REFERENCES

[1] V. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory of Probability and its Applications*, vol. 16, no. 2, pp. 264–280, 1971.

[2] T. Grobler, E. R. Ackermann, J. C. Olivier, A. J. van Zyl, and W. Kleynhans, "Land-cover separability analysis of MODIS time-series data using a combined Simple Harmonic Oscillator and a mean reverting stochastic process," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 3, pp. 857–866, 2012.

[3] J. Hermance, R. Jacob, B. Bradley, and J. Mustard, "Extracting phenological signals from multiyear AVHRR NDVI time series: Framework for applying high-order annual splines with roughness damping," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 10, pp. 3264–3276, 2007.

[4] G. Uhlenbeck and L. Ornstein, "On the theory of the Brownian motion," *Physical Review*, vol. 36, no. 5, p. 823, 1930.

[5] S. Lhermitte, J. Verbesselt, I. Jonckheere, K. Nackaerts, J. van Aardt, W. Verstraeten, and P. Coppin, "Hierarchical image segmentation based on similarity of NDVI time series," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 506–521, 2008.

[6] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 241–257, 1954.

[7] R. Lunetta, J. Knight, J. Ediriwickrema, J. Lyon, and L. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sensing of Environment*, vol. 105, no. 2, pp. 142–154, 2006.

[8] J. Foley, R. DeFries, G. Asner, C. Barford, G. Bonan, S. Carpenter, F. Chapin, M. Coe, G. Daily, H. Gibbs *et al.*, "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, 2005.

[9] G. Benítez, A. Pérez-Vázquez, M. Nava-Tablada, M. Equihua, and J. Álvarez-Palacios, "Urban expansion and the environmental effects of informal settlements on the outskirts of Xalapa city, Veracruz, Mexico," *Environment and Urbanization*, vol. 24, no. 1, pp. 149–166, 2012.

[10] W. Kleynhans, "Detecting land-cover change using MODIS time-series data," Ph.D. dissertation, University of Pretoria, 2011.

[11] "United Nations division for sustainable development – national information – South Africa," 2010. [Online]. Available: `http://www.un.org/esa/agenda21/natlinfo/countr/safrica/human_settlements`

[12] B. Salmon, "Improved hyper-temporal feature extraction methods for land cover change detection in satellite time series," Ph.D. dissertation, University of Pretoria, 2012.

[13] D. Lu, P. Mausel, E. Brondizio, and E. Moran, "Change detection techniques," *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2365–2401, 2004.

[14] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870, 2007.

[15] H. Carrão, P. Gonçalves, and M. Caetano, "Contribution of multispectral and multitemporal information from MODIS images to land cover classification," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 986–997, 2008.

[16] E. Ackermann, T. Grobler, A. van Zyl, K. Steenkamp, and J. Olivier, "Minimum error land cover separability analysis and classification of MODIS time series data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. Vancouver, Canada: IEEE, 2011, pp. 2999–3002.

[17] R. Lasaponara, "Estimating interannual variations in vegetated areas of Sardinia island using

SPOT/VEGETATION NDVI temporal series," *Geoscience and Remote Sensing Letters, IEEE*, vol. 3, no. 4, pp. 481–483, 2006.

[18] D. Alcaraz, J. Paruelo, and J. Cabello, "Identification of current ecosystem functional types in the Iberian Peninsula," *Global Ecology and Biogeography*, vol. 15, no. 2, pp. 200–212, 2006.

[19] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sensing of Environment*, vol. 114, no. 1, pp. 106–115, 2010.

[20] J. Verbesselt, R. Hyndman, A. Zeileis, and D. Culvenor, "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," *Remote Sensing of Environment*, vol. 114, no. 12, pp. 2970–2980, 2010.

[21] B. Salmon, J. Olivier, W. Kleynhans, K. Wessels, F. van den Bergh, and K. Steenkamp, "The use of a multilayer perceptron for detecting new human settlements from a time series of MODIS images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, no. 6, pp. 873–883, 2011.

[22] C. Schaaf, F. Gao, A. Strahler, W. Lucht, X. Li, T. Tsang, N. Strugnell, X. Zhang, Y. Jin, J. Muller *et al.*, "First operational BRDF, albedo nadir reflectance products from MODIS," *Remote Sensing of Environment*, vol. 83, no. 1-2, pp. 135–148, 2002.

[23] E. Ackermann, "Sequential land cover classification," Master's thesis, University of Pretoria, 2011.

[24] B. Ghosh and P. Sen, Eds., *Handbook of Sequential Analysis*. New York: Dekker, 1991, vol. 118.

[25] T. Lai, "Sequential analysis: Some classical problems and new challenges," *Statistica Sinica*, vol. 11, pp. 303–408, 2001.

[26] L. Guanter, K. Segl, and H. Kaufmann, "Simulation of optical remote-sensing scenes with application to the ENMAP hyperspectral mission," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 7, pp. 2340–2351, 2009.

[27] W. Verhoef and H. Bach, "Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data," *Remote Sensing of Environment*, vol. 109, no. 2, pp. 166–182, 2007.

[28] S. Boriah, "Time series change detection: Algorithms for land cover change," Ph.D. dissertation, The University of Minnesota, 2010.

[29] J. Verbesselt, A. Zeileis, and M. Herold, "Near real-time disturbance detection in terrestrial ecosystems using satellite image time series: drought detection in Somalia," Faculty of Economics and Statistics, University of Innsbruck, Working Papers, 2011. [Online]. Available: `http://EconPapers.repec.org/RePEc:inn:wpaper:2011-18`

[30] T. Grobler, E. Ackermann, A. van Zyl, J. Olivier, W. Kleynhans, and B. Salmon, "Using Page's Cumulative Sum test on MODIS time series to detect land-cover changes," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 2, pp. 332–336, March 2013, DOI: 10.1109/LGRS.2012.2205556.

[31] B. Jiang, S. Liang, J. Wang, and Z. Xiao, "Modeling MODIS LAI time series using three statistical methods," *Remote Sensing of Environment*, vol. 114, no. 7, pp. 1432–1444, 2010.

[32] T. Grobler, E. Ackermann, A. van Zyl, J. Olivier, W. Kleynhans, and B. Salmon, "An inductive approach to simulating multispectral MODIS surface reflectance time series," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 3, pp. 446–450, May 2013, DOI: 10.1109/LGRS.2012.2208446.

[33] T. Lillesand and R. Kiefier, *Remote Sensing and Image Interpretation*, 3rd ed. John Wiley & Sons, Inc., 1994.

[34] J. Cambell, *Introduction to Remote Sensing*, 4th ed. The Guilford Press, 2008.

[35] P. Gibson, *Introductory Remote Sensing, Principles and Concepts*, 1st ed. Routledge, 2000.

[36] E. Liebenberg and C. Vlok, "The Interpretation of maps, aerial photographs and satellite images," Study guide2 for GGH203–V, University of South Africa.

[37] B. Ramachandran, C. Justice, and A. Abrams, Eds., *Land Remote Sensing and Global Envi-*

ronmental Change, NASA's Earth Observing System and the Science of ASTER and MODIS. Springer, 2011.

[38] J. Qu, W. Gao, M. Kafatos, E. Robert, and V. V. Salomonson, Eds., *Earth Science Satellite Remote Sensing, Science and Instruments, Volume 1*. Springer, 2006.

