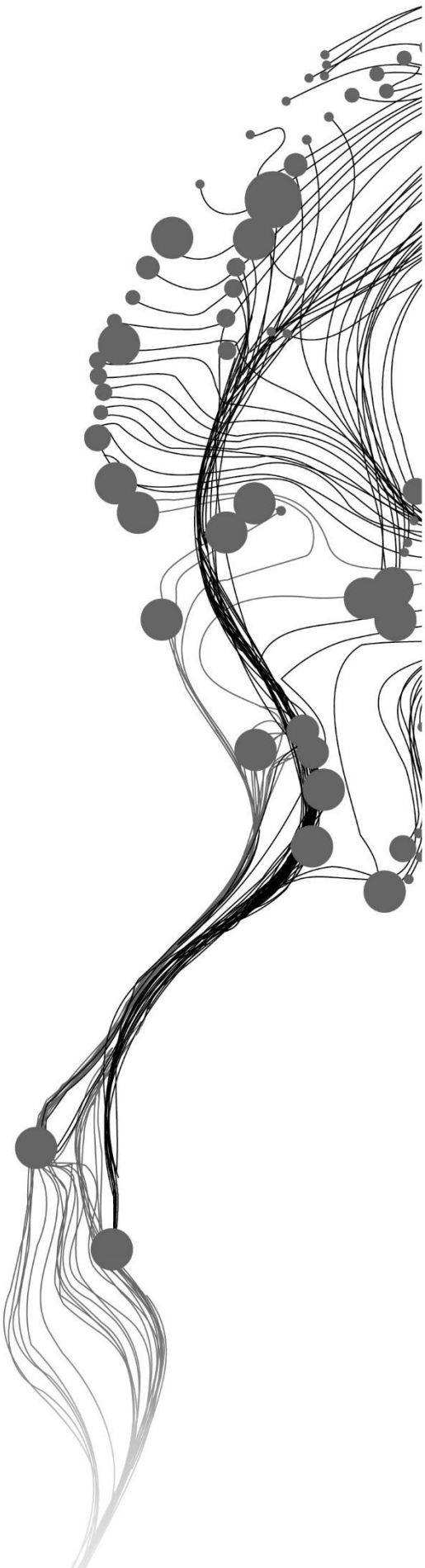
ASSESSING TEMPORAL INTERPOLATION TECHNIQUES TOWARDS GAP-FILLING CLOUDY SENTINEL-2 IMAGES

KELVIN JOSEPH KAMNDE
March, 2017

SUPERVISORS:
V. Venus (Valentijn)
Ing, G, Ashiagbor (George)
Dr Ir. C.A.J.M. de Bie (Cees)



KELVIN JOSEPH KAMNDE
Enschede, The Netherlands, March 2017

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:
V. Venus (Valentijn)
Ing, G, Ashiagbor (George)
Dr. Ir. C.A.J.M. de Bie (Cees)

THESIS ASSESSMENT BOARD:
Prof. Dr. A.D. Nelson (Chair)
Dr. B. Kumi-Boateng, University of Mines & Technology

ASSESSING TEMPORAL INTERPOLATION TECHNIQUES TOWARDS GAP-FILLING CLOUDY SENTINEL-2 IMAGES

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Sentinel-2 is the currently launched advanced revolution in medium resolution optical multispectral sensor. This satellite is important due to its advantageous characteristics such as 10m spatial resolution in visible near infrared bands, 5 days revisit time, richness in spectral bands especially four red-edge bands for monitoring of vegetation and its ability to provide important information for operational and monitoring of land and coastal processes. More importantly, all individuals and entities worldwide have free access to download Sentinel-2 images. These predict the future that Sentinel-2 will be most dependable source of satellite imageries. However, Sentinel-2 is highly affected by cloud cover that causes a lot of missing information and deterioration of information continuously. Consequently, this study recognized the importance of Sentinel-2 images and the challenges of clouds to the users. Therefore, the study offered the methods to gap-fill the information lost due to cloud cover. The study was successful to remove the clouds in Sentinel-2 images by modifying the F-mask algorithm for Landsat 8 images and used to detect clouds in Sentinel-2 images. Similarly, the study was also successful in filling the gap caused by clouds. This was done through the application of temporal interpolation algorithms Nearest neighbour, linear and Gaussian interpolations. Where by Gaussian interpolation performed better than other interpolators by attaining the general gap-filling accuracy of 90.98% and low RMSE of 0.023. Followed by linear interpolation with accuracy of 90.95% and RMSE of 0.024 and Nearest neighbour backward with accuracy of 90.29% and RMSE 0.262 and Nearest neighbour forward with accuracy of 89.83% and RMSE of 0.072. The significant difference test found that, only Nearest neighbour forward was statistically significant to other interpolators used, therefore in general gap-filling the Gaussian, linear and Nearest neighbour backward can be used for gap-filling the sentinel-2 images since there is no accuracy difference exist between them. In gap-filling the crop farms according to the different growth rates, the study found that, in Linear growth rates, the linear interpolation outperformed other interpolators by accuracy of 97.22% and RMSE of 0.010. However, significant difference test found that, there is no significant difference in accuracy between the Gaussian and linear interpolation. Therefore it can be used to gap-fill the linear growth rates. In exponential growth rates the Gaussian interpolation out performed other interpolators by accuracy of 96.38% and RMSE 0.011. However, significant difference test found that there is no significant difference in accuracy between the Gaussian and linear interpolation. Therefore it can be used to gap-fill the exponential growth rates. In Stationary growth rates the Nearest neighbour out performed other interpolators by accuracy of 92.32% and RMSE 0.024 for backward and accuracy of 92.03% and RMSE of 0.025 for forward. However, significant difference test found that there is no significant difference in accuracy between all interpolators used. Thus in gap-filling the stationary growth rates, all interpolations can be used.

Keywords: Sentinel-2, Gap-filling, Interpolation, Temporal Interpolation, Missing Information, Gaussian Interpolation, Linear Interpolation and Nearest Neighbour Interpolation.

ACKNOWLEDGEMENTS

I would like to give thanks and praise to the name of GOD Almighty for giving me health and strength until this end time of MSc programme. I thank the Netherland universities foundation for international cooperation (NUFFIC) for providing me with financial support throughout my studies.

I am deeply indebted to my supervisors; Valentijn Venus, Ing. George Ashiagbor , Dr Ir. C.A.J.M. de Bie and ; Ms. Ir. Louise van Leeuwen for the guidance and kind supervision of my thesis. l am also grateful to Mr. Alex Gaisie for teaching me Python programming language that helped to achieve the main objective of this study. Not forgetting Dr Collins Nsor (Lecturer KNUST) for his time and devotion during my data collection and Alex Abavana who assisted me during my field work data collection.

The brotherly and sisterly advice from Violeth Charles Swai, Ekow Nyamekye Tawiah, Emma Baah Agyampong, and all my relatives drew me near to GOD in situations of absolute discomfort. Jacob Onyimbi, My fellow Tanzanians and Ghanaians students at ITC and KNUST thank you for your unity and brotherly love.

Lastly is my thanks to all teachers at KNUST and ITC for international knowledge I acquired from you. I promise to be part of the positive changes with this great knowledge.

"Dedicated to my Late Mother Magreth Elia Uronu, my source of encouragement and motivation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUTIONS

LIST OF APPENDICES

LIST OF ACRONYMS

ACCA	Automatic Cloud Cover Assessment
ANOVA	Analysis Of Variance
AVHRR	Moderate Resolution Imaging Spectral-Radiometer
DEM	Digital Elevation Model
DN	Digital Numbers
ENVI	Environment for Visualizing Images
ERDAS	Earth Resource Data Analysis System
ETM	Enhanced Thematic Mapper
ESA	European Space Agency
GI	Gaussian Interpolation
GLCM	Grey Level Co-Occurrence Matrix
GLDV	Grey Level Different Vector
GLDM	Grey Level Different Matrix
GMES	Global Monitoring For Environment And Security
ITC	International Institute for Geo-information Science and Earth Observation
GPS	Global Positioning System
KNUST	Kwame Nkrumah University of Science and Technology
LAI	Leaf Area Index
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LI	Linear Interpolation
LOOCV	Leave One Out Cross Validation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLST	Mean Local Solar Time
MSE	Mean Square Error
MODIS	Moderate Resolution Imaging Spectral-Radiometer
MSI	Multispectral Instrument
NAN	Not A Number
NDSI	Normalized Different Snow Index
NDVI	Normalized Different Vegetation Index
NDWI	Normalized Different Water Index
NNBI	Nearest Neighbour Backward Interpolation
NNFI	Nearest Neighbour Forward Interpolation
SADH	Sum And Different Histogram
SPSS	Statistical Package for Social Science
SRTM	Shuttle Radar Topography Mission
SWIR	Short Wave Infrared
TOA	Top Of Atmosphere
VHR	Very High Resolution
UTM	Universal Transverse Mercator

1. INTRODUCTION

1.1. Background

Optical satellite imagery captures information from visible (0.4 μm) to thermal infrared (15 μm) portions of the electromagnetic spectrum are therefore, important in studying earth processes (Irea, 2016). Earth processes such as climate change, water pollution, melting of glaciers, desertification, vegetation growth, forest fires, land cover status and changes can be observed and monitored through optical remote sensing (Irea, 2016; J & Jojy, 2015; Segl et al., 2012; Zhu et al., 2015).

However, as optical satellite imagery continuously observe earth's phenomenon, portions of the spectrum fails to penetrate some atmospheric components such as clouds (PennState Astro 801, 2016). Clouds are great obstacle of optical remote sensing. It is estimated that, more than 66% of the earth surface is covered by clouds throughout the year. In the equatorial regions, the average cloud cover extends more than 75% annually (Arellano, 2003; Wang et al., 2009; Zhang et al., 2004; Zhu & Woodcock, 2012). In north-western Europe, climatic statistics reveals 40% cloud cover, during the clouded duration of the year (Arellano, 2003). Also in Canada, the cloud cover over land is 50% and it usually occur in the mid-mornings subject to the latitude of the area (Cihlar & Howarth, 1994).

The presence of clouds causes information lost and contributes to the deterioration in data quality of optical remote sensors e.g. Sentinel-2 (Eckardt et al., 2013). This problem brings to fore the challenges in extraction of meaningful information from satellite images hence rendering such images useless (Eckardt et al., 2013; Zhu & Woodcock, 2012).

Sentinel-2 is one of the advanced new revolution in medium resolution optical multispectral sensors. It is a very important multispectral satellite for operational and monitoring of land and coastal areas (Baillarin et al., 2012). Due to its wide field of view and altitude, it is highly affected by cloud cover, especially, for images captured around the tropics (Arellano, 2003; Wang et al., 2009; Zhang et al., 2004; Zhu & Woodcock, 2012). Clouds acts as the main causative of missing information in Sentinel-2. Clouds are greater setback for Sentinel-2 mission in operational and monitoring of land and coastal areas. This draws attention for the application of scientific approaches to remove clouds and gap-filling of the missing information.

1.2. Statement of the problem

Sentinel 2A was launched in 23rd June 2015 and 2B will be launched in 2017 (Baillarin et al., 2012; Mandanici & Bitelli, 2016). Both Sentinel 2A and Sentinel 2B contains identical Multispectral Instruments (MSI) which captures information in 13 bands with spatial resolution range between 10m and 60m (Baillarin et al., 2012; Mandanici & Bitelli, 2016). Together they will have high temporal resolution of 5 days, wide field of view of 290 km, orbital altitude of 786 km, orbit inclination is 98.62° and Mean Local Solar Time (MLST) of 10:30 (am) (Baillarin et al., 2012; Mandanici & Bitelli, 2016). Moreover, Sentinel-2 is non-commercial and available for free downloads to all individuals and entities globally. From the stated characteristics, it is obvious that, Sentinel-2 is an important multispectral satellite for operational and monitoring of land and coastal areas (Baillarin et al., 2012).

However, due to its wide field of view, high revisit time, medium spatial resolution and altitude, Sentinel-2 is highly affected by cloud cover. Clouds are the main causes of missing information in Sentinel-2. The rate of sentinel-2 images affected by clouds depends on the spatial distribution of cloud over space e.g. images that captured around tropical areas are highly affected by clouds (Arellano, 2003; Wang et al., 2009; Zhang et al., 2004; Zhu & Woodcock, 2012). The occurring of clouds in Sentinel-2 defeats its purpose of earth observation and monitoring of land and coastal areas. Hence, a problem that needs to be solved. Figure 1-1, shows the clouds coverage in the study area, these images brings imagination on amount of information lost due to cloud cover.

