Proceedings of the 9th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'23)
Brunel University, London, United Kingdom - August 03-05, 2023
Paper No. MVML 104
DOI: 10.11159/mvml23.104

Anomalous Signal Characterization Using Kalman Filter-Based Spectral Quantification and Bayesian Statistical Diagnostics

Nicholas V. Scott
Riverside Research Institute, Open Innovation Center, Dayton Research Center
2640 Hibiscus Way, Beavercreek, Ohio, 45431, USA
nscott@riversideresearch.org

Preprocessed electromagnetic signals in the form of frequency spectral groups are constantly acquired from electro-optical platforms where anomalous frequency structure is often hidden. These anomalies can be detrimental to trusted systems holding important information. There is a need to obtain vital anomalous behavioural statistical information which can be transformed into empirically driven predictive models for pattern–of-life estimation supporting trusted system protection. A two-tier algorithmic approach is employed to accomplish this using Kalman filter-based spectral quantification and Bayesian modelling. Kalman filter-based spectral deviation quantification is developed to estimate the average spectral deviation for an ensemble of frequency spectra comprising a series of spectral groups. The spectral-band based quantification of anomalous spectral group behavior over time supports the development of second stage algorithms aimed at parameterizing the statistical structure of the average spectral deviation as a random process. Algorithms here are based on Bayesian statistical estimation of mean and variance of sequentially estimated spectral deviation values for spectral groups, the application of analysis of variance (ANOVA) to mean spectral deviation values, and Markovian modelling of the mean and variance of spectral deviation values.

The development of algorithms supporting anomalous spectral deviation quantification and Bayesian diagnostics is based on the analysis of hyperspectral imagery (HSI) data. A HSI cube of land and buildings was broken up into fourteen 100 X 100 pixel image chips which were used as a generator of frequency spectral groups. A typical HSI pixel spectral signature for dirt with added noise was extracted from the first image chip representing the mode of the data set and used in a Euclidean metric for measurement of anomalous spectra within image chips. HSI spectral signatures exceeding the metric threshold of 0.2 were flagged as anomalous spectra for each image chip and Kalman filtration used to characterize flagged spectral signals in each image chip. The Kalman filter applied to frequency spectra uses a series of frequency spectral energy measurements over a finite bandwidth containing random noise to produce a statistically optimal estimate of spectral band deviation from the mode spectral signal. The spectral deviation energy estimated from the filter was averaged over the full spectral bandwidth of a flagged spectrum and then over each image chip. The array of 14 image chips were then cycled over 20 times providing a 280-point average spectral deviation time series. For large numbers of frequency spectra in a single image chip, the average spectral deviation has a probability distribution that is log normal in shape. This is not surprising given that energy deviation is what is measured by the Kalman filter algorithm.

Further insight, corroboration, and modelling of the statistical generation process for spectral deviation change was accomplished using analysis of variance (ANOVA) and Bayesian statistical modelling. Twenty F-statistic values for groups of fourteen image chips comprising the data domain were all close to 1 corroborating that the average spectral deviation values for all image chips or spectral groups do emanate from the same statistical process. Bayesian recursive estimation of the mean and variance for the average spectral deviation associated with each image chip was performed to produce a 280-point time series for each of these quantities. Periodicity of the mean spectral deviation was evident along with changes in the uncertainty intervals suggesting possible statistical structure linking mean and variance. Hidden Markov modelling, where the average spectral deviation is the state variable and the accompany variance is the observation variable, was performed to explore this idea. Preliminary analysis suggests that based on limited data, high anomalous spectral deviation structure tends to have a lower uncertainty which is useful information for the synthesis of anomalous spectral behavior characteristic of this underlying random process.

Keywords: anomalous structure, ANOVA, average spectral deviation, Bayesian recursive estimation, emission matrix, hidden Markov modelling, hyperspectral imagery, image chips, Kalman filter, spectral groups, transition matrix

1. Introduction

Pre-processed electrooptical signals in the form of frequency spectral groups are constantly acquired from electro-optical (EO) platforms where anomalous frequency structure is often hidden. This latent structure can be detrimental to systems holding important information. There is a need to obtain vital statistical information about anomalous behavior which can be transformed into empirically driven predictive models for pattern–of-life estimation. A Kalman filter-based spectral deviation quantification algorithm is applied to estimate the average spectral deviation, a quantification of anomalous behavior, for an ensemble of frequency spectra comprising spectral groups. This is the first step in addressing this need. Along with spectral-band based quantification of anomalous behavior, there is also a desire to understand how anomalous behavior, in the form of average spectral deviation, changes over time. To support this, algorithms are developed to model the structural change in average spectral deviation time series. Algorithms here are based on Bayesian statistical estimation of mean and variance of sequentially estimated spectral deviation values for spectral groups, the application of analysis of variance (ANOVA) to mean spectral deviation values, and Markovian modelling of mean and variance of the average spectral deviation values.