[39] S. Davis, D. Landgrebe, T. Phillips, P. Swain, R. Hoffer, J. Lindenlaub, and L. Silva, *Remote Sensing, the Quantitative Aproach*, 1st ed. McGraw-Hill, 1978.

[40] M. Hazewinkel, Ed., *Encyclopedia of Mathematics*. Springer, 2001.

[41] T. García-Mora, J. Mas, and E. Hinkley, "Land cover mapping applications with MODIS: a literature review," *International Journal of Digital Earth*, vol. 5, no. 1, pp. 63–87, 2012.

[42] Q. Weng, Ed., *Advances in Environmental Remote Sensing, Sensors, Algorithms, and Applications*. CRC Press, Taylor & Francis Group, 2011.

[43] R. Colditz, "Time series generation and classification of MODIS data for land cover mapping," Ph.D. dissertation, Julius-Maximilians-Universität Würzburg, September 2007.

[44] J. Neyman and E. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, pp. 289–337, 1933.

[45] A. Wald, "Sequential tests of hypotheses," *Annals of Mathematical Statistics*, vol. 16, pp. 117–186, 1945.

[46] ——, *Sequential Analysis*. John Wiley & Sons, Inc., 1947.

[47] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *The Annals of Mathematical Statistics*, vol. 19, no. 3, pp. 326–339, 1948.

[48] H. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge University Press, 2009.

[49] W. Shewhart, *Economic Control of Quality of Manufactured Product*. Princeton, NJ: D. Van Nostrand Reinhold, 1931.

[50] M. Girshick and H. Rubin, "A Bayes approach to a quality control model," *The Annals of Mathematical Statistics*, vol. 23, no. 1, pp. 114–125, 1952.

[51] A. Shiryaev, "The problem of the most rapid detection of a disturbance in a stationary process," in *Soviet Mathematics Doklady*, vol. 2, no. 795-799, 1961.

[52] G. Lorden, "Procedures for reacting to a change in distribution," *Annals of Mathematical Statistics*, vol. 42, pp. 1897–1908, 1971.

[53] G. Moustakides, "Optimal stopping times for detecting changes in distributions," *The Annals of Statistics*, vol. 14, no. 4, pp. 1379–1387, 1986.

[54] Y. Ritov, "Decision theoretic optimality of the CUSUM procedure," *The Annals of Statistics*, pp. 1464–1469, 1990.

[55] S. Roberts, "A comparison of some control chart procedures," *Technometrics*, pp. 411–430, 1966.

[56] M. Pollak, "Optimal detection of a change in distribution," *The Annals of Statistics*, pp. 206–227, 1985.

[57] A. Polunchenko, "Quickest change detection with applications to distributed multi-sensor systems," Ph.D. dissertation, University of Southern California, 2009.

[58] G. Moustakides, A. Polunchenko, and A. Tartakovsky, "A numerical approach to performance analysis of quickest change-point detection procedures," *Statistica Sinica*, vol. 21, pp. 571–596, 2011.

[59] M. Basseville, I. Nikiforov *et al.*, *Detection of Abrupt Changes: Theory and Application*. Prentice Hall Englewood Cliffs, 1993, vol. 15.

[60] F. James, *Statistical Methods in Experimental Physics*, 2nd ed. World Scientific Pub Co Inc, 2006.

[61] S. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Prentice Hall, 1998.

[62] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, 1992.

[63] A. Goel and S. Wu, "Determination of the ARL and a contour nomogram for CUSUM charts to control normal mean," *Technometrics*, vol. 13, no. 2, pp. 221–230, 1971.

[64] D. Brook and D. Evans, "An approach to the probability distribution of CUSUM run length," *Biometrika*, vol. 59, no. 3, pp. 539–549, 1972.

[65] L. Kantorovich and V. Krylov, *Approximate Methods of Higher Analysis*. Interscience Groningen, 1964.

[66] A. Shiryaev, *Optimal Stopping Rules*. New-York: Springer, 1978.

[67] ——, "On optimum methods in quickest detection problems," *Theory of Probability and its Applications*, vol. 8, no. 1, pp. 22–46, 1963.

[68] J. Schott, S. Brown, R. Raqueno, H. Gross, and G. Robinson, "An advanced synthetic image generation model and its application to multi/hyperspectral algorithm development," *Canadian Journal of Remote Sensing*, vol. 25, no. 2, pp. 99–111, 1999.

[69] A. Singh, "Digital change detection techniques using remotely-sensed data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989.

[70] D. Mouat, G. Mahin, and J. Lancaster, "Remote sensing techniques in the analysis of change detection," *Geocarto International*, vol. 8, no. 2, pp. 39–50, 1993.

[71] P. Coppin, I. Jonckheere, K. Nackaerts, B. Muys, and E. Lambin, "Digital change detection methods in ecosystem monitoring: A review," *International Journal of Remote Sensing*, vol. 25, no. 9, pp. 1565–1596, 2004.

[72] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *Image Processing, IEEE Transactions on*, vol. 14, no. 3, pp. 294–307, 2005.

[73] G. Jianya, S. Haigang, M. Guorui, and Z. Qiming, "A review of multi-temporal remote sensing data change detection algorithms," *The International Archives of the Photogrammetry, Remote*

Sensing and Spatial Information Sciences, vol. 37, pp. 757–762, 2008.

[74] H. Carrão, P. Gonalves, and M. Caetano, "A nonlinear harmonic model for fitting satellite image time series: Analysis and prediction of land cover dynamics," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 4, pp. 1919–1930, 2010.

[75] S. Jacquemoud, W. Verhoef, F. Baret, C. Bacour, P. Zarco-Tejada, G. Asner, C. François, and S. Ustin, "PROSPECT+ SAIL models: A review of use for vegetation characterization," *Remote Sensing of Environment*, vol. 113, pp. S56–S66, 2009.

[76] P. Jönsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 8, 2002.

[77] X. Zhang, M. Friedl, C. Schaaf, A. Strahler, J. Hodges, F. Gao, B. Reed, and A. Huete, "Monitoring vegetation phenology using MODIS," *Remote Sensing of Environment*, vol. 84, no. 3, pp. 471–475, 2003.

[78] W. Kleynhans, J. Olivier, K. Wessels, F. Van den Bergh, B. Salmon, and K. Steenkamp, "Improving land cover class separation using an extended Kalman filter on MODIS NDVI time-series data," *Geoscience and Remote Sensing Letters, IEEE*, vol. 7, no. 2, pp. 381–385, 2010.

[79] K. Levenberg, "A method for the solution of certain problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[80] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[81] E. Bibbona, G. Panfilo, and P. Tavella, "The Ornstein–Uhlenbeck process as a model of a low pass filtered white noise," *Metrologia*, vol. 45, p. S117, 2008.

[82] Z. Brzeźniak and T. Zastawniak, *Basic Stochastic Processes: A Course Through Exercises*. Springer Verlag, 1999.

[83] T. van den Berg, "Calibrating the Ornstein-Uhlenbeck model," May 2011, `http://www.sitmo.com/article/calibrating-the-ornstein-uhlenbeck-model/`.

[84] T. Björk, *Arbitrage Theory in Continuous Time*. Oxford University Press, USA, 2009.

[85] J. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Berlin, Germany: Springer Verlag, 2006.