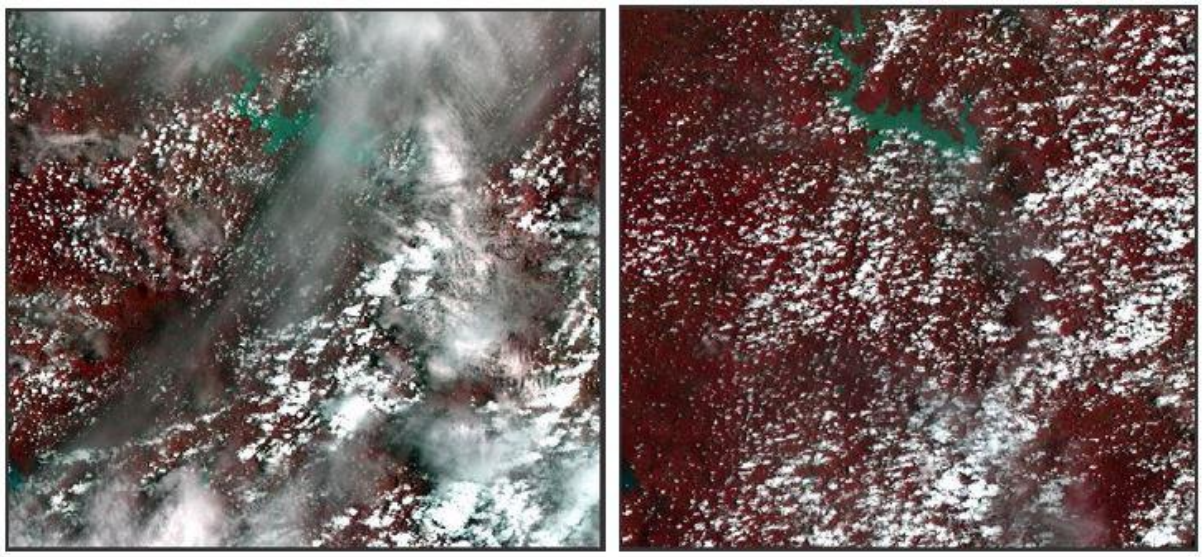


Figure 1-1 : Effects of cloud cover in Sentinel-2 Images

1.3. Research aims and objective

1.3.1. Main Aim

The aim of this research is to develop a methodology that will help to gap-fill the missing information due to cloud cover in Sentinel-2 images.

To achieve this aim, the following specific objectives are outlined.

1.3.2. Specific Objectives

i. To gap-fill the missing information in Sentinel-2 images using temporal interpolation algorithms and assessing their accuracies in gap-filling process.

ii. To assess the accuracy of interpolators in gap-filling of missing information in different crop growth rates.

1.4. Research Questions

i. Which interpolation technique performs better in gap-filling of missing information in sentinel-2 images?

ii. Which interpolation technique out perform others in gap-filling of missing information in terms of linear crop growth rates, stationary crop growth rates and exponential crop growth rates?

1.5. Hypotheses

i. **H_0:** There is no significant different in accuracies of interpolators in gap-filling Sentinel-2 images

H_1: *There is significant different in accuracies of interpolators in gap-filling Sentinel-2 images*

ii. **H_0:** There is no significant different in accuracies of interpolators in gap-filling the linear crop growth rates.

H_1: *There is significant different in accuracies of interpolators in gap-filling the linear crop growth rates.*

iii. **H_0:** There is no significant different in accuracies of interpolators in gap-filling the stationary crop growth rates.

H_1: *There is significant different in accuracies of interpolators in gap-filling the stationary crop growth rates.*

iv. **H_0:** There is no significant different in accuracies of interpolators in gap-filling the exponential crop growth rates.

H_1: *There is significant different in accuracies of interpolators in gap-filling the exponential crop growth rates*

2. LITERATURE REVIEW

2.1. Review for cloud removal and information Gap-filling in optical satellite imageries

Optical satellite imagery, they captures information from visible 0.4 micrometres to thermal infrared 15 micrometres portion of electromagnetic spectrum, are very important in studying earth processes (Irea, 2016). Very important processes like climate change, water pollution, melting of glaciers, desertification, vegetation growth, forest fires, land cover status and changes can be observed, captured and stored through optical remote sensing (Irea, 2016; J & Jojy, 2015; Segl et al., 2012; Zhu et al., 2015).

As optical imagery being appreciated for continuously earth observation, part of its wavelength fails to penetrate on atmospheric components, one of them is clouds (PennState Astro 801, 2016). By way of optical remote sensing and electromagnetic spectrum, the visible light covers small portion from 0.4 to 0.7 micrometres, while the average size of cloud droplet is 10 micrometres (CMMAP, 2016). As the cloud droplet is bigger than any portion of visible light, when it meets cloud droplets it doesn't penetrating but it scatters equally to all direction which appears and recorded as white features in optical satellite imagery (CMMAP, 2016).

Clouds are great obstacle of optical remote sensing. It has been estimated that, more than 66% of the earth surface is covered by clouds throughout the year more in equatorial region where the average cloud cover is more than 75% annually (Arellano, 2003; Wang et al., 2009; Zhang et al., 2004; Zhu & Woodcock, 2012). While in north-western Europe climate statistics shows that, during the slightest clouded duration of the year the cloud cover percentage is 40%(Arellano, 2003). Furthermore, in Canada the estimates states that the cloud cover over land is 50% and found more than 80% in the mid-morning subjected to an area latitude (Cihlar & Howarth, 1994).

Presence of clouds causes loss of important information and contributes to the deterioration of data quality captured by the optical sensors (Eckardt et al., 2013). Due to the circumstance it brings challenges in extraction of meaningful information and cause the images captured to be of non-useful (Eckardt et al., 2013; Zhu & Woodcock, 2012). The presence of clouds in optical images causes failure to perform different imagery analysis. Even when the analysis performed in images with less clouds causes biased and incorrect estimations such as mistakes and false detections in classification of land cover, NDVI and land cover changes (Growth & Stein, 2015; Zhu & Woodcock, 2012).

As explained how the clouds causes loss of information and challenges in using of optical images, the situation brought attention to science community that paved a way to think and develop several methodologies and approaches to remove clouds and gap-filling the missing information caused by clouds existence.

Detection of clouds and cloud shadows is very difficult task because of similarities in reflectance of clouds, cloud shadows and surface features (Hagolle et al., 2010; Koren, 2016.; Zhu et al., 2015; Zhu & Woodcock, 2012). Example differentiation of clouds to snow, white sands and white roofs is difficult because they are similar in reflectance recorded in an image, also identifying cloud shadows over water bodies is difficult because they have similar reflectance in an image (Amin et al., 2011; Koren, 2016; Zhu et al., 2015; Zhu & Woodcock, 2012)

Initial campaigns for cloud detection, it started with traditional methodologies which involves both spectral and spatial procedures to detects clouds in the optical satellite imageries (Jojy, 2015; Tian et al. , 1999). The methodologies were manually performed by using algorithms which used to thresholds test of several spectral bands and identify the one which captures better the surface and atmospheric effects and set fixed threshold for cloud detection (Jojy, 2015; Tian et al., 1999). Therefore, detecting clouds by using manual interpretations it was difficult, time consuming and results were biased, this situation of biasness and other challenges opened the room for the development of automated and user friendly algorithms for processing and screening of clouds in huge amount of optical satellite imageries (Jojy, 2015; Tian et al., 1999; Zhu & Woodcock, 2012)

The development of automated algorithms for cloud detections deals with the cloud classification area which depends on cloud features such as spectral and textural features (Jojy, 2015; Tian et al., 1999). The spectral features are mostly used for information extraction of cloud radiance in different spectral bands (Jojy, 2015; Tian et al., 1999). To extract information on how cloud radiance varies in different spectral bands uses methods like histogram scheme, threshold based scheme, multispectral approaches and also spectral features which related to their physical characteristics such as albedo and temperature (Jojy, 2015; Tian et al., 1999). Spectral features in cloud classification are confirmed to be simple and successful though it encounters challenges in cloud detection due to spectral similarities with surface features like snow, ice, white sand and roofs (Jojy, 2015; Tian et al., 1999). Also presence of moisture in atmosphere may affect the multispectral characteristics and affects final cloud detections (Jojy, 2015; Tian et al., 1999).

Turning to the cloud classification by textural features, this differentiates types of clouds depending on their spatial distribution characteristics of grey levels (Jojy, 2015; Tian et al., 1999). Textural features are recommended compared to spectral features, because the spectral characteristics of clouds may change while, the textural characteristics are clearly different and be likely less affected by detector noise and atmospheric attenuation (Jojy, 2015; Tian et al., 1999). Furthermost textural based cloud classification techniques using statistical computation based on grey level co-occurrence matrix (GLCM) and its variant like grey level different matrix (GLDM), sum and difference histogram (SADH) and grey level different vector (GLDV)(Jojy, 2015; Tian et al., 1999).

All of the methodologies developed for cloud removal tried to remove clouds and acquired different accuracies. After removing clouds, what remains in images, is pixels with no information, and that is what referred as the missing information. This situation calls for attention that the pixels with no values in cloud masked images should be gap-filled with information so as can be used to perform different analysis, to fulfil the objective of image capturing.

Gap-filling the missing information in satellite images involves development of different methodologies, ranges from methods that depends on only spatial information available, methods which depends only on temporal information available from image time-series and the methods which comprise both spatial and temporal in the information gap-filling process (Weiss et al., 2014).

Spatial methodologies for Gap-filling the missing information, incorporates geo-statistical approaches example long time used kriging approach for image Gap-filling, where by it uses the information available in surrounding pixels (pixels with information) and interpolate it to the pixels with no information (Addink, 1999; Kandasamy et al., 2013; Weiss et al., 2014). In addition, there is Co-kriging approach that uses the image of the same sensor of the area of interest, which acquired at different date to Gap-filling the no values pixels. it look like temporal

approach but (Zhu et al., 2012) declared it was developed by geo-statistical theory and fall under spatial approach (Weiss et al., 2014; Zhang et al., 2009; Zhu et al., 2012)

Temporal methodologies for Gap-filling the missing information in satellite images models the situation whereby missing pixels are Gap-filled with different pixels of different point in time (Jönsson & Eklundh, 2004; Weiss et al., 2014) . Example Jönsson & Eklundh (2004) they developed TIMESAT software which uses temporal approach to Gap-filling (smooth) time-series images by built-in filters such as Savitzky-Golay and asymmetric Gaussian filters (Jönsson & Eklundh, 2004; Weiss et al., 2014). Many more application of temporal gap filling has been used in hyper-temporal imageries such MODIS (Kandasamy et al., 2013; Moody et al., 2005).

Spatial-temporal approaches for Gap-filling the missing information, is developed by modelling the Gap-filling process to involve both spatial and temporal approach in Gap-filling the information in images (Kandasamy et al., 2013; Weiss et al., 2014). Example Kang et al. (2005) they develop spatial-temporal approach to Gap-filling missing information in ecosystem metrics such as LAI, net photosynthesis and fPAR from MODIS data (Kang et al., 2005; Weiss et al., 2014). They developed their algorithm in such a way that, it starts with spatial Gap-filling by filter of 5 by 5-pixel window filters the image. After filtering if still there are no values in pixels it proceeding by using the temporal approach by taking the values of different pixels of different point in time to fill the no value pixels (Kang et al., 2005; Weiss et al., 2014). Many more application of spatial-temporal gap filling has been used in hyper-temporal imageries such MODIS (Kandasamy et al., 2013; Moody et al., 2005; Moreno et al., 2014; Yin et al., 2016)

As from the review, every published methodology for both cloud removals in optical images and methodologies for Gap-filling the missing information acquired significant results. This depending on the type of images and data used. Because different satellite images have different spectral bands and different approaches used in both cloud removal and Gap-filling of missing information.

2.2. Sentinel-2

This is the Satellite developed by ESA in partnership with the European Commission under GMES program (Baillarin et al., 2012). Involves the constellation of two Satellites Sentinel 2A which was launched 23th June 2015 and 2B will be launched in 2017, the mission for Sentinel-2 installation is to "provide high resolution optical imagery for operational monitoring of land and coastal areas" (Baillarin et al., 2012; ESA, 2015). Sentinel-2 provides very important information in vegetation parameters such as leaf chlorophyll content, leaf water content and leaf area index which are very important parameters in identification, mapping, predicting crop yield and harvesting time of food crops (Baillarin et al., 2012).

