MOMENT-BASED GOODNESS-OF-FIT TESTS FOR POLARIMETRIC RADAR DATA

Stian Normann Anfinsen, Torbjørn Eltoft, and Anthony Doulgeris

Department of Physics and Technology, University of Tromsø

1. INTRODUCTION

The advent of polarimetric synthetic aperture radar (PolSAR) data has spurred a growing interest in statistical models for complex covariance matrices, which is the common representation of multilooked radar data. In this paper, we respond to an emergent need by proposing a set of goodness-of-fit tests for certain complex matrix distributions, and use these to compare radar data models with observations. We evaluate the strength of our tests with simulated data, and present test results for data acquired with several airborne PolSAR instruments.

2. STATISTICAL MODELS OF POLSAR DATA

The complex Wishart distribution [1] was the first model proposed for multilooked polarimetric radar data, and is still the most common, largely due to its mathematical tractability. It provides a simplistic analysis based on the assumption of Gaussian statistics for the complex scattering coefficients. Its shortcomings has been amended with alternative compound distributions, such as the polarimetric K distribution [2], G_0 distribution [3], and others [4] that account for non-Gaussianity and thus provide a more realistic model for high resolution data covering inhomogeneous areas and pronounced texture.

3. GOODNESS-OF-FIT TESTS

Most studies where new models are proposed justify them only by visually comparing fitted model densities to histograms of empirical data for a single polarimetric channel at the time. One exception is [4], where the likelihood function was used to assess model fit for a set of different compound models. The use of the likelihood function as a goodness-of-fit measure is generally discouraged, since it does not carry much information as a test statistic [5]. However, none of the standard goodness-of-fit tests (such as the Kolmogorov-Smirnov or Anderson-Darling test statistics) are well suited for the matrix distributions under study. They require binning and ordering of data points, or access to the cumulative distribution function, which is not feasible when the data samples are defined on the cone of positive definite matrices. No adequate replacement has, to the best of our knowledge, been suggested in the literature, which is what we try to remedy.

In [6], Li and Papodopoulos provide a general framework for the design of moment-based goodness-of-fit tests. Their simple idea is to compare sampling moments with population moments, and to combine these in a test statistic which is asymptotically normal, which makes it easy to perform hypothesis testing or to obtain a p-value, i.e., the probability of the given observation under the hypothesized model. We have applied their theory to a set of compound distribution models for polarimetric radar data, and derived a set of candidate test statistics using certain moment expressions for the polarimetric covariance matrix.

4. MOMENTS OF MATRIX DISTRIBUTIONS

Specifically, we look at trace moments and log-determinant moments of the polarimetric covariance matrix. The trace moments represent the standard definition of matrix moments. The log-determinant moments are multivariate generalisations of the log-cumulants derived and successfully applied to the analysis of single polarisation SAR data in a series of papers by Nicolas et al. (See e.g. [7]). We further use a diagram of the space spanned by certain log-determinant moments to illustrate the statistical distance between the given models and empirical data. The diagram provides intuition about the modelling capabilities of the different models, and how well they can adapt to the non-Gaussian data found in textured and heterogeneous areas. This is also a polarimetric extension of the visualisation tool introduced by Nicolas et al.

5. RESULTS

We evaluate the strength of the goodness-of-fit tests based on individual matrix moments, to rate their applicability to model fit assessment. We then compare these tests to a test that combines several moments. These examinations are performed with simulated data generated from the candidate distribution models. We further demonstrate how our method behaves when it is used to compare model fit for real PolSAR data acquired by the NASA/JPL AIRSAR instrument, the DLR E-SAR instrument, and the DCRS EMISAR instrument. Numerical results are illustrated with log-determinant moment diagrams.

6. REFERENCES

[1] N. Goodman, “Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction),” Ann. Math. Statist., vol. 34, pp. 152-177, Mar. 1963.

[2] J.-S. Lee, D.L. Schuler, R.H. Lang, and K.J. Ranson, “K-distribution for multi-look processed polarimetric SAR imagery,” Proc. IEEE Int. Geosci. Remote Sensing Symp., IGARSS 1994, Pasadena, USA, vol. 4, pp. 2179-2181, Aug. 1994.

[3] C.C. Freitas, A.C. Frery, and A.H. Correia, “The polarimetric G distribution for SAR data analysis,” Environmetrics, vol. 16, pp. 13 - 31, 2005.

[4] A.P. Doulgeris and T. Eltoft, “Scale mixture of Gaussians modelling of polarimetric SAR data,” Proc. IEEE Int. Geosci. Remote Sensing Symp., IGARSS 2007, Barcelona, Spain, pp. 3486-3489, Jul. 2007.

[5] J. Heinrich, “Pitfalls of goodness-of-fit from likelihood,” Proc. PHYSTAT 2003 Conf. on Statist. Problems in Particle Phys., Astrophys., and Cosmology, Stanford, USA, pp. 52-55 , Oct. 2003.

[6] G. Li and A. Papadopoulos, “A note on goodness-of-fit tests using moments,” Statistica, no. 1, pp. 71-86, 2002.

[7] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre, “A new statistical model for Markovian classification of urban areas in high-resolution SAR images,” vol. 42, no. 10, pp. 2046-2057, Oct. 2004.