[86] J. Paola and R. Schowengerdt, "A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery," *International Journal of Remote Sensing*, vol. 16, no. 16, pp. 3033–3058, 1995.

[87] P. Swain and H. Hauska, "The decision tree classifier design and potential," *IEEE Transactions on Geoscience Electronics*, vol. 15, no. 3, pp. 142–147, 1977.

[88] R. De Fries, M. Hansen, J. Townshend, and R. Sohlberg, "Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers," *International Journal of Remote Sensing*, vol. 19, no. 16, pp. 3141–3168, 1998.

[89] C. Huang, L. Davis, and J. Townshend, "An assessment of Support Vector Machines for land cover classification," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 725–749, 2002.

[90] G. Foody and A. Mathur, "A relative evaluation of multiclass image classification by Support Vector Machines," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 6, pp. 1335–1343, 2004.

[91] G. Mountrakis, J. Im, and C. Ogole, "Support Vector Machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, 2011.

[92] Y. Sohn and N. Rebello, "Supervised and unsupervised spectral angle classifiers," *Photogrammetric Engineering and Remote Sensing*, vol. 68, no. 12, pp. 1271–1282, 2002.

[93] T. Lee, M. Lewicki, and T. Sejnowski, "ICA mixture models for unsupervised classification of non-Gaussian classes and automatic context switching in blind signal separation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 10, pp. 1078–1089, 2000.

[94] C. Shah, M. Arora, and P. Corresponding, "Unsupervised classification of hyperspectral data:

An ICA mixture model based approach," *International Journal of Remote Sensing*, vol. 25, no. 2, pp. 481–487, 2004.

[95] A. Koltunov and E. Ben-Dor, "A new approach for spectral feature extraction and for unsupervised classification of hyperspectral data based on the Gaussian mixture model," *Remote Sensing Reviews*, vol. 20, no. 2, pp. 123–167, 2001.

[96] ——, "Mixture density separation as a tool for high-quality interpretation of multi-source remote sensing data and related issues," *International Journal of Remote Sensing*, vol. 25, pp. 3275–3299, 2004.

[97] M. Collins, C. Dymond, and E. Johnson, "Mapping subalpine forest types using networks of nearest neighbour classifiers," *International Journal of Remote Sensing*, vol. 25, no. 9, pp. 1701–1721, 2004.

[98] P. Hardin, "Parametric and nearest-neighbor methods for hybrid classification: A comparison of pixel assignment accuracy," *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 12, pp. 1439–1448, 1994.

[99] R. Haapanen, A. Ek, M. Bauer, and A. Finley, "Delineation of forest/nonforest land use classes using nearest neighbor methods," *Remote Sensing of Environment*, vol. 89, no. 3, pp. 265–271, 2004.

[100] G. Foody, "Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data," *International Journal of Remote Sensing*, vol. 17, no. 7, pp. 1317–1340, 1996.

[101] J. Zhang and G. Foody, "Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and Artificial Neural Network approaches," *International Journal of Remote Sensing*, vol. 22, no. 4, pp. 615–628, 2001.

[102] J. Adams, D. Sabol, V. Kapos, R. Almeida Filho, D. Roberts, M. Smith, and A. Gillespie, "Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon," *Remote Sensing of Environment*, vol. 52, no. 2, pp. 137–154, 1995.

[103] D. Lu, E. Moran, and M. Batistella, "Linear mixture model applied to Amazonian vegetation classification," *Remote Sensing of Environment*, vol. 87, no. 4, pp. 456–469, 2003.

[104] D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, "Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models," *Remote Sensing of Environment*, vol. 65, no. 3, pp. 267–279, 1998.

[105] B. Mannan and A. Ray, "Crisp and fuzzy competitive learning networks for supervised classification of multispectral IRS scenes," *International Journal of Remote Sensing*, vol. 24, no. 17, pp. 3491–3502, 2003.

[106] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 26, no. 1, pp. 52–67, 1996.

[107] R. Schowengerdt, "On the estimation of spatial-spectral mixing with classifier likelihood functions," *Pattern Recognition Letters*, vol. 17, no. 13, pp. 1379–1387, 1996.

[108] P. Aplin and P. Atkinson, "Sub-pixel land cover mapping for per-field classification," *International Journal of Remote Sensing*, vol. 22, no. 14, pp. 2853–2858, 2001.

[109] C. Lloyd, S. Berberoglu, P. Curran, and P. Atkinson, "A comparison of texture measures for the per-field classification of Mediterranean land cover," *International Journal of Remote Sensing*, vol. 25, no. 19, pp. 3943–3965, 2004.

[110] U. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, no. 3, pp. 239–258, 2004.

[111] I. Gitas, G. Mitri, and G. Ventura, "Object-based image classification for burned area mapping of Creus Cape, Spain, using NOAA-AVHRR imagery," *Remote Sensing of Environment*, vol. 92, no. 3, pp. 409–413, 2004.

[112] V. Walter, "Object-based classification of remote sensing data for change detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, no. 3, pp. 225–238, 2004.

[113] L. Wang, W. Sousa, P. Gong, and G. Biging, "Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of Panama," *Remote Sensing of Environment*, vol. 91, no. 3, pp. 432–440, 2004.

[114] S. Magnussen, P. Boudewyn, and M. Wulder, "Contextual classification of Landsat TM images to forest inventory cover types," *International Journal of Remote Sensing*, vol. 25, no. 12, pp. 2421–2440, 2004.

[115] E. Mohn, N. Hjort, and G. Storvik, "A simulation study of some contextual classification methods for remotely sensed data," *Geoscience and Remote Sensing, IEEE Transactions on*, no. 6, pp. 796–804, 1987.

[116] B. Jeon and D. Landgrebe, "Classification with spatio-temporal interpixel class dependency contexts," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 30, no. 4, pp. 663–672, 1992.

[117] Y. Jhung and P. Swain, "Bayesian contextual classification based on modified M-estimates and Markov random fields," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 34, no. 1, pp. 67–75, 1996.

[118] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 721–741, 1984.

[119] J. Stuckens, P. Coppin, and M. Bauer, "Integrating contextual information with per-pixel classification for improved land cover classification," *Remote Sensing of Environment*, vol. 71, no. 3, pp. 282–296, 2000.

[120] M. Hodgson, J. Jensen, J. Tullis, K. Riordan, and C. Archer, "Synergistic use of LIDAR and color aerial photography for mapping urban parcel imperviousness," *Photogrammetric Engineering and Remote Sensing*, vol. 69, no. 9, pp. 973–980, 2003.

[121] Q. Du and C. Chang, "A linear constrained distance-based discriminant analysis for hyperspectral image classification," *Pattern Recognition*, vol. 34, no. 2, pp. 361–373, 2001.

[122] S. Myint, "A robust texture analysis and classification approach for urban land-use and land-

cover feature discrimination," *Geocarto International*, vol. 16, no. 4, pp. 29–40, 2001.

[123] G. Okin, D. Roberts, B. Murray, and W. Okin, "Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments," *Remote Sensing of Environment*, vol. 77, no. 2, pp. 212–225, 2001.

[124] J. Benediktsson, J. Sveinsson, and K. Amason, "Classification and feature extraction of AVIRIS data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 33, no. 5, pp. 1194–1205, 1995.