Furthermore, Sentinel-2 is very powerful due to its high spatial resolution and multi-spectral information with 13 bands (Baillarin et al., 2012). The visible and near infrared bands have 10 m spatial resolution for land applications. 4 vegetation red edge bands has 20m spatial resolution and 2 SWIR bands which have 20m spatial resolution for snow, ice, cloud detection and vegetation moisture stress assessment, 3 bands of 60m spatial resolution for atmospheric correction and cirrus detection (Baillarin et al., 2012; ESA, 2015). In addition, it has high temporal resolution revisit in every 5 days with two satellites with wide field of view 290km. Systematically it captures the information covers all coastal and land areas between 84°N and 56°S and crossing the equator at 10:30 am the local time (Baillarin et al., 2012; Eox, 2016; ESA, 2015; Mandanici & Bitelli, 2016)

Figure 2-1, shows both spatial and spectral characteristics of Sentinel-2 images. This involves bands arrangements, their corresponding spatial resolution, and their use.

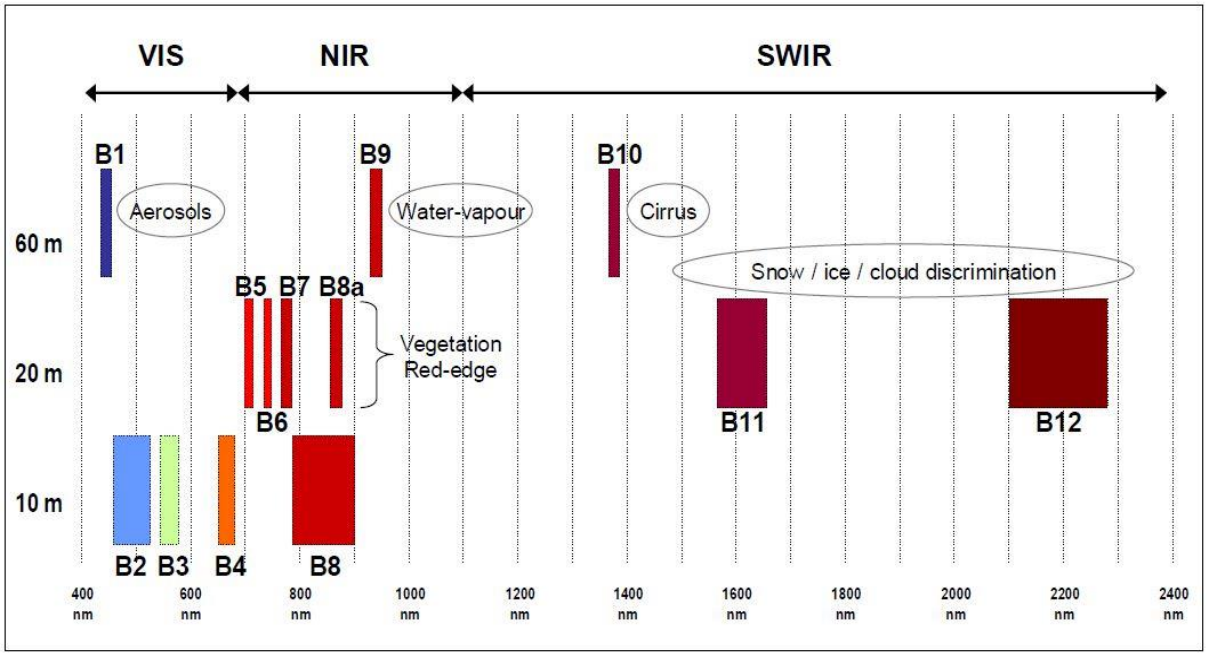


Figure 2-1: Sentinel-2 spectral and spatial resolution characteristics. Source: (Baillarin et al., 2012)

Sentinel-2 as the one of advanced medium resolution optical multispectral satellite, it's images can be used in vast range of applications from vegetation monitoring, crop management and monitoring, forest monitoring by calculating vegetation parameters like chlorophyll content, leaf area index, NDVI and carbon mass estimation (EOX, 2016; Surya & Simon, 2014). Also apart from vegetation, Sentinel-2 can be used in monitoring of coastal areas such as mapping of coastal resources and monitoring of marine environment, can be used to monitor land cover changes, glacier monitoring, inland water monitoring and mapping of snow cover and ice extent (EOX, 2016; Surya & Simon, 2014).

2.3. Sentinel-2 image preparation procedures for cloud removal

2.3.1. Ortho-rectification

Is the geometric correction whereby involve presenting the image as if was captured in Nadir or captured vertically (Baillarin et al., 2008). It corrects the terrain effects and effect of image tilt, it incorporates using of DEM and ground control points to correct the image. After Ortho-rectification the accurate measurement such as distance, angles and area can take place (OSSIM, 2016; Satpalda, 2016)

Figure 2-2, shows the Ortho-rectification process how it is conducted, where, the image with the terrain distortions and scale anomalies are corrected by using digital elevation model.

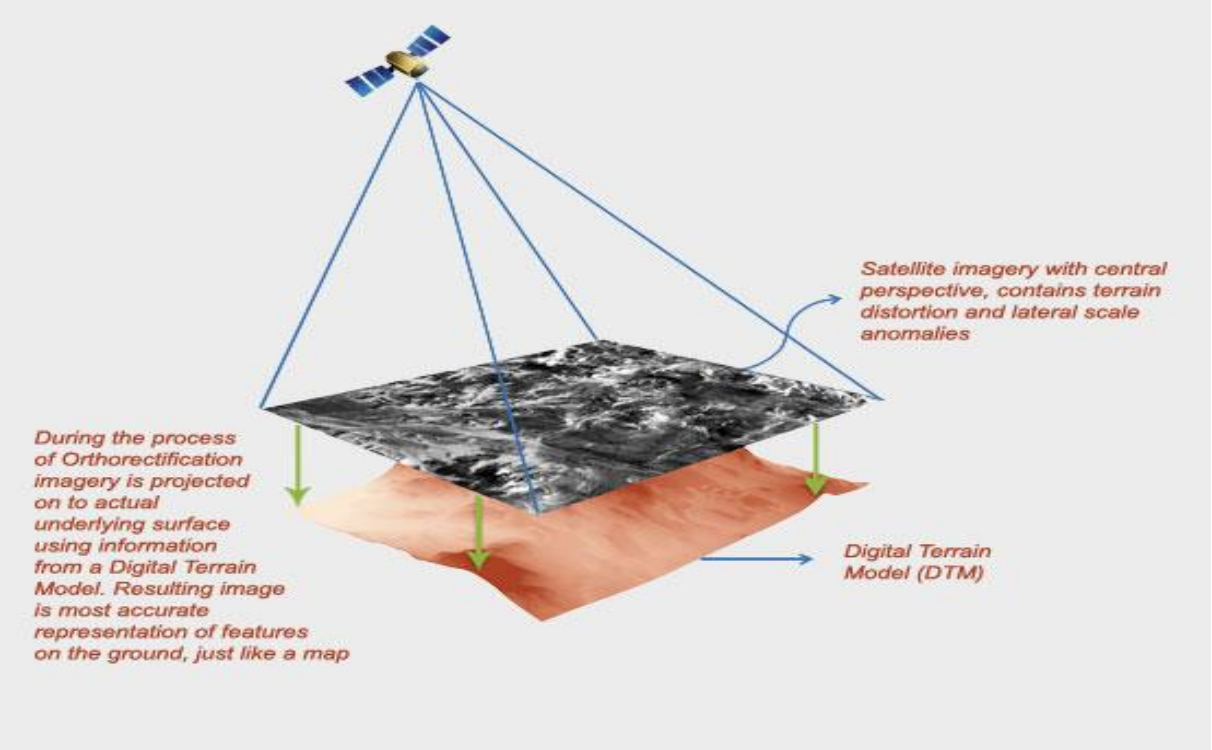


Figure 2-2: Image Ortho-rectification procedure. Source: (Satpalda, 2016)

2.3.2. Sensor calibration (DN values to Radiance)

Digital number (DN) is the common word used for pixel values that have not yet been changed into physically meaningful units (Harris, 2016). Therefore, changing DN to radiance in summary is changing raw DN values into the amount of energy that the sensor received. This involves changing of digital values which are in (Bits) to the physical meaningful unit which are proportional to up welling electro-magnetic radiation (radiance) and their units are in watt per steradia per square meter ($W \cdot sr^{-1} \cdot m^{-2}$), which represent amount of radiation come from an area (Harris, 2016; Kiyoshi, 2014; Turks, 1990). To perform the radiance conversion, it needs gain and off set values which are provided in image Meta data from the image provider. Also the method it is automated in remote sensing software such as ENVI (Harris, 2016) The Equation 2-1 is the equation that used for conversion of DN values to radiance, it uses offset and gain values from the metadata as input values.

Equation 2-1: DN to radiance conversion

$$\mathbf{L\lambda = "gain" * (DN) + "offset"}$$

Where:

$\mathbf{L_\lambda}$ = Radiance

2.3.3. Radiance to reflectance conversion

Reflectance

Is the proportion of the radiation reflected from the surface to the amount radiation hit the surface, this helps to identify surface features due to their spectral reflectance characteristics, so it is very important to convert the radiance to reflectance to understand surface feature characteristics from image (Harris, 2016; Kiyoshi, 2014; Turks, 1990). The Equation 2-2 is the equation that used for conversion of radiance to both top of atmosphere and surface reflectance.

Equation 2-2: Equation for computing surface and top of atmosphere reflectance

$$\rho_p = \frac{\pi \bullet L_\lambda \bullet d^2}{ESUN_\lambda \bullet \cos\theta_s}$$

Where:

ρ_p = Unit less planetary reflectance

L_λ = Spectral radiance at the sensor's aperture

d = Earth-Sun distance in astronomical units

$ESUN_\lambda$ = Mean solar exoatmospheric irradiances

θ_s = Solar zenith angle in degrees

2.3.3.1. Top of Atmosphere Reflectance named as (TOA) reflectance

Is the reflectance measured on top of atmosphere where the sensor is flying higher than the earth's atmosphere, and the reflectance values contains the contributions from atmospheric aerosols, gases and clouds and it is unit less planetary reflectance. Mostly is expressed in percentage and its value ranges from 0-1. (Harris, 2016; Kiyoshi, 2014) Also conversion of radiance to reflectance it is automated in remote sensing software such as ENVI but it is sensor specific (Harris, 2016).

2.3.3.2. Surface Reflectance

This is the reflectance at the surface of the earth where by all the atmospheric components such as aerosols, gases and clouds have been removed and do not add to the spectral reflectance, it is unit less planetary reflectance and mostly is expressed in percentage and Its value ranges from 0-1 (Harris, 2016).

2.4. Sentinel-2 Images processing levels

Detecting clouds in optical remote sensing is not an easy mission, it is very difficult to differentiate clouds and surface features such as snow, ice and sand (Griggin et al., 2003; Wang et al., 2009) Until now most of images have been supplied without corrections to remove haze and clouds (Hagolle et al., 2010)

Therefore, image-processing levels have been developed in order to indicate correction levels image attained during image supply. The production levels start from level 0 to level 3A (Hagolle et al., 2015; Piwowar, 2014).

2.4.1. Image Level 0

These are the images directly collected from the sensor, no any correction has attained and most suppliers are not supplying level 0 products (Piwowar, 2014).

2.4.2. Image Level 1A

Is the image whereby, some few corrections like detector equalization (radiometric correction) to correct problems caused by detector variation within the sensor. It is a basic product, provided with metadata which allow one to calibrate and Ortho-rectify images (Hagolle et al., 2015; Piwowar, 2014)

2.4.3. Image Level 1B

These are Ortho-rectified images whereby, Involves Ortho-rectification to correct the geometry distortions caused by sensor movement while was scanning the earth, the distortions corrected are Mis-aligned scan lines and non-uniform pixel size to improve geometric qualities (Piwowar, 2014).

2.4.4. Image Level 1C

Involves conversion of Ortho-rectified DN values images to TOA reflectance (Hagolle et al., 2015). Therefore, the Level 1C images are TOA reflectance images

2.4.5. Image Level 2A

These are Ortho-rectified surface reflectance images whereby clouds has been masked (Hagolle et al., 2015; Hagolle et al., 2016). Therefore, in level 2A the pixels where clouds has been masked out, remains with no values (Olivier Hagolle et al., 2015).