2. Data Structure and Methodology

A hyperspectral imagery cube, shown in Fig. 1a, broken up into a series of fourteen 100 X 100 pixel image chips was used as an ensemble frequency spectrum generator. This HSI data comes from the Rochester Institute of Technology (RIT) hyperspectral imagery data set which was part of the Target Detection Blind Test Project [1]. A typical HSI spectral signature for dirt with added noise was extracted from the first image chip (labeled as α_1 in Fig. 1b) representing the mode of the data set.

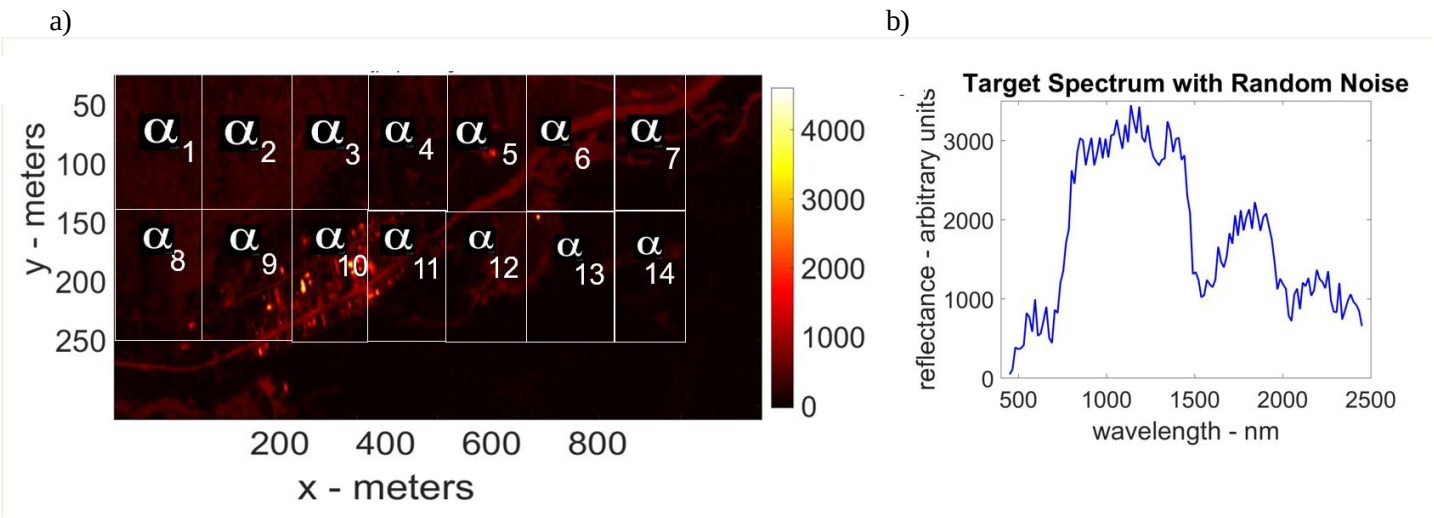


Figure 1: a) RIT hyperspectral imagery (HSI) cube from the Target Detection Blind Test project which is used to generate frequency spectral groups. Image chips denoted by white boxes represent demarcations of spectral groups with labels representing average spectral deviation values associated with the spectral groups. b) HSI signal for dirt representing the mode for the entire HSI cube used in the Euclidean metric to flag anomalous spectral signals in image chip spectral groups. Random noise added to spectrum.

It is used in a Euclidean metric for measurement of anomalous spectra within image chips. HSI spectral signatures exceeding the metric threshold of 0.2 were flagged as anomalous spectra for the image chip spectral group.