[125] M. Ulfarsson, J. Benediktsson, and J. Sveinsson, "Data fusion and feature extraction in the wavelet domain," *International Journal of Remote Sensing*, vol. 24, no. 20, pp. 3933–3945, 2003.

[126] T. Rashed, J. Weeks, M. Gadalla, and A. Hill, "Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: A case study of the Greater Cairo region, Egypt." *Geocarto International*, vol. 16, no. 4, pp. 7–18, 2001.

[127] R. Lunetta and M. Balogh, "Application of multi-temporal Landsat 5 TM imagery for wetland identification," *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 11, pp. 1303–1310, 1999.

[128] M. Heinl and U. Tappeiner, "The benefits of considering land cover seasonality in multispectral image classification," *Journal of Land Use Science*, vol. 1, pp. 1–19, 2011.

[129] J. Townshend, T. Goff, and C. Tucker, "Multitemporal dimensionality of images of Normalized Difference Vegetation Index at continental scales," *Geoscience and Remote Sensing, IEEE Transactions on*, no. 6, pp. 888–895, 1985.

[130] M. Hall-Beyer, "Comparison of single-year and multiyear NDVI time series principal components in cold temperate biomes," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 11, pp. 2568–2574, 2003.

[131] R. Colditz, M. Schmidt, C. Conrad, M. Hansen, and S. Dech, "Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions," *Remote Sensing of Environment*, vol. 115, pp. 3264–3275, 2011.

[132] V. Simonneaux, B. Duchemin, D. Helson, S. Er-Raki, A. Olioso, and A. Chehbouni, "The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco," *International Journal of Remote Sensing*, vol. 29, no. 1, pp. 95–116, 2008.

[133] N. Viovy, "Automatic classification of time series (ACTS): A new clustering method for remote sensing time series," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1537–1560, 2000.

[134] R. Juárez and W. Liu, "FFT analysis on NDVI annual cycle and climatic regionality in Northeast Brazil," *International Journal of Climatology*, vol. 21, no. 14, pp. 1803–1820, 2001.

[135] M. Jakubauskas, D. Legates, and J. Kastens, "Crop identification using harmonic analysis of time-series AVHRR NDVI data," *Computers and Electronics in Agriculture*, vol. 37, no. 1, pp. 127–139, 2002.

[136] D. de Castro Victoria, A. da Paz, A. Coutinho, J. Kastens, and J. Brown, "Cropland area estimates using MODIS NDVI time series in the state of Mato Grosso, Brazil," *Pesquisa Agropecuária Brasileira, Brasília*, vol. 47, no. 9, pp. 1270–1278, 2012.

[137] G. Galford, J. Mustard, J. Melillo, A. Gendrin, C. Cerri, and C. Cerri, "Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 576–587, 2008.

[138] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag New York Inc, 2000.

[139] C. Burges, "A tutorial on Support Vector Machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[140] J. Gualtieri and R. Cromp, "Support Vector Machines for hyperspectal remote sensing classification," in *Proceedings of the SPIE, 27th AIPR Workshop: Advances in Computer Assisted Recognition*, Washington, DC, October 1998, pp. 221–232.

[141] G. Zhu and D. Blumberg, "Classification using ASTER data and SVM algorithms;: The case study of Beer Sheva, Israel," *Remote Sensing of Environment*, vol. 80, no. 2, pp. 233–240, 2002.

[142] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with Support Vector Machines," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 8, pp. 1778–1790, 2004.

[143] M. Pal and P. Mather, "Support Vector Machines for classification in remote sensing," *International Journal of Remote Sensing*, vol. 26, no. 5, pp. 1007–1011, 2005.

[144] Y. Liu, Y. Xu, R. Shi, and Z. Niu, "Evaluation of various classifiers on regional land cover classification using MODIS data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2005 IEEE International*, vol. 2, July 2005, pp. 1281–1283.

[145] J. Guo, J. Zhang, Y. Zhang, and Y. Cao, "Study on the comparison of land cover classification for multitemporal MODIS images," in *International Workshop on Earth Observation and Remote Sensing Applications*, Beijing, June 2008, pp. 1–6.

[146] C. Ye, Y. Liu, J. Peng, P. Song, and D. Zhao, "Improving MODIS land cover classification using NDVI time-series and Support Vector Machines in the Poyang Lake Basin, China," in *Wireless Communications Networking and Mobile Computing*, Chengdu, September 2010, pp. 1–4.

[147] H. Cai and S. Zhang, "Regional land cover classification from MODIS time-series and geographical data using Support Vector Machines," in *IEEE Youth Conference on Information Computing and Telecommunications*, Beijing, November 2010, pp. 102–105.

[148] S. Li, Y. Zhu, J. Feng, P. Ai, and X. Chen, "Comparative study of three feature selection algorithms for regional land cover classification using MODIS data," in *Congress on Image and Signal Processing*, 2008, pp. 565–569.

[149] B. Salmon, J. Olivier, K. Wessels, W. Kleynhans, F. Van den Bergh, and K. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, no. 99, pp. 1–9, 2011.

[150] R. Lasaponara and A. Lanorte, "Satellite time-series analysis," *International Journal of Remote Sensing*, vol. 33, no. 15, pp. 4649–4652, 2012.

[151] A. Srivastava, "A novel approach of land cover change mapping using hypertemporal images," Master's thesis, University of Twente, 2011.

[152] T. Sohl, "Change analysis in the United Arab Emirates: an investigation of techniques," *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 4, pp. 475–484, 1999.

[153] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to Tsunami-damage assessment," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 6, pp. 1658–1670, 2007.

[154] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Processing*, vol. 90, no. 5, pp. 1471–1485, 2010.

[155] T. Celik and K. Ma, "Unsupervised change detection for satellite images using dual-tree complex wavelet transform," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 3, pp. 1199–1210, 2010.

[156] C. Jha and N. Unni, "Digital change detection of forest conversion of a dry tropical Indian forest region," *International Journal of Remote Sensing*, vol. 15, no. 13, pp. 2543–2552, 1994.

[157] A. Prakash and R. Gupta, "Land-use mapping and change detection in a coal mining area-a case study in the Jharia coalfield, India," *International Journal of Remote Sensing*, vol. 19, no. 3, pp. 391–410, 1998.

[158] J. Townshend and C. Justice, "Spatial variability of images and the monitoring of changes in the Normalized Difference Vegetation Index," *International Journal of Remote Sensing*, vol. 16, no. 12, pp. 2187–2195, 1995.

[159] E. Lambin and A. Strahler, "Indicators of land-cover change for Change Vector Analysis in multitemporal space at coarse spatial scales," *International Journal of Remote Sensing*, vol. 15, no. 10, pp. 2099–2119, 1994.

[160] J. Deng, K. Wang, Y. Deng, and G. Qi, "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data," *International Journal of Remote Sensing*, vol. 29, no. 16, pp. 4823–4838, 2008.

[161] J. Collins and C. Woodcock, "An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data," *Remote Sensing of Environment*, vol. 56, no. 1, pp. 66–77, 1996.

[162] M. Ridd and J. Liu, "A comparison of four algorithms for change detection in an urban environment," *Remote Sensing of Environment*, vol. 63, no. 2, pp. 95–100, 1998.