2.4.6. Image Level 3A

Expressed as composite Ortho-rectified surface reflectance images with no clouds (Hagolle et al., 2015; Inglada et al., 2015). The pixels where clouds removed, has been filled with values interpolated from the image composite. Therefore, there is no missing information and the image is good for extracting information and signal to noise ratio has been secured.

Figure, 2-3, shows the complete process for gap-filling, where level1C is the preparation for cloud removal. Level 2A shows the cloud has been removed ready for gap filling, and level 3A shows the gap in image has been filled.

Figure 2-3: Complete Gap-filling process. Source: (Hagolle et al., 2016)

2.5. Possible spectral cloud detection methodologies for Sentinel-2 images

Separating clouds from cloud free surface features captured by image based on reflectance from spectral bands it is bit difficult, especially thin clouds which their reflectance look similar to cloud free surface features beneath (Koren, 2016.; Zhu et al., 2015; Zhu & Woodcock, 2012)

Currently, most of the method developed for cloud detection are designed for moderate spatial resolution satellites such as Moderate Resolution Imaging Spectral-radiometer (MODIS) and Advanced Very High Resolution Radiometer (AVHRR) (Ackerman et al., 2008; Martins, 2002; Zhu et al., 2012; Zhu et al., 2015; Zhu & Woodcock, 2012). In these sensors cloud detection is of high accuracy, because the sensors are launched with more than one thermal band and water vapour or carbon dioxide absorption bands which helps to improve the detection of different types of clouds (Ackerman et al., 2008; Martins, 2002; Zhu et al., 2012; Zhu et al., 2015; Zhu & Woodcock, 2012). Example thermal bands helps to identify clouds that are colder than earth surface features, which in that way, it detect all types of clouds except the lower clouds which their temperature bit related to the surface temperature (Ackerman et al., 2008; Martins, 2002; Zhu et al., 2012; Zhu et al., 2015; Zhu & Woodcock, 2012). Therefore, detecting clouds in satellite images depends on the sensor characteristics especially spectral bands of the sensor

In order to identify methodology for cloud detection in Sentinel-2, one has to understand the characteristics of Sentinel-2. Figure 2-1 shows some spectral band characteristics of Sentinel-2 that can be studied for cloud detection in Sentinel-2 images.

Band	Band range (nm) / Band Center (nm)	Spatial Resolution (m)	Purpose in L2 processing context
B1	433-453 / 443	60	Atmospheric Correction
B2	458-523 / 490	10	Sensitive to Vegetation Aerosol Scattering
B3	543-578 / 560	10	Green peak, sensitive to total chlorophyll in vegetation
B4	650-680 / 665	10	Max Chlorophyll absorption
B5	698-713 / 705	20	Not used in L2A context
B6	734-748 / 740	20	Not used in L2A context
B7	765-785 / 783	20	Not used in L2A context
B8	785-900 / 842	10	Leaf Area Index (LAI)
B8a	855-875 / 865	20	Used for water vapor absorption reference
B9	930-950 / 945	60	Water vapor absorption atmospheric correction
B10	1365-1385 / 1375	60	Detection of thin cirrus for atmospheric correction
B11	1565-1655 /1610	20	Soils detection
B12	2100-2280 / 2190	20	AOT determination

Figure 2-4: Sentinel-2 spectral characteristics Source (EOX, 2016)

From the spectral characteristics of Sentinel-2, and as it misses thermal band it makes the cloud detection process bit challenging (Zhu et al., 2015). Therefore, possible methodologies can be applied for cloud detection

in Sentinel-2 ranging from thresholds the reflectance in blue band, thresholds the reflectance in SWIR cirrus band and multi-temporal cloud detection methods.

2.5.1. Threshold the reflectance in blue band and prevent false detection using other bands

If someone uses spectral characteristics of an image to detect clouds, usually the first identification step to take during differentiation of clouds and other cloud free surface features in an image, is the changes in reflectance values (ESA, 2016; GSU, 2016; Zhu et al., 2015; Zhu & Woodcock, 2012). The land surface reflectance usually changes slowly, therefore, when clouds happens, there is rapid increase in reflectance (Koren, 2016.; Zhu et al., 2015; Zhu & Woodcock, 2012).

For rapidness increase of reflectance for cloud identification can be detected well in blue band because of Rayleigh scattering where by surface features have low reflectance in blue except cloud and snow, therefore the threshold of the significant increase of reflectance in blue band improves well the separation of cloudy pixels from clear pixels (GSU, 2016). But the obstacle of threshold reflectance in blue band can identify only Dense clouds but not cirrus clouds, this is because the cirrus clouds are transparent or Semi-transparent (ESA, 2016).

Though detecting dense cloud depends on the blue band reflectance thresholds, false detections can be conducted, example cloud can be detected as snow because both their reflectance appears rapidly (ESA, 2016). Therefore, to prevent false detections probably snow cloud detection in Sentinel-2 SWIR band 11 and band 12 reflectance can be used because Sentinel-2 it lacks thermal bands (ESA, 2016). From the experiments they detected that, still clouds have high reflectance in SWIR band 11 and band 12 than the snow which have low reflectance in these bands (ESA, 2016).

Cirrus clouds characteristics, they are thin, transparent or semi-transparent they appear in high altitudes approximately 6-7 kilometres from the surface (ESA, 2016; Zhu et al., 2015). Therefore, detecting Cirrus cloud in Sentinel-2, band 10 can be used because it is high in absorption of atmospheric components consequently, only the clouds form at high altitude can be detected (ESA, 2016; Zhu et al., 2015). Also the cirrus clouds have low reflectance in blue band, but it has high reflectance in band 10, therefore the pixels with low reflectance in blue band and high reflectance in band 10 it has high probability to be cirrus clouds (Zhu et al., 2015). But the assumption has challenge whereby some dense clouds might have low reflectance in blue band and detected as cirrus cloud, therefore to prevent false detection, it needs morphological-based operation filter to be applied during detection of clouds and differentiate between dense and cirrus cloud (ESA, 2016)

2.5.2. Multi-Temporal Cloud Detection (MTCD) Method

This is the method for cloud detection which relies on the multi-dated images for cloud detection, it also relies on the increase reflectance in blue spectral band (Hagolle et al., 2010). In order to detect clouds, it needs an image that is cloud free to detect the pixel where the cloud is available. Cloud free image used as reference, but to obtain the full cloud free image is impossible, so it compositing the near dated images and using the values of the cloud free pixels to detected cloud on the image which will be corrected (Hagolle et al., 2010). Equation 2-3 is the equation used in multi-temporal cloud detection method to detect cloud from images that captured in different dates.

Equation 2-3: Multi temporal cloud detection

$$[Pblue(D) - Pblue(D_r)] > 0.03 * (1 + (D - D_r)/30)$$

Where, $Pblue(D)$= Pixel reflectance in blue band at date D corrected for Rayleigh scattering

D_r= Most recent cloud free date before date D

$D - D_r$ = Difference dates expressed in days

To detect clouds for image date D, the image Dr should be acquired before the date D. and the threshold of 0.03 will be used, this depending on the number of the days between D and Dr. 0.03 is the threshold when the difference is 30 days and it double when it is more than 30 days to allow changes in surface reflectance (Hagolle et al., 2010; Tobergte & Curtis, 2013)

2.5.3. F-mask algorithm

One challenge of thresholds blue band reflectance, is failure of detection or false detection when the surface is bright, therefore, as Sentinel-2 it lacks thermal band, this part discusses how Cirrus band can be used to detect clouds in Sentinel-2 (Zhu et al., 2015).

The algorithm developed by Zhu & Woodcock (2012). This algorithm developed first for cloud detection in Landsat images TM and ETM+. To use the algorithm, it needs to convert band 1,2,3,4,5 and band 7 to TOA reflectance and band 6 to brightness temperature(Zhu et al., 2015; Zhu & Woodcock, 2012). Then it extracts cloud physical properties from the inserted bands to detect clouds and cloud shadows (Zhu et al., 2015; Zhu & Woodcock, 2012). After the clouds has been detected it uses the geometrical relationship to match cloud and its possible shadow and finally it produces the layer which detected both clouds and its possible shadows(Zhu et al., 2015; Zhu & Woodcock, 2012). The algorithm uses the following formula which thresholds the inserted TOA bands and brightness temperature for Landsat images.

Equation 2-4: F-mask for cloud detection in Landsat images

Cloud pixel = Band 7 > 0.03 and BT < 27 and NDSI < 0.8 and NDVI < 0.8

Where by,

NDSI= (Band 2−Band 5)/ (Band 2+Band 5)

NDVI= (Band 4−Band 3)/ (Band 4+Band 3).

The equation explains that for all type of clouds will be identified should have the reflectance in band 7 TOA should be greater than 0.03 this threshold has been taken from LEDAPS, and Brightness temperature should be less than 27 °C this adopted from ACCA (Zhu et al., 2015; Zhu & Woodcock, 2012). Also NDSI and NDVI should be less than 0.8 this is because of thick clouds around forested or vegetated areas, likewise for differentiation between clouds and snow(Zhu et al., 2015; Zhu & Woodcock, 2012). Therefore, for cloud detection F-mask achieved the overall accuracy of 96.41% with standard deviation 3.2% which is very small (Zhu et al., 2015; Zhu & Woodcock, 2012).

Zhu et al (2015) they expanded the algorithm for Landsat 8 and possibly here it will be tested for Sentinel-2. The expanded version still using the same TOA bands, Brightness temperature as was in past F-mask, the improvement here is the using of Cirrus band (Zhu et al., 2015). The cirrus band is applied to improve the identification of cloud pixels (Zhu et al., 2015). In order to use the cirrus band, it needed to be converted to TOA and if the reflectance values of cirrus TOA is greater than 0.01 then the pixel will be labelled as the cloud pixel (Zhu et al., 2015). By joining the older F-mask that have the ability of detecting both thin and thick clouds and new F-mask that identifies the cirrus clouds and high altitude cirrus clouds it increased the accuracy of the methodology. This is by the overall accuracy of more than 99% with the challenge of over estimates cloud detection in cloud free pixels who's the pixel value are bright, white and cold (Zhu et al., 2015).

F-mask can be applied to remove clouds in sentine-2 images by excluding thermal band, Sentinel-2 has all bands which available in Landsat missions from Landsat 4-landsat 8 (Zhu et al., 2015). Therefore, the cloud detection

of Sentinel-2 can use the expanded F-mask version for Landsat 8, where by it will exclude brightness temperature as one of the input because Sentinel-2 it lacks thermal band (Zhu et al., 2015).

2.6. Gap-fillinging by using interpolation algorithms

As the main objective, to gap-filling missing information in Sentinel-2 after cloud masking. Here is the elaboration of multi-temporal interpolation techniques that can be used to Gap-filling the missing information in the Sentinel-2 satellite images.

The word Interpolation originated from Latin verb "interpolare" which is the combination of "inter" means "between" and "polare" means "to polish" in combination it means to smooth in between pieces of information given (Abdulla, 2014). Other meaning Interpolation: referred as the creation or construction of new data points which are generated from the discrete set of existing known data points (Levy, 2010).

There are different types of interpolation algorithms, but for temporal aspect three interpolation algorithms; Nearest Neighbor Interpolation, Linear Interpolation and Gaussian Interpolation will be discussed. Where Figure 2-5, shows how each interpolation algorithm function in gap filling.