Kalman filtration was used to characterize flagged spectral signals in each image chip. The Kalman filter uses a series of frequency spectral energy measurements over a finite bandwidth containing random noise to produce a statistically optimal estimate of spectral band deviation from the mode spectral signal [2]. The estimated state variable is the noise filtered spectral band deviation obtained using a predictor-corrector formulism. E.g. The filter estimates the spectral deviation at some frequency (the predictor) and obtains feedback in the form of noisy observations which updates these values (the corrector). The Kalman filter-

based spectral quantifier assumes the full bandwidth spectral energy deviation can be modeled as a Markovian process [3]. E.g. Only the estimated spectral abundance or deviation from the previous frequency band and the current spectral deviation measurement value are needed to compute the current spectral deviation estimate. The Kalman filter has two parameters which influence estimation of the spectral abundance. The process noise for the Kalman filter spectral quantifier was set to 0 and the measurement noise set to 0.1 [2].

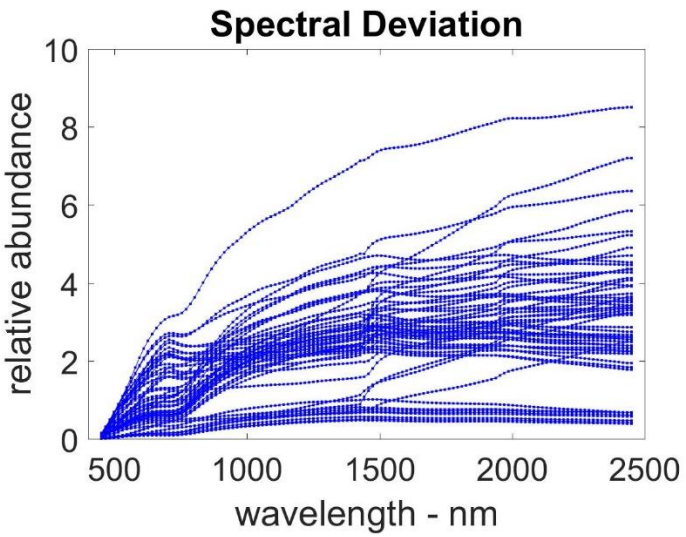
3. Kalman Filter-Based Spectral Quantification and Bayesian Statistical Diagnostics Results

Examples of the spectral deviation for the first image chip spectral group are shown in Fig. 2a. and Fig. 2b. The spectral deviation energy is averaged over the full spectral bandwidth providing an average spectral deviation time series shown in Fig. 2c. For extremely large N, where N is the number of pixel spectrum realizations, the average spectral deviation time series has a probabilistic distribution that is log normal in shape. This is not surprising given that energy deviation is what is measured by the Kalman filter algorithm.

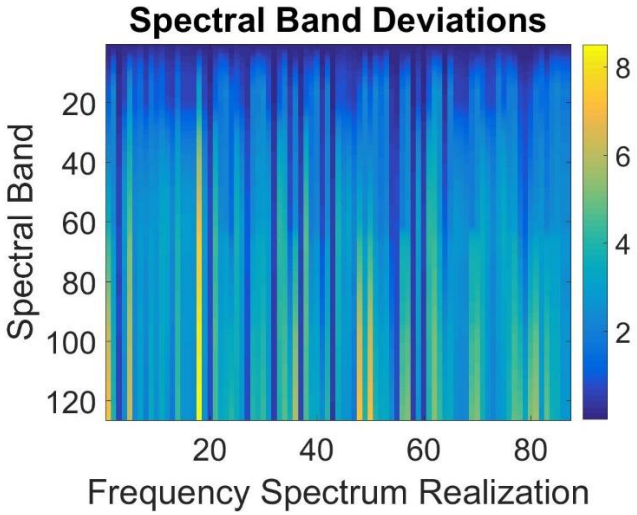
Average spectral deviation time series were calculated using the 14 image chips which were cycled over 20 times. The resulting 280-point average spectral deviation time series was calculated using the same process carried out on the first image chip for the series of image chips shown in Fig. 1a. The statistical generation process for state spectral deviation change was addressed using analysis of variance (ANOVA) [4]. Twenty F-statistic values for each of the 14 image chips were all close to 1 suggesting that the average spectral deviation values for the image chips do not deviate much from each other. This is consistent with the spectral deviations all emanating from the same statistical process.