[163] C. Munyati and T. Kabanda, "Using multitemporal Landsat TM imagery to establish land use pressure induced trends in forest and woodland cover in sections of the Soutpansberg Mountains of Venda region, Limpopo Province, South Africa," *Regional Environmental Change*, vol. 9, no. 1, pp. 41–56, 2009.

[164] L. Bruzzone and S. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 35, no. 4, pp. 858–867, 1997.

[165] G. Moser, S. Serpico, and G. Vernazza, "Unsupervised change detection from multichannel SAR images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 4, no. 2, pp. 278–282, 2007.

[166] S. Gopal and C. Woodcock, "Remote sensing of forest change using Artificial Neural Networks," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 34, no. 2, pp. 398–404, 1996.

[167] S. Macomber and C. Woodcock, "Mapping and monitoring conifer mortality using remote sensing in the Lake Tahoe Basin," *Remote Sensing of Environment*, vol. 50, no. 3, pp. 255–266, 1994.

[168] O. Samain, J. Roujean, and B. Geiger, "Use of a Kalman filter for the retrieval of surface BRDF coefficients with a time-evolving model based on the ECOCLIMAP land cover classification," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1337–1346, 2008.

[169] M. Chen, S. Liu, L. Tieszen, and D. Hollinger, "An improved state-parameter analysis of ecosystem models using data assimilation," *Ecological Modelling*, vol. 219, no. 3-4, pp. 317–326, 2008.

[170] J. Slater and R. Brown, "Changing landscapes: Monitoring environmentally sensitive areas using satellite imagery," *International Journal of Remote Sensing*, vol. 21, no. 13-14, pp. 2753–2767, 2000.

[171] W. Michalak, "GIS in land use change analysis: Integration of remotely sensed data into GIS," *Applied Geography*, vol. 13, no. 1, pp. 28–44, 1993.

[172] T. Stone and P. Lefebvre, "Using multi-temporal satellite data to evaluate selective logging in Para, Brazil," *International Journal of Remote Sensing*, vol. 19, no. 13, pp. 2517–2526, 1998.

[173] G. Henebry, "Detecting change in grasslands using measures of spatial dependence with Landsat TM data," *Remote Sensing of Environment*, vol. 46, no. 2, pp. 223–234, 1993.

[174] F. Wang, "A knowledge-based vision system for detecting land changes at urban fringes," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 31, no. 1, pp. 136–145, 1993.

[175] R. Lawrence and W. Ripple, "Calculating change curves for multitemporal satellite imagery: Mount St. Helens 1980–1995," *Remote Sensing of Environment*, vol. 67, no. 3, pp. 309–319, 1999.

[176] J. Morisette, S. Khorram, and T. Mace, "Land-cover change detection enhanced with generalized linear models," *International Journal of Remote Sensing*, vol. 20, no. 14, pp. 2703–2721, 1999.

[177] T. Yue, S. Chen, B. Xu, Q. Liu, H. Li, G. Liu, and Q. Ye, "A curve-theorem based approach for change detection and its application to Yellow River Delta," *International Journal of Remote Sensing*, vol. 23, no. 11, pp. 2283–2292, 2002.

[178] Q. Zhang, J. Wang, X. Peng, P. Gong, and P. Shi, "Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data," *International Journal of Remote Sensing*, vol. 23, no. 15, pp. 3057–3078, 2002.

[179] J. Read and N. S.-. N. Lam, "Spatial methods for characterising land cover and detecting land-cover changes for the tropics," *International Journal of Remote Sensing*, vol. 23, no. 12, pp. 2457–2474, 2002.

[180] J. Borak, E. Lambin, and A. Strahler, "The use of temporal metrics for land cover change detection at coarse spatial scales," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1415–1432, 2000.

[181] R. Kennedy, W. Cohen, and T. Schroeder, "Trajectory-based change detection for automated characterization of forest disturbance dynamics," *Remote Sensing of Environment*, vol. 110, no. 3, pp. 370–386, 2007.

[182] A. Ramoela, "An innovative method to map land cover changes at a country level utilising hyper-temporal satellite images: A case study of Portugal," Master's thesis, International Institute for Geo-Information Science and Earth Observation, 2007.

[183] C. De Bie, M. Khan, A. Toxopeus, V. Venus, and A. Skidmore, "Hypertemporal image analysis for crop mapping and change detection," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 803–814, 2008.

[184] D. Mildrexler, M. Zhao, and S. Running, "Testing a MODIS global Disturbance Index across North America," *Remote Sensing of Environment*, vol. 113, no. 10, pp. 2103–2117, 2009.

[185] N. Coops, M. Wulder, and D. Iwanicka, "Large area monitoring with a MODIS-based Disturbance Index (DI) sensitive to annual and seasonal variations," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1250–1261, 2009.

[186] J. Beltran-Abaunza, "Method development to process hyper-temporal remote sensing images for change mapping," Master's thesis, International Institute for Geo-Information Science and Earth Observation, 2009.

[187] S. Boriah, V. Mithal, A. Garg, V. Kumar, M. Steinbach, C. Potter, and S. Klooster, "A comparative study of algorithms for land cover change," in *Proceedings of the Conference on Intelligent Data Understanding*, 2010.

[188] F. Van Den Bergh, K. Wessels, S. Miteff, T. Van Zyl, A. Gazendam, and A. Bachoo, "HiTempo: A platform for time-series analysis of remote-sensing satellite data in a high-performance computing environment," *International Journal of Remote Sensing*, vol. 33, no. 15, pp. 4720–4740, 2012.

[189] W. Kleynhans, B. Salmon, J. Olivier, K. Wessels, and F. Van den Bergh, "An autocorrelation analysis approach to detecting land cover change using hyper-temporal time-series data," in *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, 2011, pp. 94–97.

[190] W. Kleynhans, B. Salmon, J. Olivier, F. van den Bergh, K. Wessels, T. Grobler, and K. Steenkamp, "Land cover change detection using autocorrelation analysis on MODIS time-series data: Detection of new human settlements in the Gauteng province of South Africa," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 3, pp. 777–783, 2012.

[191] W. Kleynhans, J. Olivier, K. Wessels, B. Salmon, F. Van den Bergh, and K. Steenkamp, "Detecting land cover change using an extended Kalman filter on MODIS NDVI time-series data," *Geoscience and Remote Sensing Letters, IEEE*, no. 99, pp. 506–510, 2011.

[192] V. Chandola and R. Vatsavai, "A Gaussian process based online change detection algorithm for monitoring periodic time series," in *Proceedings of SIAM Data Mining Conference*, 2011, pp. 95–116.

[193] Y. Kang, "Real-time change detection in time series based on growing feature quantization," in *The 2012 International Joint Conference on Neural Networks*. IEEE, 2012, pp. 1–6.

[194] J. Kučera, P. Barbosa, and P. Strobl, "Cumulative Sum charts - A novel technique for processing daily time series of MODIS data for burnt area mapping in Portugal," in *International Workshop on the Analysis of Multi-temporal Remote Sensing Images*, Leuven, August 2007.

[195] B. W. Silverman, *Density Estimation*. Chapman and Hall, 1986.

[196] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, 1992.

[197] A. Ihler and M. Mandel, "Kernel Density Estimation Toolbox for MATLAB," 2003. [Online]. Available: `http://www.ics.uci.edu/ ihler/code/kde.html`

[198] T. Grobler *et al.*, "Sequential classification of MODIS time series," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. Munchen, Germany: IEEE, 2012,

pp. 6236–6239.