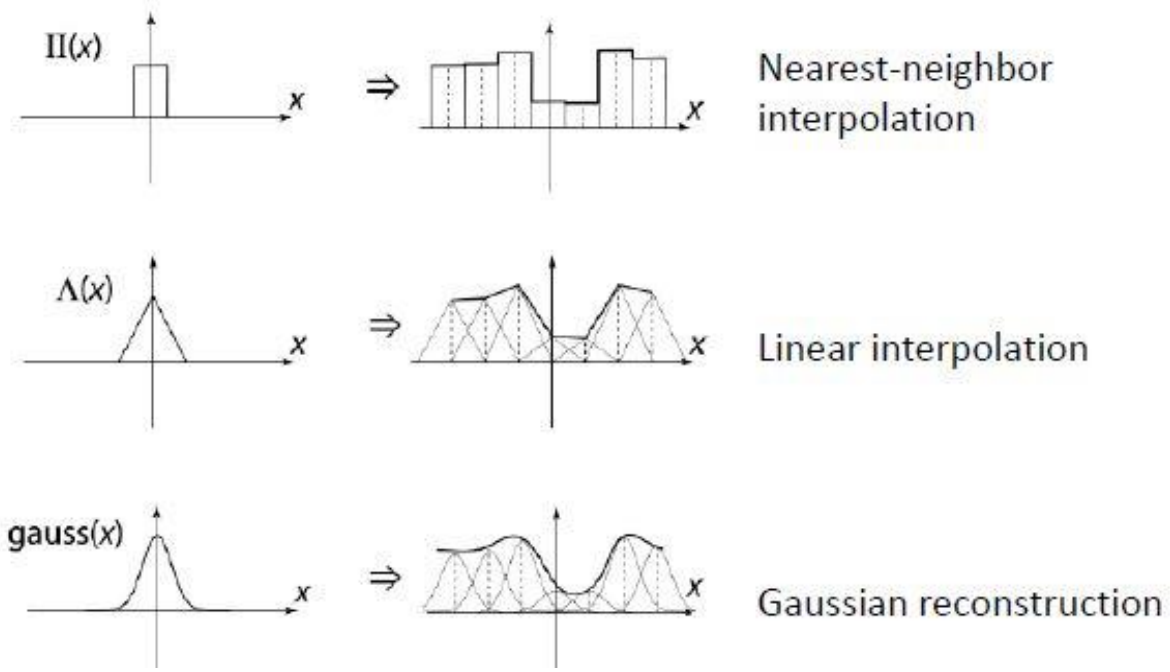


Figure 2-5: Interpolation algorithms function.

2.6.1. Nearest Neighbour Interpolation Algorithm

Involves the construction of new data point from the existing known data points that are more closest or more nearest, example in image the pixel with no values will be given the value of the closest neighbour pixel either in temporal or spatial aspect (Movebank, 2016). Nearest Neighbour does not involve modification of existing known data points to create new values, but the existing data values will be taken as the way it is to the interpolated point (Movebank, 2016).

Equation 2-5: Nearest neighbour new value prediction

$$(x_S, y_S): D(x_D, y_D) = S(\text{round}(x_S), \text{round}(y_S))$$

From the Figure 2-6, E is a data with no value. It needs to take values from either points A, B, C or D. Therefore, from nearest neighbour interpolation concept, A point seems to be closest neighbour to E, so it takes the values from A as the way it is (Movebank, 2016).

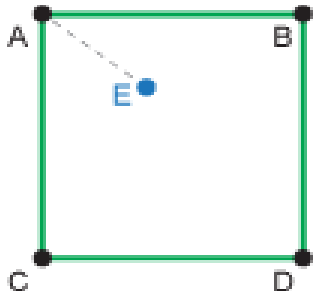


Figure 2-6 : Nearest Neighbour interpolation concept (Movebank, 2016)

For the context of temporal nearest neighbour interpolation, as explained in Figure 2-7, it used image pixel values captured in a month that form the basis of temporal aspect. Whereby, in date 8th and date 20th there was no value captured in a pixel. Therefore, the nearest neighbour interpolation context uses the closest date values to fill the gap of no value pixel. As observed from the graph image of date 8th has been filled with date 13th because is closest than date 1st, likewise image date 20th has been filled with image date 25th because is closer than image 13th.

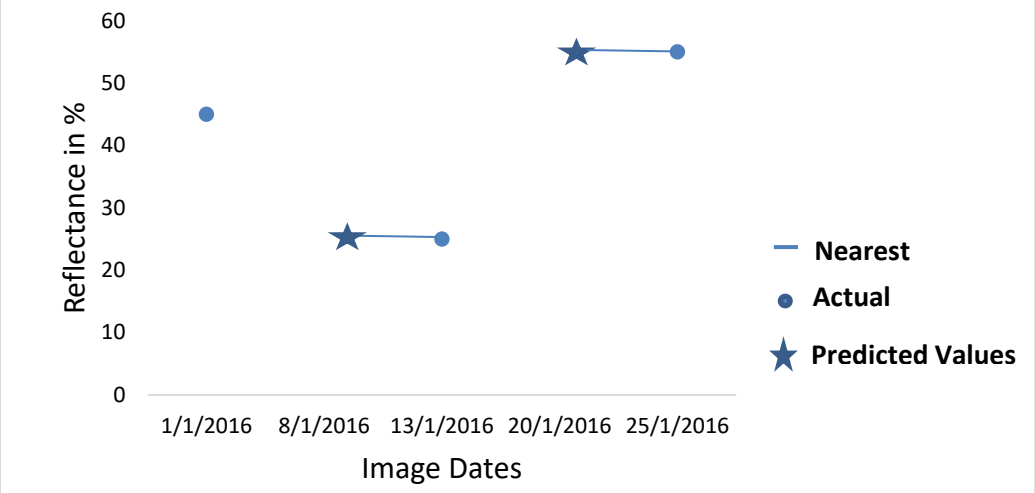


Figure 2-7: Temporal nearest neighbour interpolation function

For spatial nearest neighbour interpolation, can be used mostly for qualitative data for situations that defined by categories and it is not possible to calculate average as needed by other interpolation techniques, such as land use (Li & Heap, 2014; Movebank, 2016). Other situation for use of spatial nearest neighbour interpolation is when the data values are far, therefore, the best way is to take closest existing measurement and ignore other data values that are far away (Li & Heap, 2014; Movebank, 2016).

Advantage of using nearest neighbour interpolation considered as simple and very easy to clarify, also it does not requiring calculation of new values, it is only use the values from original dataset (Li & Heap, 2014;

Movebank, 2016). The disadvantage of the technique, only it considers one value in the range of data set (Li & Heap, 2014; Movebank, 2016).

2.6.2. Linear Interpolation Algorithm

This is one of interpolation technique where by, it uses straight line to connect two adjacent known data values to estimate the new value for unknown data (Blue leaf, 2016; Kendall & Han, 2003; Paulbourke, 1999) The Equation 2-6, shows the equation used in linear interpolation algorithm. Whereby, if the values of known two data points are (x1, y1) and (x2, y2), then the y value for some point x is:

Equation 2-6: Linear interpolation equation

$$y = y_1 + (x - x_1)\frac{y_2 - y_1}{x_2 - x_1}$$

Figure 2-8, elaborates how does the linear interpolation algorithm can predict new data point by straight line fit from the existing two data points in between.

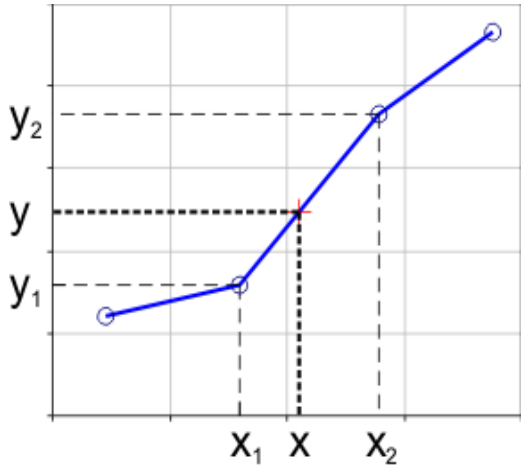


Figure 2-8: Shows straight line fit between two data points

Figure 2-9, elaborates the context of temporal linear interpolation, as explained in an image, it used image pixel values captured within a month that form the basis of temporal aspect. Whereby, in date 8th and date 20th there was no value captured in a pixel. Therefore, the linear interpolation context it takes the closest front and back pixels and use the straight line to connect and estimate new value to fill the gap of no value pixel. As observed from the graph image of date 8th has been filled by taking closer pixel value from back image date 1st and front date 13th and estimate values at image date 8th.

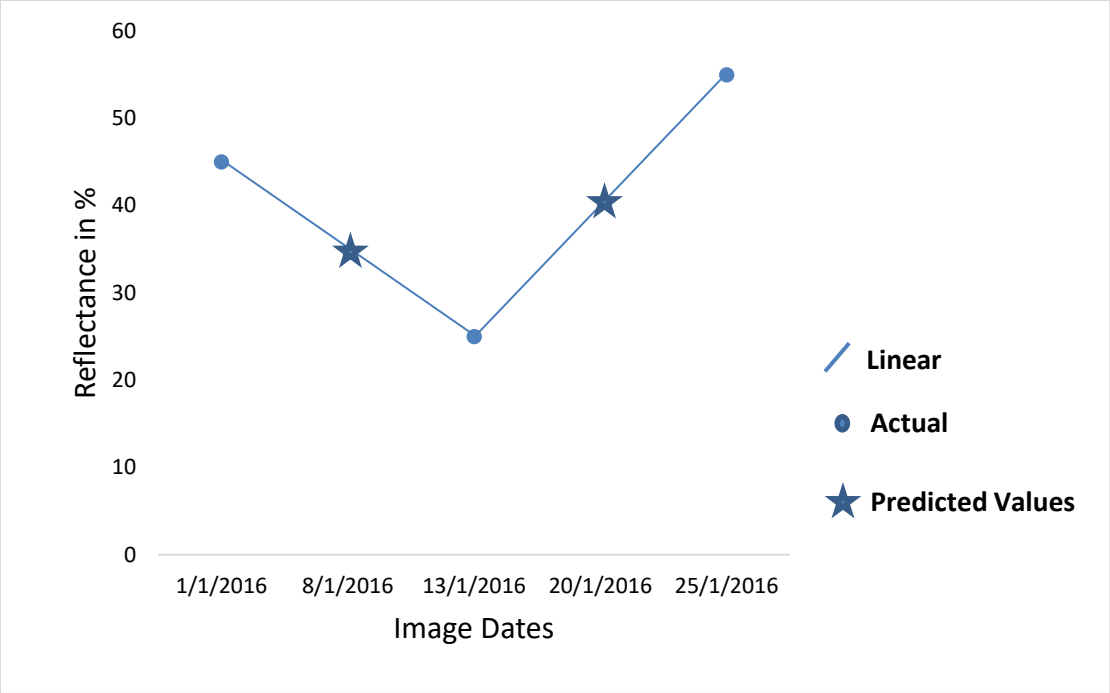


Figure 2-9: Linear interpolation function

2.6.3. Gaussian Interpolation Algorithm

It is named Gaussian, bell shaped or normal distribution function, where by it associated with two parameters mean and standard deviation N *(μ; σ)* and the Gaussian graph is characterized with symmetric (bell shape) curve (Gan, 2013; Hamilton, 2010; Revolvy, 2016; Ribeiro, 2004). One of Gaussian properties elaborates that Gaussian function consists of exponential function that contain the concave quadratic function. Therefore, the Gaussian function it algorithm is a concave quadratic function that have the ability to capture narrow peaks, noise, rise and fall within the data set. (Revolvy, 2016). Equation 2-7 shows how the Gaussian normal distribution function works whereby, it uses mean to control position of the peak and the standard deviation to control the width of the bell curve

The probability function f (x) of N *(μ;σ)*

$$f(\mathrm{x}|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad \text{Equation 2-7}$$

Where,

μ=Mean
σ=Standard deviation
σ^2=Variance = V(X)
Expected Value: E(X) = μ for a normal random variable X.
Symmetry= f (μ -X) = f (μ + X)

The parameter, mean *(μ)* is the one controls the location or position of the distribution or the centre of the peak while the standard deviation *(σ)* it controls the width of the bell curve (Gan, 2013; Hamilton, 2010; Ribeiro,

2004). As shown, Figure 2-10, the Gaussian function that forms normal distribution curve that is controlled by mean and standard deviation.