Bayesian recursive estimation of the mean and variance for the average spectral deviation associated with each image chip over the 20 cycles is shown in Fig. 2d. [5,6]. This was done after taking the exponential of the mean spectral deviation for each image chip. The time series shows an anticipated 14-image chip periodicity for the average spectral deviation with accompanying variance bounds. Mean values and errors bounds fluctuate erratically and it is instructive to examine if there is a pattern existing within the time series that is not immediately evident. This can be accomplished using hidden Markov modelling.

a)



b)



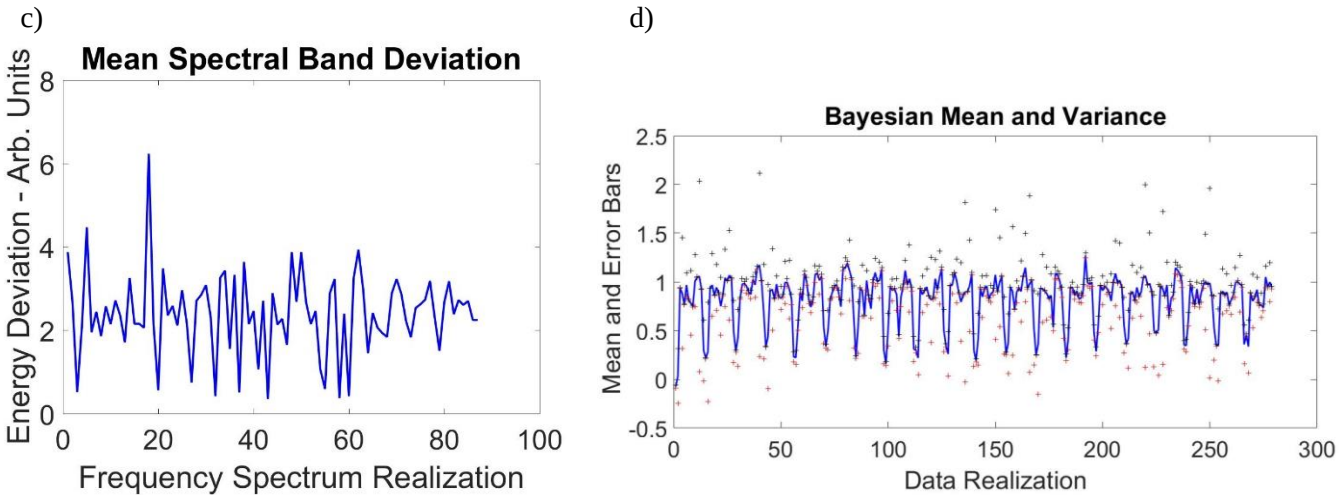


Figure 2: a) Spectral deviation values taken over a finite bandwidth for flagged anomalous spectra in the first image chip. b) Same spectral deviation values plotted as an image plot. c) Spectral deviation values averaged over the complete spectral bandwidth for each HSI anomalous spectral signal in the first image chip. d) Bayesian recursive estimation of mean and variance assuming a normal distribution. Red plus signs signify the 95% confidence-based lower error bar and black plus signs signify the upper error bar. Blue curve is the mean spectral deviation.

Hidden Markov modelling of the 280-point Bayesian mean spectral deviation and variance data series provides transition structural insight into the relationship of the average spectral deviation state variable and the accompany variance observation variable [6]. The transition matrix (Fig. 3a) shows that the mean spectral deviation value intervals, spanning the range of 0.63-1.28 units (blocks 3-6 on the y-axis), tends to transition into the interval of 0.84-1.1 units (block 5 on the x-axis) while the mean value intervals spanning 0.21-0.63 units (blocks 2 and 3 on the y-axis) tend to transition into itself. Overall, this structure suggests that moderate and high mean spectral deviation values transition into high mean spectral values. On the other hand, low mean spectral values tend to transition into low mean spectral values. This pattern of medium and high spectral mean values transitioning into high mean spectral values is a distinctive characteristic of this process which is not observed in many random dynamic variables such as turbulent kinetic energy in fluid turbulence [7]. It is noted that land-based energy changes are not bound by any dynamical energy constraint, allowing for the observed spectral dynamics.

The emission matrix has a complete mean spectral deviation value interval spanning the range of 0-1.28 units for the state transition variable and a complete observation variable range for variance spanning 0.0006-0.98 units. The emission matrix (Fig. 3b) results show that most medium to high mean spectral deviation intervals (blocks 4-6 on the y-axis) tend to emit associated variances in the lowest variance range spanning 0.0006-0.16 units (block 1 on the x-axis). The lowest mean spectral deviation interval (block 1 with an interval of 0-0.21 units) as well as the third deviation interval (block 3 with an interval of 0.42-0.63 units) tend to have variances spanning the first 2 blocks on the x-axis (0.0006-0.33 units). High anomalous spectral deviations tend to have low uncertainty or variance suggesting that when spectral deviations are large they are significant. Large deviations are often associated with noise, making this finding atypical but useful in the synthesis of anomalous spectral behavior for this random process.

