

## Clustering and Classification of Cotton Lint Using Principle Component Analysis, Agglomerative Hierarchical Clustering, and K-Means Clustering

Edwin Kamalha <sup>a</sup>, Jovan Kiberu Ildephonse Nibikora<sup>b</sup>, Josphat Igadwa Mwasiagi<sup>c</sup> & Edison Omollo<sup>d\*</sup>

<sup>a</sup>Busitema University .Tororo <sup>b</sup>University of Science and Technology France <sup>c</sup>Politecnico di Torino, Moi University <sup>d\*</sup>Department of Fashion and Textiles Technical University of Kenya

### ABSTRACT

Cotton from the three cotton growing regions of Uganda was characterized for 13 quality parameters using the High Volume Instrument (HVI). Principal Component Analysis (PCA), Agglomerative Hierarchical Clustering (AHC) and k-means clustering were used to model cotton quality parameters. Using factor analysis, cotton yellowness and short fiber index were found to account for the highest variability. At 5% significance level, the highest correlation (0.73) was found between short fiber index and yellowness. Based on Cotton Outlook's world classification and USDA Standards, the cotton under test was deemed of high and uniform quality, falling between *Middling* and *Good Middling* grades. Our suggested classification integrates all lint quality parameters, unlike the traditional methods that consider selected parameters.

Journal of Natural Fibers Vol.15 (3) pp. 425-435 (2018 )

See more at:

<https://www.tandfonline.com/doi/full/10.1080/15440478.2017.1340220?scroll=top&needAccess=true>