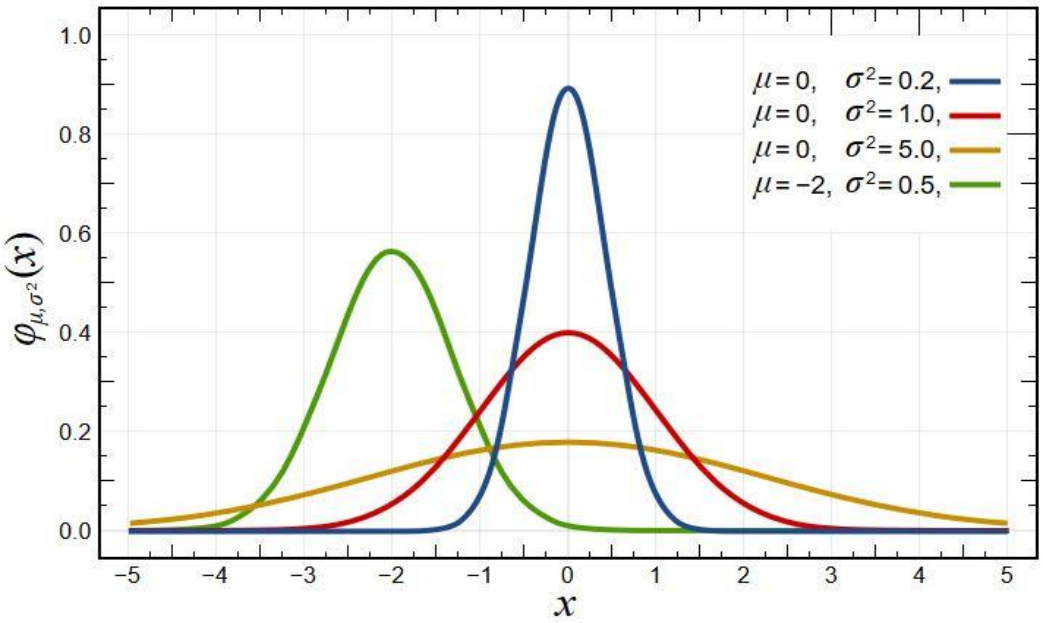


Figure 2-10: Gaussian function that forms normal distribution curve. (Revolvy, 2016

Therefore the Gaussian interpolation it is derived from the central difference interpolation formula, which involves linear combination of Gaussian bell-shaped curves and fits the interpolated data points from measured values $z(x, y)$ to follow the normal distribution curve (MathWorks, 2016; Revolvy, 2016; Spring, 2016)

- **Procedures for Curve fitting in Gaussian interpolation by (Spring, 2016)**

Equation 2-8 shows the Gaussian interpolation developed based on Gaussian normal distribution function (fundamental form g_i) for the measured values (x_i, y_i)

$$g_i(x,y) = \frac{1}{1+\frac{\sqrt{(x-x_i)^2+(y-y_i)^2}}{m^2}}, i = 1,2,\dots\dots\dots,n \qquad \text{Equation 2-8}$$

Where

m = average distance between all measuring points.

g_i represent surfaces of rotation which is linearly linked to surface function by measuring point (x_i, y_i).

$$z(x, y) = b_1 g_1(x, y) + b_2 g_2(x, y) + \dots + b_n g_n(x, y)$$

The coefficients b_j for the linear combination are unknown, they represent parameter of surface function. Therefore, in order to lie the measured values on defined surface, the following equation must come for each measuring point.

$$z_i = z(x_i, y_i) = b_1 g_1(x_i, y_i) + b_2 g_2(x_i, y_i) + \dots + b_n g_n(x_i, y_i)$$

Then follows a symmetric $(n \times n)$ equation system for the n unknowns b_i , $i=1,...n$.

Resulting values of bi complete the surface function z (x, y)

Therefore, the Replacement of x and y values returns the value at every node or rather element center, so to prevent fluctuations of the Gaussian surface function $z(x,y)$,the additional for surface function is required

$$Z = Z\,(x, y) = B_1\, g_1(x, y) + B_2\, g_2(x, y) + \ldots + B_n\, g_n\,(x, y)$$

The coefficients B_i , $i = 1, \ldots, n$ are determined using

$$Z_i = Z\,(x_i, y_i) = B_1\, g_1(x_i, y_i) + B_2\, g_2(x_i, y_i) + \ldots + B_n\, g_n\,(x_i, y_i) = 1 \text{ for } i=1,..,n.$$

Finally, the interpolation value I at a given point (x, y) defined as

$$I\,(x, y) = z\,(x, y) \,/\, Z\,(x, y).$$

2.7. Evaluation of different interpolation Algorithms

For testing the accuracy and comparing the performance of interpolation algorithms, it needs the evaluation techniques which measures the strength of statistical relationship between measured (actual) value and estimated (predicted) values (Piazza et al., 2015; Robinson & Metternicht, 2006)

For the aim of testing the quality of the model and the accuracy in model predictability, performed by checking the error, which is different between the actual or real value and the predicted value. Whereby, the lower the error the better the model in predicting the values (Piazza et al., 2015; Robinson & Metternicht, 2006)

For the nature of this study, assessing the accuracy of interpolation algorithms in predicting new values, Leave one out cross validation has been suggested by (Arlot & Celisse, 2010). This is because within LOOCV there are several statistical test to analyse the strength of interpolation algorithm in predicting new values, which can take place such as MSE, RMSE, MAE and MAPE.

2.7.1. Leave One Out Cross Validation (LOOCV)

According to the Encyclopaedia of machine learning Sammut & Webb (2010) defined cross validation as the process of creating distribution pairs of training set and testing test from a single data set (Arlot & Celisse, 2010; Christensen, 2015; Landy, 2015). Cross validation suggested being powerful technique for evaluating how well or the accuracy of the prediction model that may perform in an independent dataset, it is good because is not biasing the predictions (Arlot & Celisse, 2010; Christensen, 2015; Landy, 2015).

Leave one out cross validation is a form of cross validation whereby, in a dataset number of training is equal to the number of testing, this is by involving all member of dataset to be both a training sample and testing sample (Arlot & Celisse, 2010; Christensen, 2015; Landy, 2015). When implementing the algorithm it uses all member

of dataset as training except one member, which is used as testing set. Leave-one-out cross-validation is fundamentally an estimate of the generalization performance of a model trained on $n-1$ samples of dataset. The process is closely similar to statistical method jack-knife estimation (Arlot & Celisse, 2010; Christensen, 2015; Landy, 2015; Sammut & Webb, 2010). In LOOCV, the model will run until every member in the data set will be applied as both the training and testing member in a data set. Table, 2-1 shows how the Leave One-Out Cross Validation functions. Where for every fold the test are the values left out for validation (prediction statistics).

Table 2-1: The Leave One-Out Cross Validation function

FOLD 1	FOLD 2	FOLD 3	FOLD 4	FOLD 5		Prediction Statistics
Test	Training	Training	Training	Training	→	Test
Training	Test	Training	Training	Training	→	Test
Training	Training	Test	Training	Training	→	Test
Training	Training	Training	Test	Training	→	Test
Training	Training	Training	Training	Test	→	Test

Therefore, when implementing the test for each member in a data set the values will be taken for prediction statistics for testing the accuracy of the predictor. The following prediction statistics that may be conducted are:

2.7.2. Square mean Error and Root mean Square Error (RSME)

It starts with calculation of Mean Square Error (MSE), this tires to measure how close the fitted line is to the actual values (data points) (Barnston, 1992; Piazza et al., 2015; Robinson & Metternicht, 2006; Vernier, 2016). Involves measuring of the vertical distance from the actual point to the estimated in order to find the different between them on the curve fit (the error), and then square the values in order to remove the negative signs and give extra weight to huge differences (Barnston, 1992; Statistics how to, 2016; Vernier, 2016). Therefore, the smaller the SME the smaller the error between the actual values and estimated values and closer you find the line of best fit (Barnston, 1992; Statistics how to, 2016; Vernier, 2016)

After calculating the MSE, it followed by calculating RMSE, this is only the square root of the MSE. Thus, the RSME is the average distance of the actual data point from the fitted line (estimated value) (Barnston, 1992; Statistics how to, 2016; Vernier, 2016). RMSE is a better measure of predictor accuracy or the goodness of fit than even correlation coefficient (Barnston, 1992; Statistics how to, 2016; Vernier, 2016). Equation 2-9, is the equation used for calculation of root mean square error.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\{z(x_1) - \hat{Z}(x_1)\}^2} \qquad \text{Equation 2-9}$$

Where;

$z(x_1)$ = Actual Value

$\hat{Z}(x_1)$ = Predicted Value

N = Number of observation

The basis of using RMSE, is to judge the quality of predictions.

2.7.3. Mean Absolute Error (MAE)

This is a measure to quantify how predictions or forecasts are close to the actual values. The absolute error means the absolute value which obtained by identifying the different between actual value and forecasted value, it express how large an error can be expected from the predicted on average (Chai & Draxler, 2014; Hyndman & Koehler, 2005; kaggle, 2016; Tadd Wood, 2016.). Equation 2-10, is the equation used for calculation of mean absolute error.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}[y_i - \hat{y}_i] = \frac{1}{n}\sum_{i=1}^{n}[e_i] \qquad \text{Equation 2-10}$$

$$AE = [e_i] = [y_i - \hat{y}_i]$$

Where;

y_i = Actual Value

$\hat{y}_i$ = Predicted Value

n = Number of observation

2.7.4. Mean Absolute Percentage Error (MAPE)

MAPE is one of the method in statistic which used to measure the prediction accuracy of the predictor, it measures the size of an error and tell the accuracy in percentage terms (Harvey, 1989; Hyndman & Koehler, 2005). Equation 2-10, is the equation used for calculation of mean absolute percentage error.

$$M = \frac{100}{n}\sum_{i=1}^{n}\left[\frac{A_t - F_t}{A_t}\right] \qquad \text{Equation 2-11}$$

Where;

A_t = Actual Value

F_t = Predicted Value

n = Number of observation

MAPE is claimed to achieve scale independence whereby, it takes the different between actual A_t and predicted F_t and dividing it with actual A_t. But this situation brings some disadvantages like limitation of using data when there is zero because will be divided by zero (Harvey, 1989; Hyndman & Koehler, 2005).

2.7.5. Relevance of using RMSE and (MAE&MAPE)

The two common used statistics to measure the prediction accuracy of the predictor, is RMSE and MAE (Chai & Draxler., 2014; Roy et al., 2016; Willmott & Matsuura, 2005). However, every measure it has its own

advantages and disadvantages. As generally, the RMSE express highest error deviations while MAE and MAPE it comprise easiest clarification and it is helpful in evaluating the accuracies between predictors.

The RMSE has been given the priority and used for long time to evaluate the accuracy of predictors performance because it uses the standard error, that helps to provide the understanding of the error distribution (Chai & Draxler, 2014). The best statistic measure to be used has to present and give both performance and error distribution (Chai & Draxler, 2014). Therefore, the relevance of using the RSME to evaluate the predictor performance was due to its ability to assess the performance of the predictor and the ability to show how does the error distributed. In addition, the RMSE is unbiased, has the ability to detect noise and outliers in the data, and in addition it follows the normal distribution data format.

According to Roy et al, (2016); Willmott & Matsuura (2005) they realized that, even though RMSE has been used most and reported widely, there are some error within it and misinterpretation of RMSE. This is due to use of the squared expression of the predicted errors during explanation of RMSE whereby the error variance can be affected by the set of data. The error predicted with higher value it will be given more weight than the predicted error with low values, therefore the higher values errors, have greater control on the total square error than the smaller error values. (Willmott & Matsuura, 2005). In addition, the information presented by RMSE has been doubted to be very complicated compare to MAE and MAPE are the same statistic measure which quantify the absolute error. Their small difference comes only where MAPE quantify the deviation from the actual values in terms of percentage.

MAE and MAPE are declared as straightforward and simpler determinant of errors prediction (Willmott & Matsuura, 2005). This is due to determination of the central tendency that comprise average error prediction together with error spreading causes the measures to be direct and straightforward (Roy et al., 2016; Willmott & Matsuura, 2005). In addition, they propose MAE and MAPE, because there is no squaring of error, it involves absolute values, and maintain the sign of error.

The study opted to involve both the RMSE, MAE and MAPE for evaluating the accuracy of predictor performances. Because there was strong argumentation between scholars that denied the dependency of one measure. Chai & Draxler (2014) confidently concluded that, RMSE is most appropriate statistical measure to use than MAE and MAPE, because RMSE is unbiased, has the ability to detect noise and outliers in the data, and in addition, it follows the normal distribution data format. According to Roy et al., (2016) and Willmott & Matsuura (2005) they concluded confidently that, the MAE and MAPE are the most natural statistic measure of average error magnitude than RMSE. Therefore, RMSE should not only depended to measure the average error. In MAE and MAPE there is no squaring of an error, therefore the magnitude of the error is maintained and not influence the average error measures like how RMSE is performing.

