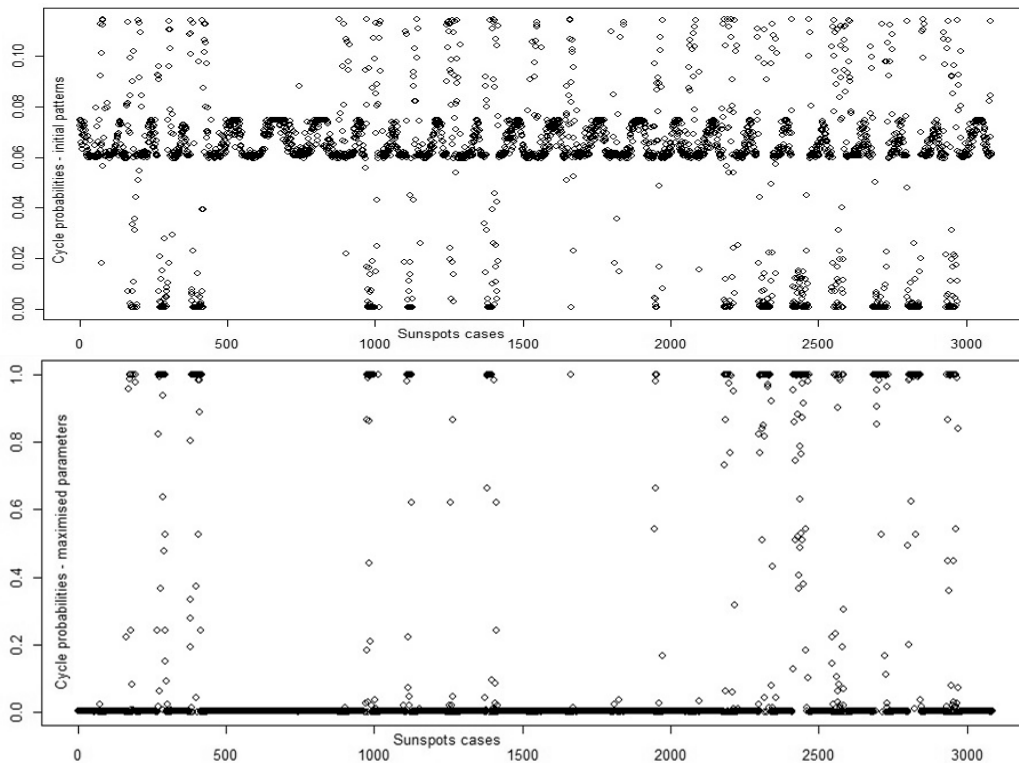


Figure6. Sun cycles class probabilities.

naturally arising structures in the data with the hope that if uncovered they may provide potentially useful information. This paper adopted the foregoing philosophy and sought to develop a predictive framework for modelling sunspots data using inherent distributional properties in the data. The paper relied on a continuous flow of data for prediction, but rather than assessing model accuracy on the NOAA benchmark, an SVM model was trained and tested on a notionally infinite dataset of cycles.

By examining multiple sets of observations from the onset of each cycle via graphical visualisation early patterns of sun cycles and their binary nature were determined. Comparing multiple early patterns for each recorded cycle extracted at different time periods to the corresponding full cycles revealed that the first 3 years provide a sufficient basis for predicting the cycle's peak. The patterns were then adapted as inputs into an integrated unsupervised and supervised modelling algorithm. The novel method's mechanics are geared towards simultaneously tracing anomalies via an adaptive approach. Repeated SVM runs using repeatedly improved parameters showed that the approach yields greater accuracy and reliability than conventional approaches. Multiple simulations of this type can be generated based on the algorithm above to assist in selecting the most consistent model. The paper's main substance can be described as an enhancement of algorithmic methods for learning underlying rules from data.

Finally, it is worth noting that while the study was confined to the conventional periodicity of 11.11 years [22] with a binary pattern of cycles, the definition implies that the periodicities can differ according to definitions. Further, while we assumed a binary scenario of the cycles in **Figure 1**, different bandwidths are likely to yield different patterns. To address this limitation, the paper's findings highlight potential investigations paths into such variations. Further, the current study, based on a single application and a single method, could not confirm the algorithm's robustness. Although we adopted SVM for implementation, the approach is amenable to any domain-partitioning method. Thus, for model enhancement purposes, it will be useful to provide a comparative study using other learning algorithms such as neural networks and decision trees.