[199] S. Lhermitte, J. Verbesselt, W. Verstraeten, and P. Coppin, "A comparison of time series similarity measures for classification and change detection of ecosystem dynamics," *Remote Sensing of Environment*, 2011.

[200] J. Cohen *et al.*, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[201] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, New Jersey, USA: Prentice Hall, Inc., 1996.

[202] J. Landis and G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, pp. 159–174, 1977.

[203] S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy, "SVM and Kernel Methods Matlab Toolbox," Perception Systèmes et Information, INSA de Rouen, Rouen, France, 2005. [Online]. Available: `http://asi.insa-rouen.fr/enseignants/ arakotom/toolbox/index.html`

APPENDIX A

MATHEMATICAL BACKGROUND

This appendix contains the mathematical background that is needed to understand Chapter 3, Chapter 4 and Chapter 5. In Section A.1 an overview is given of stochastic calculus. In the last five sections of the appendix the following mathematical concepts or ideas are summarised: Gaussian quadrature, Cholesky factorisation, the method of Langragian multipliers and KDE.

A.1 STOCHASTIC CALCULUS

The basic definitions used in Chapter 3 are defined in this section and can also be found in [48, 59, 82].

Definition 1 (A σ-field $\mathscr{F}$) *Let Ω be a non-empty set. A σ-field $\mathscr{F}$ on Ω is a family of subsets of Ω such that*

1. *the empty set $\emptyset$ belongs to $\mathscr{F}$;*
2. *if A belongs to $\mathscr{F}$, then so does the complement $\Omega \backslash A$;*
3. *if $A_1, A_2, \cdots$ is a sequence of sets in $\mathscr{F}$, then their union $A_1 \cup A_2 \cup \cdots$ also belongs to $\mathscr{F}$.*

Definition 2 (Family of Borel sets) *The family of Borel sets $\mathbb{B} = \mathscr{B}(\mathbb{R})$ is a σ-field on $\mathbb{R}$. The Borel σ-field in $\mathbb{R}$ is the smallest σ-field containing all intervals in $\mathbb{R}$.*

Definition 3 (Probability measure) *Let $\mathscr{F}$ be a σ-field on Ω. A probability measure P is a function*

$$P : \mathscr{F} \to [0,1]$$

such that

1. $P(\Omega) = 1$;

2. *if* $A_1, A_2, \cdots$ *are pairwise disjoint sets (that is,* $A_i \cap A_j = \emptyset$ *for* $i \neq j$*) belonging to* $\mathscr{F}$*, then*

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots .$$

The triple $(\Omega, \mathscr{F}, P)$ is known as a *probability space*. Each set A contained in $\mathscr{F}$ is referred to as an *event*. Whenever the $P(A) = 1$ then event A is said to occur *almost surely*.

Definition 4 (Random variable) *If* $\mathscr{F}$ *is a* σ*-field on* Ω*, then a function* $\xi : \Omega \to \mathbb{R}$ *is said to be* $\mathscr{F}$*-measurable if*

$$\{\xi \in B\} \in \mathscr{F}$$

for every Borel set $B \in \mathscr{B}(\mathbb{R})$*. If* $(\Omega, \mathscr{F}, P)$ *is a probability space, then such a function* ξ *is called a random variable.*

The short-hand notation $\{\xi \in B\}$ in the above definition represents the inverse image $\xi^{-1}(B)$. In expanded form $\{\xi \in B\}$ is written as

$$\{\omega \in \Omega : \xi(\omega) \in B\}.$$

Definition 5 (The σ field generated by a random variable) *The* σ*-field* $\sigma(\xi)$ *generated by a random variable* $\xi : \Omega \to \mathbb{R}$ *consists of all sets of the form* $\{\xi \in B\}$*, where* B *is a Borel set in* $\mathbb{R}$*.*

Definition 6 (The σ field generated by a family of random variables) *The* σ*-field* $\sigma\ \{\xi_i : i \in I\}$ *generated by a family* $\{\xi_i : i \in I\}$ *of random variables is defined to be the smallest* σ*-field containing all the events of the form* $\{\xi_i \in B\}$*, where* B *is a Borel set in* $\mathbb{R}$ *and* $i \in I$ *(*I *is an index set).*

Definition 7 (The distribution of a random variable and the cumulative distribution function) *Every random variable* $\xi : \Omega \to \mathbb{R}$ *gives rise to a probability measure*

$$P_\xi(B) = P\{\xi \in B\}$$

on $\mathbb{R}$ *defined on the* σ*-field of Borel sets* $B \in \mathscr{B}(\mathbb{R})$*. The probability measure* P_ξ *is called the distribution of* ξ*. The function* $F_\xi : \mathbb{R} \to [0,1]$ *defined by*

$$F_\xi(x) = P\{\xi \leq x\} = P_\xi((-\infty, x]),\ x \in \mathbb{R}$$

is called the cumulative distribution function of ξ.

Definition 7 implies that the measurability of ξ ensures that ξ generates a probability measure P_ξ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, such that $(\mathbb{R}, \mathscr{B}(\mathbb{R}), P_\xi)$ is also a probability space. If P_ξ is known, then the structure of $(\Omega, \mathscr{F}, P)$ is no longer needed to describe the behavior of ξ.

Definition 8 (Probability density function) *If there is a Borel function* $f_\xi : \mathbb{R} \to \mathbb{R}$ *such that for any Borel set* $B \subset \mathbb{R}$

$$P\{\xi \in B\} = \int_B f_\xi(x)dx,$$

then ξ *is said to be a random variable with absolutely continuous distribution and* f_ξ *is called the density of* ξ. *If there is a (finite or infinite) sequence of pairwise distinct real number* $x_1, x_2, \cdots$ *such that for any Borel set* $B \subset \mathbb{R}$

$$P\{\xi \in B\} = \sum_{x_i \in B} P\{\xi = x_i\},$$

then ξ *is said to have discrete distribution with values* $x_1, x_2, \cdots$ *and mass* $P\{\xi = x_i\}$ *at* x_i.

Definition 9 (Essential supremum) *Let* $(\xi_i)_{i \in I}$ *(where* I *is an index set) be a family of real-valued random variables on* $(\Omega, \mathscr{F}, P)$, *bounded by another variable. The essential supremum of* $(\xi_i)_{i \in I}$ *is* Ξ *(denoted by* $\Xi = \ ess\ sup_I\ \xi_i$*) if*

$$(\forall i \in I)\ \xi_i \leq \mathscr{X}\ P - almost\ surely \Leftrightarrow \Xi \leq \mathscr{X}\ P - almost\ surely.$$

Definition 10 (Expectation) *A random variable* $\xi : \Omega \to \mathbb{R}$ *is said to be integrable if*

$$\int_\Omega |\xi| dP < \infty.$$

Then

$$\mathbb{E}[\xi] = \int_\Omega \xi dP$$

exists and is called the expectation of ξ.

The expectation $\mathbb{E}[\xi]$ can also be expressed as a Riemann integral as follows [48]:

$$\begin{aligned}
\mathbb{E}[\xi] &= \int_\Omega \xi dP \\
&= \int_{\mathbb{R}} x dP_\xi \\
&= \int_{-\infty}^{\infty} x dF_\xi(x) \\
&= \int_{-\infty}^{\infty} x f_\xi(x) dx.
\end{aligned}$$

It is also important to note that $\frac{d}{dx}F_\xi(x) = f_\xi(x)$.