3. MATERIALS AND METHODS

3.1. Materials

3.1.1. Study area

The study area is located in the Upper East region of Ghana. It lies between latitudes 10°30' and 11°15' North and longitudes 0° and 1°30' West (Figure 3-1). The region is dominated by Guinea Savannah zones except the corner at North-eastern part where is occupied by Sahelian Savannah land mass (Daniel, 2015; Kusi, 2013; Seidu, 2011). The harsh Guinea condition dominates almost throughout the year. The climatic condition of the region is poor and inconsistency, dry land patterns and poor soil, poor vegetation cover, bad agriculture practices, unreliable rainfall pattern and prolong dry season (Daniel, 2015; Kusi, 2013; Seidu, 2011)

In the study area, are two large irrigation schemes. The Tono and VEA where all data was collected for the study. The Tono irrigation scheme is the irrigation project with 3840ha potential area and 2490ha developed area (Tono irrigation scheme, 2017) While its source of water is from Tono River, the project started in 1965 and completed in 1980. The scheme lies between latitude 10° 45' N and longitude 1° W. The project was inaugurated by Government of Ghana to ensure that small-scale farmer's irrigation schemes are well managed and organized. Rice is the dominance food crop cultivated throughout the year in the project lowland area. Other food crops cultivated in a project zone are pepper, Garden-eggs, Groundnuts, onions, tomatoes, maize and millet.

VEA irrigation scheme has 1197ha potential area with only 852ha developed (VEA irrigation scheme, 2017). Its source of water is from Yarigatanga River. The project started in 1975 and completed in 1985, the scheme lies between latitude 10° 45' N and longitude 1° W in Bongo near Bolgatanga town (Vea irrigation scheme, 2017.). The food crops cultivated in a project zone are rice, pepper, Garden-eggs, Groundnuts, onions, tomatoes, maize and millet.

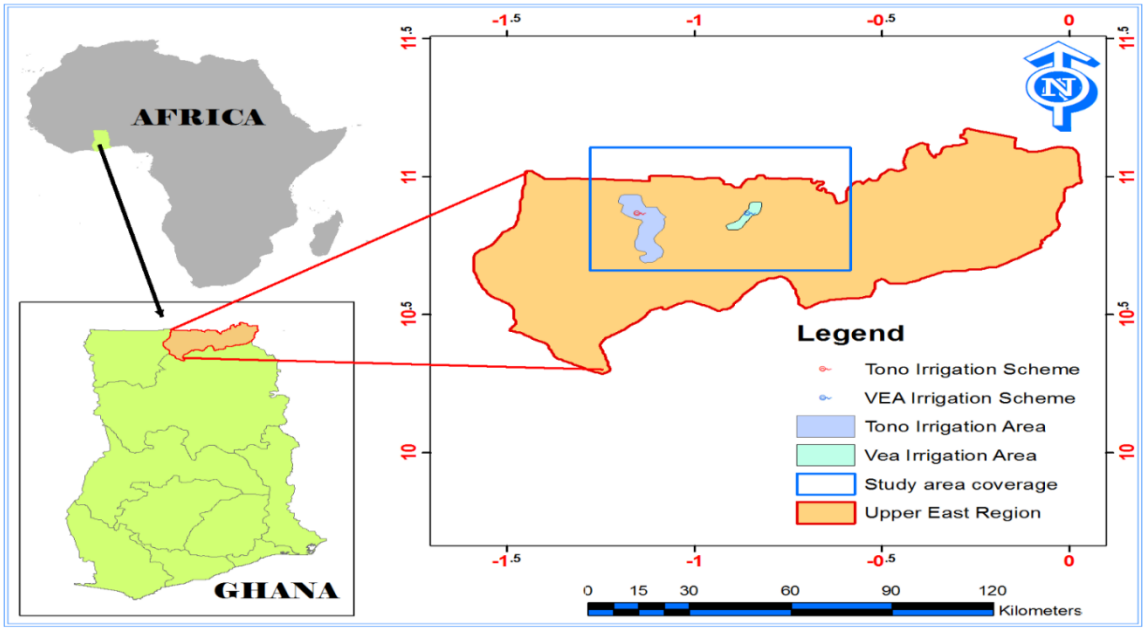


Figure 3-1: Study area location map

3.1.2. Software

During this study, different and various software packages were used for processing and analysis of datasets. As elaborated in Table 3-1 are the software used and the purpose of use.

Table 3-1: List of software and its purposes/use

Software	Purposes/Use
ArcGIS 10.3.1 and ERDAS 2015	GIS analysis, map making and Image preparations
ESA SNAP 4.0	Sentinel-2 image preparation and cloud masking
ENVI 5.3.1 SP 1	Image preparations level 1B to 1C
F-Mask algorithm in ENVI	Cloud masking to Level 2A
EditIX xml editor	Granule metadata preparation
Microsoft Visual Studio	Python Script writing
F-mask Matlab Script	Cloud detection
Python 2.7 and Python Script	Gap-filling
SPSS	Statistical analysis
MS Office 2013 (Excel)	Statistical analysis
MS Office 2013 (Word)	Reports and Thesis writing
Mapc2mapc V564	Preparation of mobile map
Orux navigation app	Tracking farms

3.2. Methods

1. Download sentinel-2 Images from January to November 2016
2. Subseting Sentinel-2 images by creating new metadata for every image.
3. Performing geometric, radiometric and atmospheric corrections
4. Detecting and masking clouds in all Sentinel-2 images downloaded
5. Categorize land cover into static classes (water & built-up) and dynamic classes (crop farms)
6. Selecting of cloud free farms, tracking and collecting their information
7. Arranging all images, both cloud free and contaminated (NaN)
8. Extracting selected pixels in a time-series data format using 5x5 pixel window
9. Gap-filling missing information in 'clouded' image dates for farms, built-up and water bodies using interpolation algorithms in excel and test for accuracy of the interpolation algorithms
10. Gap-filling for crop farms using the crop growth rate as a guide and test for accuracy of the interpolation algorithms in gap-filling according to growth rate.
11. Significance difference test of interpolators in gap-filling
12. Write Python script and performing actual Gap-filling

Figure 3-2 is the method flow chart, it explains the methods, data and steps that the study used to achieve its main aim.

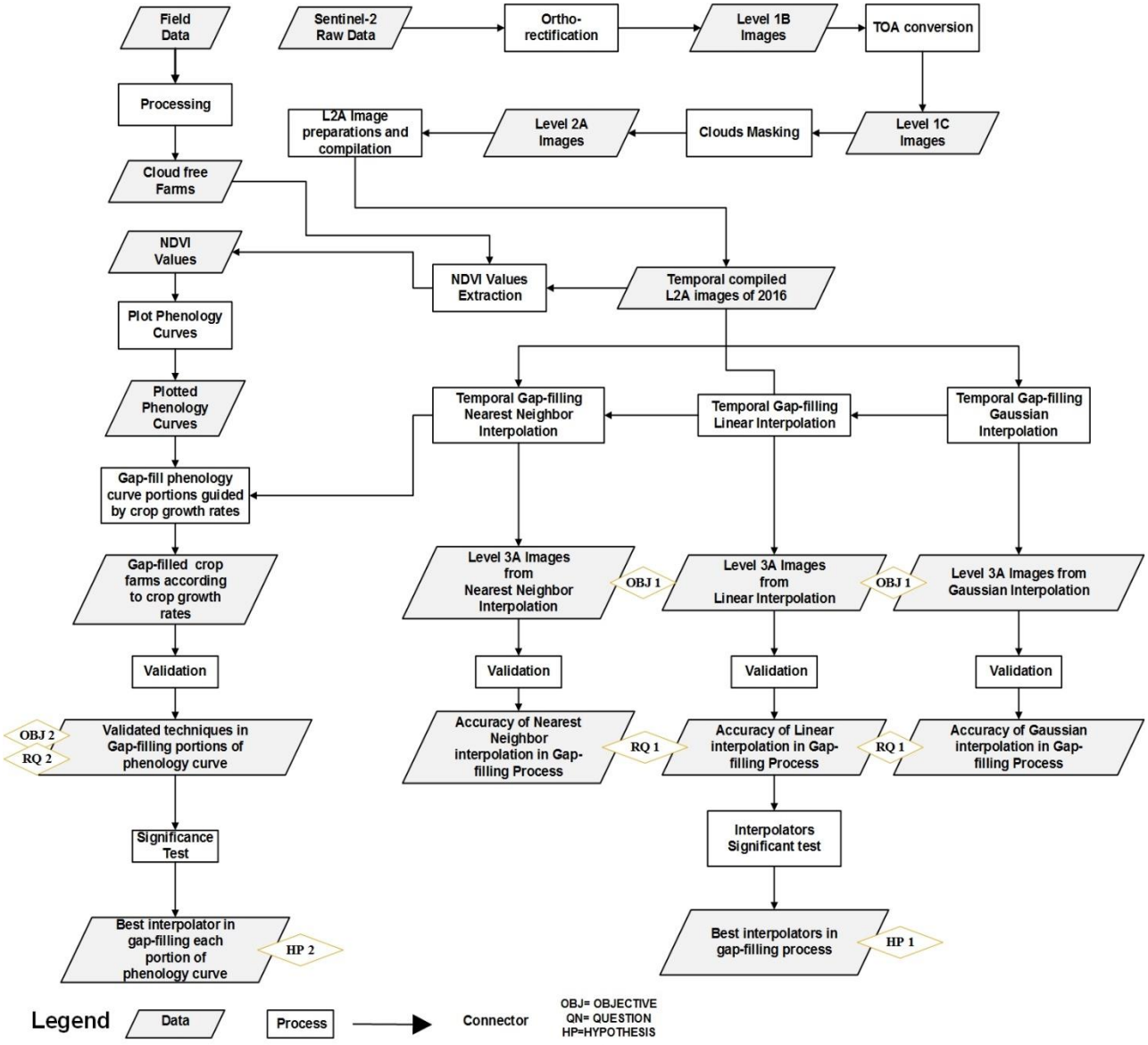


Figure 3-2 Flowchart showing the methods used in the study

3.2.1. Download of Sentinel-2 images.

Data used for this study was Sentinel-2-time series images. All images for the study area, dating from January, 2016 to November, 2016 were downloaded from European Space Agency (ESA) portal (https://scihub.copernicus.eu/). Sentinel-2 images are free for download to all individuals and entities worldwide through ESA online portals such as Scientific Data Hub and Copernicus Services Data Hub. These are for non-commercial use. To access the Sentinel data hubs, the entities and individuals are supposed to be registered, user will be given the user name and password to access the data. The Copernicus programs is committed to deliver upscale services of Sentinel-2 images to its user for more than 20 years. This has been observed from the very start date where by, the system has stable reliability and good performances. In addition, the commitment of Copernicus program been observed through major improvement concerning calibration and accuracy. Particularly for active applications range from crop, forest, vegetation and land-management

monitoring, this is one of the exceptional promotion of Sentinel-2. The information about the image dates and the identification number is found in Appendix 1.

3.2.2. Subseting of Sentinel 2 Imagery

The 290x290 km main image files and its metadata was critically analysed and the granules identification numbers that formed the main image was identified. The Edit-IX-xml-editor 2016 software, was used to edit the 290x290 km main image metadata and an xml script metadata for 100x100 km granule was developed. The sample of xml metadata script developed is attached in Appendix II. The 100 x100km granule metadata file developed was used to subset the main image file using ENVI software. As shown in Figure 3-3, the granule was subseted from the main image.