## REFERENCES

- [1] Bugrien, J., Mwitondi, K. and Shuweihi, F. (2013). A Kernel Density Smoothing Method for Determining an Optimal Number of Clusters in Continuous Data; The 16th International Conference on Computational Methods and Experimental Measurements; 2 – 4 July, 2013, A Coruña, Spain.
- [2] Choudhuri, A. R., Chatterjee, P. and Jiang, J. (2007). Predicting Solar Cycle 24 with a Solar Dynamo Model; *Physical Review Letters*, Vol. 98, Issue 13, American Phys. Society.
- [3] Cuevas, A., Febrero, M. and Fraiman, R. (2000). Estimating the number of clusters; *The Canadian Journal of Statistics*, Vol. 28, No. 2, pp 367-382.
- [4] Cortes and Vapnik, (1995). Support-vector networks; *Machine Learning*, Vol. 20, No. 3, pp. 273-297, Kluwer Academic Publishers.
- [5] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm; *Journal of the Royal Statistical Society*, Vol. 39, pages 1-38.
- [6] Hand, D., Mannila, H. and Smyth, P. (2001). Principles of Data Mining (Adaptive Computation and Machine Learning); A Bradford Book; ISBN-13: 978-0262082907.
- [7] Kane, R. P. (2007). Solar Cycle Predictions Based on Extrapolation of Spectral Components: An Update; *A Journal for Solar and Solar-Stellar Research and the Study of Solar Terrestrial Physics*; Vol. 246, Issue 2, pp 487-493, ISSN: 0038-0938.
- [8] Kitiashvili, I. and Kosovichev, A. (2009). Prediction of solar magnetic cycles by a data assimilation method; *Cosmic Magnetic Fields: From Planets, to Stars and Galaxies*; Proceedings IAU Symposium, No. 259, Edited by Strassmeier, K, Kosovichev, A. and Beckman, J. (2009) - International Astronomical Union.
- [9] McLachlan, G. Krishnan, T. (1996). The EM Algorithm and Extensions; John Wiley.
- [10] Mwitondi, K., Said, R. and Yousif, A.: A sequential data mining method for modelling solar magnetic cycles; *Neural Information Processing, LNCS*, Vol. 7663, pp 296-304, Springer (2012).
- [11] NOAA (2012). <http://www.ngdc.noaa.gov>
- [12] Pohtila, E. (1980). Climatic fluctuations and forestry in Lapland; *Ecography*, Vol. 3, Issue 2, pp 65–136, ISSN: 1600-0587.
- [13] Pielke, R., Avissar, R., Raupach, M., Dolman, A., Zeng, X. and Denning, A. (1998). Interactions between the atmosphere and terrestrial ecosystems: Influence on weather and climate; *Global Change Biology*, Vol 4, Issue 5, pp 461–475.
- [14] Qahwaji, R. and Colak, T. (2007). Automatic Short-Term Solar Flare Prediction Using Machine Learning and Sunspot Associations; *SOLAR PHYSICS*, Vol. 241, No. 1, pp 195-211, ISBN 11207-006-0272-5.
- [15] R (2011). R Version 2.13.0 for Windows; R Foundation for Statistical Computing.
- [16] Reames, D. (2002). Magnetic topology of impulsive and gradual solar energetic particle events; *The Astrophysical Journal*, Vol. 571, pp 63–66.

- [17] Roberts, S. J.: Parametric and Non-parametric Un-supervised Cluster Analysis, *Pattern Recognition*, 30, 5, pp 261-272 (1997).
- [18] Rycroft, M. J, Israelsson, S. and Price, C. (2000). The global atmospheric electric circuit, solar activity and climate change; *Journal of Atmospheric and Solar-Terrestrial Physics*, Vol. 62, Issues 17–18, pp 1563–1576.
- [19] Schwabe, S.H. (1843). *Astronomische Nachrichten*, 20, No. 495, 234-235
- [20] Siscoe, G. L. (1978). Solar–terrestrial influences on weather and climate; *Climatology Supplement, Nature*, Vol. 276, pp 348-352.
- [21] Silverman, B. W. (1981). Using Kernel Density Estimates to Investigate Multimodality, *Journal of the Royal Statistical Society, B*, 43, pp 97-99.
- [22] Wolf, J. R. (1852). New studies of the period of Sunspots and their meanings; *Communications of Natural History; Society in Bern*, 255, pp 249-270.