Definition 11 (Conditioning on an event) *For an integrable random variable ξ and any event $B \in \mathscr{F}$ such that $P(B) \neq 0$ the conditional expectation of ξ given B is defined by*

$$\mathbb{E}[\xi|B] = \frac{1}{P(B)}\int_B \xi dP.$$

Definition 12 (Conditioning on a discrete random variable) *Let ξ be an integrable random variable and let η be a discrete random variable. Then the conditional expectation of ξ given η is defined to be a random variable $\mathbb{E}[\xi|\eta]$ such that*

$$\mathbb{E}[\xi|\eta](\omega) = \mathbb{E}[\xi|\{\eta = y_n\}] \text{ if } \eta(\omega) = y_n$$

for any $n = 1, 2, \cdots$.

By using Definition 12 the following proposition can be derived (stated here without proof) [82]:

Proposition 1 *If ξ is an integrable random variable and η is a discrete random variable, then*

1. *$\mathbb{E}[\xi|\eta]$ is $\sigma(\eta)$-measurable;*

2. *For any $A \in \sigma(\eta)$*

$$\int_A \mathbb{E}[\xi|\eta]dP = \int_A \xi dP. \tag{A.1}$$

Definition 13 (Conditioning on an arbitrary random variable) *Let ξ be an integrable random variable and let η be an arbitrary random variable. Then the conditional expectation of ξ given η is defined to be a random variable $\mathbb{E}[\xi|\eta]$ such that*

1. *$\mathbb{E}[\xi|\eta]$ is $\sigma(\eta)$-measurable;*

2. *For any $A \in \sigma(\eta)$*

$$\int_A \mathbb{E}[\xi|\eta]dP = \int_A \xi dP. \tag{A.2}$$

Definition 14 (Conditioning on a σ-field) *Let ξ be an integrable random variable on a probability space $(\Omega, \mathscr{F}, P)$, and let $\mathscr{G}$ be a σ-field contained in $\mathscr{F}$. Then the conditional expectation of ξ given $\mathscr{G}$ is defined to be a random variable $\mathbb{E}[\xi|\mathscr{G}]$ such that*

1. $\mathbb{E}[\xi|\mathscr{G}]$ *is* $\mathscr{G}$*-measurable*

2. *For any* $A \in \mathscr{G}$

$$\int_A \mathbb{E}[\xi|\mathscr{G}]dP = \int_A \xi dP.$$

Definition 15 (Sample path) *The sequence of number* $\xi_1(\omega), \xi_2(\omega), \cdots$ *for any fixed* $\omega \in \Omega$ *is called a sample path.*

Definition 16 (Filtration) *A sequence of* σ*-fields* $\mathscr{F}_1, \mathscr{F}_2, \cdots$ *on* Ω *such that*

$$\mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{F}$$

is called a filtration.

Definition 17 (Adapted to a filtration) *A sequence of random variables* $\xi_1, \xi_2, \cdots$ *is adapted to a filtration* $\mathscr{F}_1, \mathscr{F}_2 \cdots$ *if* ξ_n *is* $\mathscr{F}_n$*-measurable for each* $n = 1, 2, \cdots$.

Definition 18 (Stopping time) *A random variable* T *with values in the set* $\{1, 2, \cdots\} \cup \{\infty\}$ *is called a stopping time (with respect to a filtration* $\mathscr{F}_n$*) if for each* $n = 1, 2, \cdots$

$$\{T = n\} \in \mathscr{F}_n.$$

At this point Wald's identities can be presented (and is stated here without proof) [48]. Wald's identities were used in Chapter 3 to derive the OC and ASN functions.

Theorem 6 (Wald's identities) *Suppose* $\{s_k; k = 1, 2, \cdots\}$ *is an i.i.d. sequence adapted to the filtration* $\{\mathscr{F}_k\}$*, and let* S_k *denote the sequence of cumulative sums,* $S_k = \sum_{i=1}^{k} s_i$*. Then the following statements are true:*

1. *Suppose* $\mathbb{E}[s_1]$ *is finite, then for every stopping time* T *satisfying* $\mathbb{E}[T] < \infty$, $\mathbb{E}[S_T] = \mathbb{E}[s_1]\mathbb{E}[T] \implies \mathbb{E}[T] = \frac{\mathbb{E}[S_T]}{\mathbb{E}[s_1]}$ *when* $\mathbb{E}[s_1] \neq 0$.

2. *Suppose* $\mathbb{E}[s_1^2]$ *is finite, then for ever stopping time* T *satisfying* $\mathbb{E}[T] < \infty$, $\mathbb{E}[S_T - T\mathbb{E}[s_1]]^2 = \mathbb{E}[T]\mathbb{E}[s_1 - \mathbb{E}[s_1]]^2 \implies \mathbb{E}[T] = \frac{\mathbb{E}[S_T^2]}{\mathbb{E}[s_1^2]}$ *when* $\mathbb{E}[s_1] = 0$.

3. *For scalars* $a, h > 0$ *define the stopping time* $T_{-a}^{h} = \inf\{k | S_k \notin (-b, a)\}$*. Suppose* $\omega \neq 0$ *is such that* $\mathbb{E}[e^{-\omega s_1}] < \infty$*, then* $\mathbb{E}[e^{\omega S_T}(\mathbb{E}[e^{-\omega s_1}])^{-T}] = 1$*, holds for any stopping time* T *such that* $P(T \leq T_{-a}^{h}) = 1$.

Definition 19 (Stochastic process) *A stochastic process is a family of random variables $\xi(t)$ parameterised by $t \in \mathscr{T}$, where $\mathscr{T} \subset \mathbb{R}$. When $\mathscr{T} = \{1,2,\cdots\}$ then $\xi(t)$ is a stochastic process in discrete time (a sequence of random variables). When $\mathscr{T}$ is an interval in $\mathbb{R}$ (typically $\mathscr{T} = [0,\infty]$) then $\xi(t)$ is a stochastic process in continuous time. Moreover, for every $\omega \in \Omega$ the function*

$$\mathscr{T} \ni t \to \xi(t,\omega)$$

is called a sample path of $\xi(t)$.

Definition 20 (Brownian motion) *The Wiener process (or Brownian motion) is a stochastic process $W(t)$ with values in $\mathbb{R}$ defined for $t \in [0,\infty)$ such that*

1. *$W(0) = 0$ almost surely;*

2. *the sample paths $t \to W(t)$ are almost surely continuous;*

3. *for any finite sequence of times $0 < t_1 < \cdots < t_n$ and Borel sets $A_1, \cdots, A_n \subset \mathbb{R}$*

$$P\{W(t_1) \in A_1, \cdots, W(t_n) \in A_n\} = \int_{A_1} \cdots \int_{A_n} p(t_1,0,x_1)p(t_2-t_1,x_1,x_2)\cdots p(t_n-t_{n-1},x_{n-1},x_n)dx_1\cdots dx_2,$$

where

$$p(t,w,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}$$

defined for any $x,y \in \mathbb{R}$ and $t > 0$ is called the transition density.