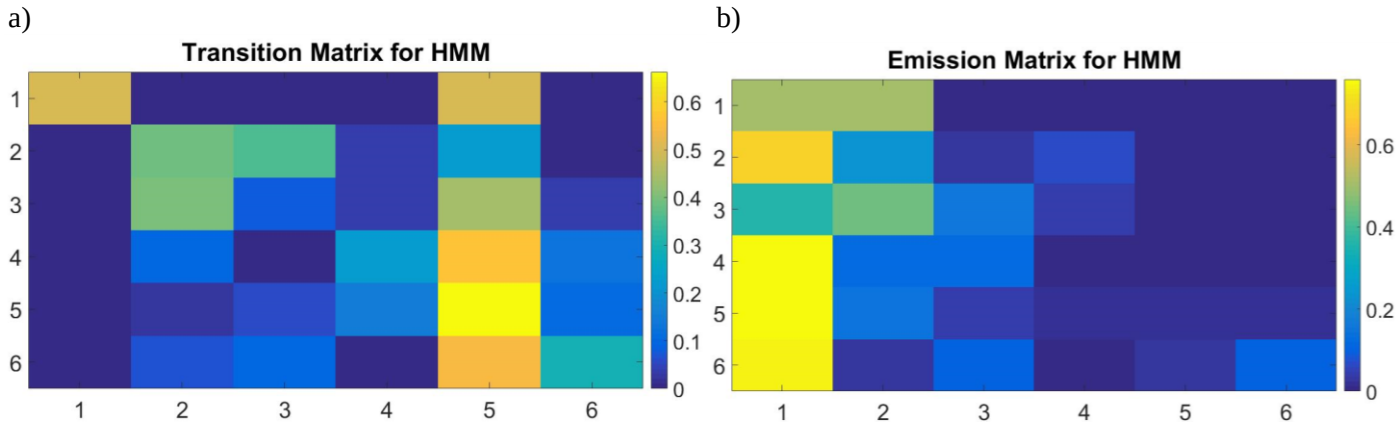


Figure 3: a) Transition matrix for average spectral deviation time series. The 6-transition intervals span 0-1.28 units with equidistant intervals of 0.21 units. b) The emission matrix for mean spectral deviation interval state and the emitted variance observation. It also has 6-transition intervals which span 0-1.28 units with equidistant intervals of 0.21 units and 6-emission observation intervals which span the interval of 0.0006-0.98 units with equidistant intervals of 0.16 units. Units are arbitrary.

4. Conclusions

A Kalman filtration-based signal processing formulism is demonstrated for the characterization of anomalous spectral signals existing in spectral groups. It is a generator of mean spectral deviations and variance values for frequency spectral groups which are in turn used in the characterization of the underlying statistical spectral deviation process. The application of ANOVA and hidden Markov modelling are shown to be not only different ways of characterizing the underlying random process responsible for spectral group observations, but also are the starting point for robust predictive power. The illustrated algorithms are meant to be the first steps towards the construction of geo-intelligence processors operating in a full geo-intelligence processing pipeline for processing of data streams containing anomalies. It is conjectured that an ensemble of developed processors, such as these, represent a viable way for performing statistical signal modelling and characterization for establishing statistical priors for neural network-based processing. This is because data hungry, deep learning methods are often in dire need of guidance in making robust function approximations which can be furnished by the type of statistical processors illustrated here.

References

[1] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a Web-based Application to Evaluate Target Finding Algorithms," *Proceedings of the 2008 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2, pp. 915-918, Boston, MA, 2008.

[2] C. -I. Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*. Hoboken, NJ: John Wiley and Sons, 2013.

[3] G. A. Fink, *Markov Models for Pattern Recognition: From Theory to Applications*. London, UK: Springer-Verlag, 2014.

[4] F.M. Dekking, C. Kraaikamp, H.P. Lopuhaä, and L.E. Meester, *A Modern Introduction to Probability and Statistics: Understanding Why and How*. London, UK: Springer: Verlag, 2005.

[5] W. M. Bolstad, *Introduction to Bayesian Statistics*. Hoboken, NJ: John Wiley and Sons, 2007

[6] O. Ibe, *Markov Processes for Stochastic Modeling*, Boston, MA: Elsevier Academic Press, 2013.

[7] Davidson, P. A., *Turbulence: An Introduction for Scientists and Engineers*. London, UK: Oxford University Press, 2015.