From Definition 20 the following theorem can be derived (stated here without proof) [82]:

Theorem 7 *A stochastic process $W(t)$, $t \geq 0$, is a Wiener process if and only if the following conditions hold:*

1. *$W(0) = 0$ almost surely;*

2. *the sample paths $t \to W(t)$ are almost surely continuous;*

3. *$W(t)$ has stationary (the distribution of $X(s+t) - X(s)$ does not depend on s for all $s,t > 0$) independent ($\mathbb{E}[(W(u) - W(t))(W(s) - W(r))] = 0$ for any $0 \leq r \leq s \leq t \leq u$) increments;*

4. *the increment $W(t) - W(s)$ has the normal distribution with mean 0 and variance $t - s$ for any $0 \leq s \leq t$.*

Definition 21 (Random step process) *A process $f(t), t \geq 0$ is a random step process if there is a finite sequence of numbers $0 = t_0 < t_1 < \cdots < t_n$ and square integrable random variables $\eta_0, \eta_1, \cdots, \eta_{n-1}$ such that*

$$f(t) = \sum_{j=0}^{n-1} \eta_j \mathbf{1}_{[t_j, t_{j+1}]}(t),$$

where η_j is $\mathscr{F}_{t_j}$-measurable for $j = 0, 1, \cdots, n-1$. The set of random step processes will be denoted by M^2_{step}. In Equation A.3, $\mathbf{1}(t)$ represents the indicator function.

Definition 22 (Stochastic integral of a random step process) *The stochastic integral of a random step process $f \in M^2_{step}$ is defined by*

$$I(f) = \sum_{j=0}^{n-1} \eta_j (W(t_{j+1}) - W(t_j)).$$

Definition 23 *M^2 is the class of stochastic processes $f(t)$, $t \geq 0$ that satisfy*

$$\mathbb{E}\left[\int_0^\infty |f(t)|^2\right] < \infty$$

and

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^\infty |f(t) - f_n(t)|^2\right] = 0. \tag{A.3}$$

In this case the sequence of step processes $f_1, f_2, \cdots$ approximates f in M^2.

Definition 24 (Itô stochastic integral) *$I(f) \in L^2$ (L^2 is the space of square integrable random variables) is called the Itô stochastic integral [from 0 to ∞] of $f \in M^2$ if*

$$\lim_{n\to\infty} \mathbb{E}[|I(f) - I(f_n)|^2] = 0$$

for any sequence $f_1, f_2, \cdots \in M^2_{step}$ of random step processes that approximates f in M^2 (i.e. such that Equation A.3 is satisfied). $I(f)$ and $\int_0^\infty f(t)dW(t)$ are interchangeable.

A.2 GAUSSIAN QUADRATURE

The Gauss-Legendre quadrature rule is formally expressed as [40]

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2} \sum_{i=1}^{n} w_i f\left(\frac{b-a}{2} z_i + \frac{a+b}{2}\right),$$

where z_i is the i-th root of the Legendre polynomial $P_n(z) = \frac{1}{2^n n!} \frac{d^n}{dz^n}\left[(z^2-1)^n\right]$ and

$$w_i = \frac{2}{\left(1 - z_i^2\right)\left[P_n'(z_i)\right]^2}.$$

A.3 CHOLESKY FACTORISATION

If $\mathbf{A}$ has real entries and is symmetric ($\mathbf{A} = \mathbf{A}^T$) and positive definite ($\mathbf{z}^T\mathbf{A}\mathbf{z}$ is positive, for any column vector $\mathbf{z}$), then $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{L}\mathbf{L}^*,$$

where $\mathbf{L}$ has positive diagonal entries and is a lower triangular matrix, and $\mathbf{L}^*$ is equal to the conjugate transpose of $\mathbf{L}$. Writing $\mathbf{A}$ as the product $\mathbf{L}\mathbf{L}^*$ is known as Cholesky decomposition [40].

The Cholesky algorithm, used to calculate the decomposition matrix $\mathbf{L}$ is described next. The recursive Cholesky algorithm starts by setting $i = 1$ and $\mathbf{A}(1) = \mathbf{A}$. At step i, the matrix $\mathbf{A}(i)$ then has the following form:

$$\mathbf{A}^{(i)} = \begin{pmatrix} \mathbf{I}_{i-1} & 0 & 0 \\ 0 & a_{i,i} & \mathbf{b}_i^* \\ 0 & \mathbf{b}_i & \mathbf{B}^{(i)} \end{pmatrix},$$

where $\mathbf{I}_{i-1}$ is the identity matrix of dimension $i-1$.

If matrix $\mathbf{L}_i$ is defined as

$$\mathbf{L}_i = \begin{pmatrix} \mathbf{I}_{i-1} & 0 & 0 \\ 0 & \sqrt{a_{i,i}} & 0 \\ 0 & \frac{1}{\sqrt{a_{i,i}}}\mathbf{b}_i & \mathbf{I}_{n-i} \end{pmatrix},$$

then $\mathbf{A}(i)$ can be written as

$$\mathbf{A}^{(i)} = \mathbf{L}_i\mathbf{A}^{(i+1)}\mathbf{L}_i^*$$

where

$$\mathbf{A}^{(i+1)} = \begin{pmatrix} \mathbf{I}_{i-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbf{B}^{(i)} - \frac{1}{a_{i,i}}\mathbf{b}_i\mathbf{b}_i^* \end{pmatrix}.$$

Note that $\mathbf{b}_i\mathbf{b}_i^*$ is an outer product. If the above is repeated enough times then at step n (which is also the dimension of the matrix $\mathbf{A}$), $\mathbf{A}(n+1) = \mathbf{I}$. Hence, the lower triangular matrix $\mathbf{L}$ is equal to

$$\mathbf{L} = \mathbf{L}_1\mathbf{L}_2\ldots\mathbf{L}_n.$$

A.4 LAGRANGE MULTIPLIERS

The method of Lagrange multipliers is employed to solve the following optimization problem:

$$\begin{aligned} &\max_{x,y} f(x,y) \\ &\text{s.t. } g(x,y) = c, \end{aligned} \tag{A.4}$$

where f and g are functions that have continuous first order partial derivatives and c is a constant. The Lagrange function is derived from $f(x,y), g(x,y), c$ and a new variable λ (the Lagrange multiplier) and is defined as

$$\mathscr{L}(x,y,\lambda) = f(x,y) + \lambda \cdot \Big(g(x,y) - c\Big).$$

If x_0 and y_0 are solutions of Equation A.4, then there exists an λ_0 such that (x_0, y_0, λ_0) is a stationary point of $\mathscr{L}(x,y,\lambda)$ [40].

A.5 KERNEL DENSITY ESTIMATION

If $(x_1, x_2, \cdots, x_n)$ are i.i.d. samples drawn from a distribution with an unknown density f, then the kernel density estimator of f is

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) \quad = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

where $K(x)$ is a symmetric but not necessarily positive function that integrates to one (and is known as the kernel), $h > 0$ is called the bandwidth (functions as a smoothing parameter) and $K_h(x) = 1/hK(x/h)$ (and is known as the scaled kernel). If Gaussian kernels are used to approximate univariate data, and the underlying density being estimated is itself a Gaussian then Silverman's rule of thumb is the optimal choice of h. Silverman's rule of thumb is:

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}n^{-1/5},$$

where $\hat{\sigma}$ is the standard deviation of the samples $(x_1, x_2, \cdots, x_n)$. The multivariate case is approached in a similar way [195].