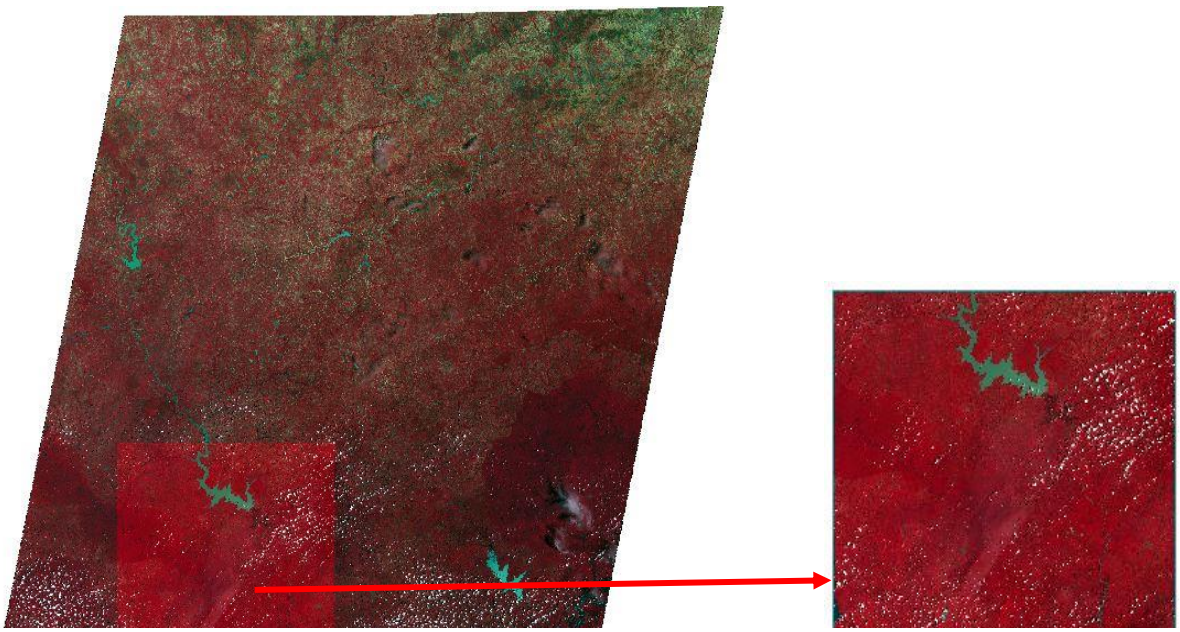


Figure 3-3 Granule subset from the main image by using metadata created

3.2.3. Ortho-rectification of Level 1A images

An SRTM 30m resolution DEM of study area was downloaded from https://earthexplorer.usgs.gov/, mosaicked and resampled to 10m spatial resolution. The resampled DEM was used to Ortho-rectify the Sentinel-2 images (Level 1A) using the ENVI software version 5.3.1. In the process, the Sentinel-2 image was resampled to 10m resolution and projected to UTM WGS 84 Zone 30. This resulted in a level 1B image.

3.2.4. Conversion of DN values to Radiance and TOA Reflectance.

The Ortho-rectified images (Level 1B) were converted to Radiance then to TOA reflectance by using ENVI software. Figure 3-4 shows Image A, B and C. Image A shows raw level 1B image, with no correction (radiometric and atmospheric) applied. In image B, the DN values has been converted to physical meaningful units (Radiance). In image C, the radiance has been converted to top of atmosphere (TOA) reflectance to form a Level 1C image.

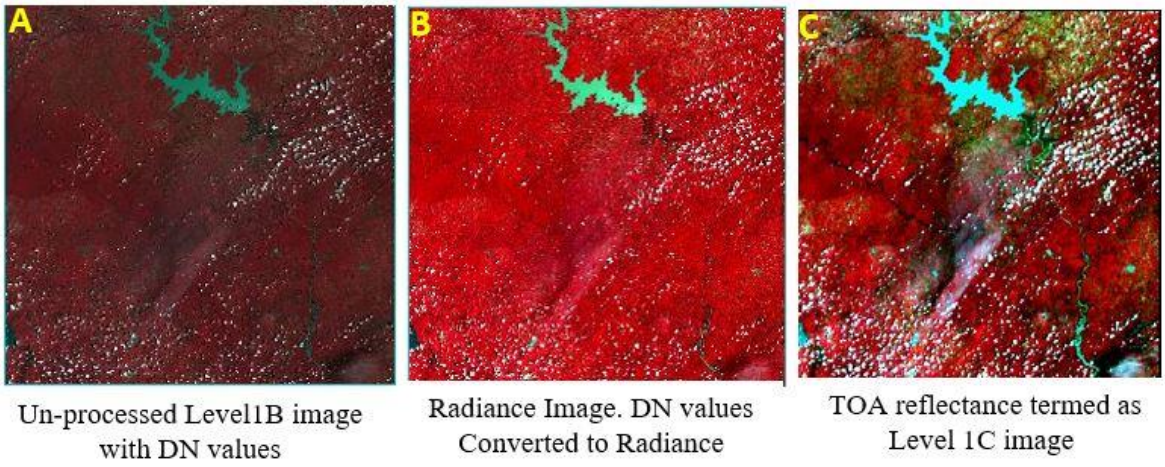


Figure 3-4 : Image processing that form Level 1C Images

3.2.5. Cloud detection and masking of Sentinel-2 images

F-mask expanded version designed for cloud detection of Landsat 8 images was used for the detection and cloud masking in Sentinel-2 images. In this process, the 10 m resolution TOA bands 2, 3, 4, 5, 6, 7, 8, 8a, 11 and 12 were stacked and input as the base image in F-Mask. Bands 1, 9, and 10 were resampled to 10m TOA, stacked and used as the input layers to detect clouds in the base image. This resulted in the detection of all clouds in the base image. Erdas Imagine was used to mask the clouds

Figure 3-5 shows the process of cloud detection and masking in the Sentinel-2 images. Figure 3-5 (a) shows the level 1C images with clouds cover. Figure 3-5 (b) is the cloud cover layer detected after run the algorithm in F-mask. Black regions represents cloud and the white shade are the surface feature. Figure 3-5 (c) shows the final Level 2A image with all clouds been masked.

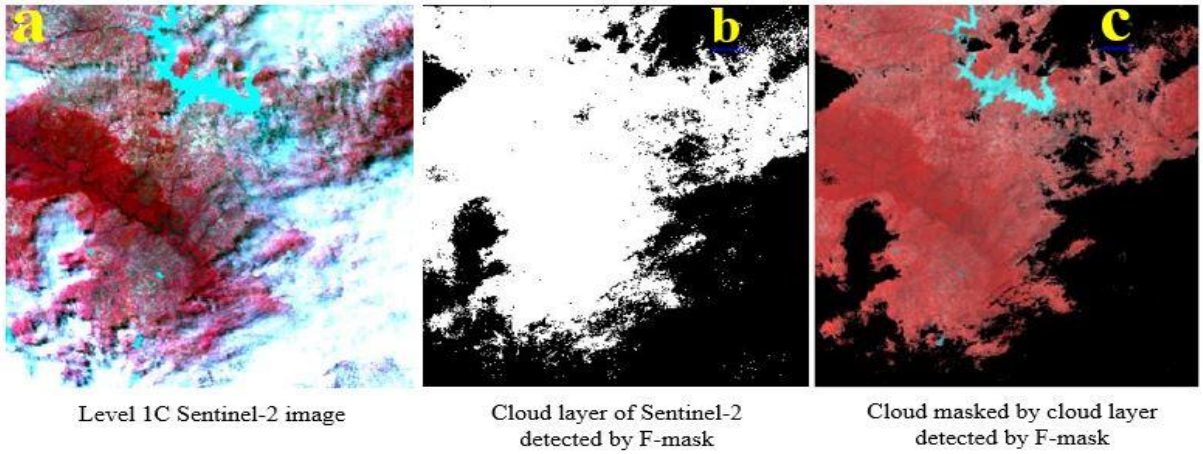


Figure 3-5: The results for cloud detection of Sentinel-2 images by using the F-mask algorithm.

3.2.6. Selection of signature points for training and accuracy assessment

Signature selection was based on temporally static (built-up, open water surface etc.) and temporally dynamic (i.e. crop farms) land cover classes. Very high-resolution satellite images from google earth were used in the identification and selection of signature points for the study. The google earth images were used as map base for tracking the land cover classes (farms, built-up, water etc.). Farms and other land cover classes with more

than 10 cloud free temporal images (image dates) were selected and tracked using orux map. Figure 3-6 shows farms (representing temporally dynamic land cover classes) that were identified from google earth image. The identified farms were tracked for field work data collection.

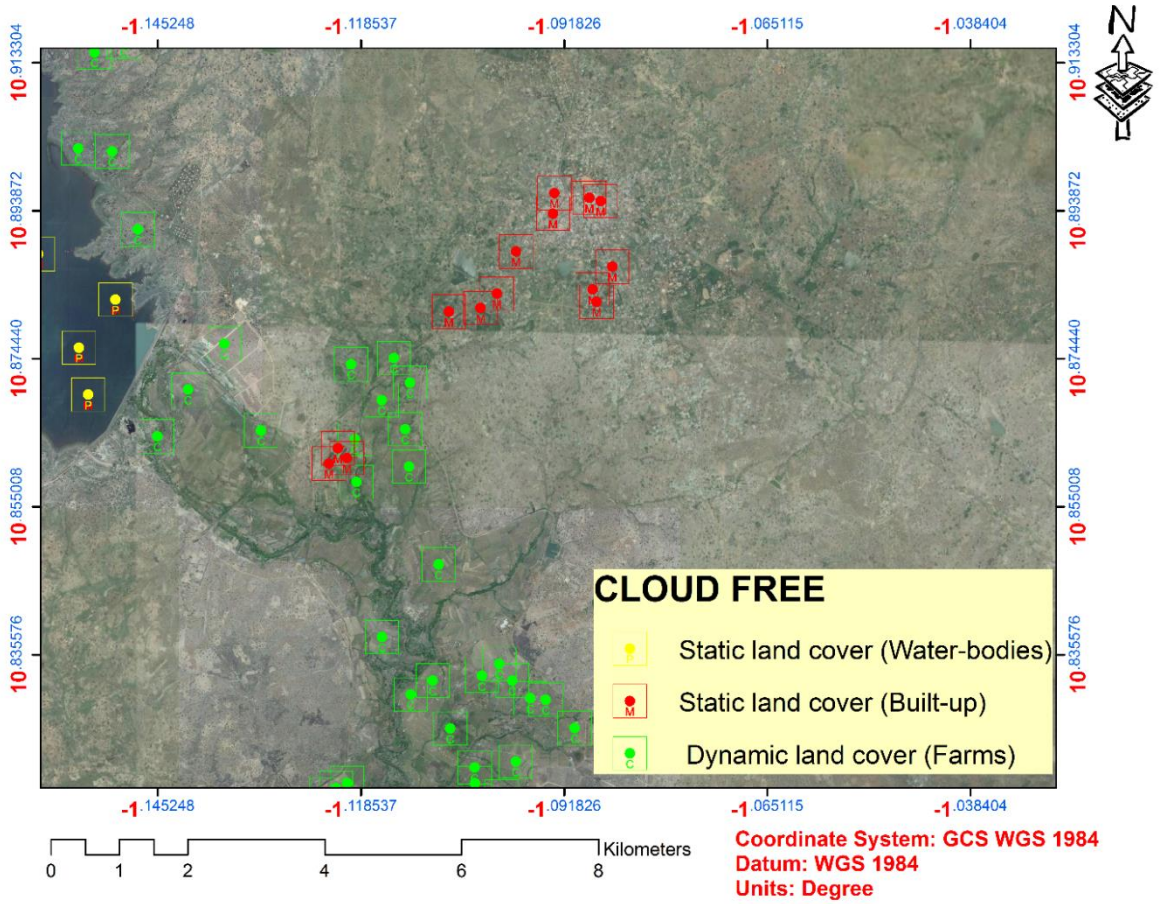


Figure 3-6: Farms, water body and built-up Identification from VHR (google earth)

After identifying signature points from VHR google earth, all cloud masked satellite images were composited into a single image file. Identified signatures were overlaid with the cloud free composite image to help identify the points (pixels) that fell under the cloud free portions of the image. Values for the cloud free pixels farms, water body and built-up were extracted for analysis. Appendix III showed image values extracted from some selected farms, water body and built-up.

3.2.7. Data collection and data processing

In other to construct a phenology curve for the temporally dynamic signatures (crop farms), farmers of 'cloud free' farms, were identified and interviewed based on the following information requirements:

- Type of crops in the farm
- Month of planting
- Month of harvesting

In all, 37 farmers were interviewed. Appendix IV shows the summary of the information collected from the farms. The minimum size of farms tracked was one hectare.

3.2.8. Preparation for application of Gap-filling techniques

A 5x5-pixel window representing 25 pixels (about a quarter of the average farm size of 1 Ha in the study area) was used to extract reflectance values from centre of selected farms, water bodies and built-up land cover classes. The average surface reflectance value for a selection was calculated and used as representative value for the land cover class.

Figure 3-7 and Table 3-2 shows sample rice farm selected from the study catchment for Gap-filling of missing information. From the figure and table, it can be seen that reflectance values were only present for 12 image dates (Cloud free images) out of the 34 Sentinel-2 images downloaded for the scene. The values were extracted from the near infrared band. From the table the maximum and minimum values extracted were 0.248 and 0.359 respectively, with a mean value of 0.283±0.028.

Figure 3-7: Farm values extracted, preparation for gap-filling techniques application

Table 3-2: Description of farm values extracted for application of gap-filling techniques

Descriptive Statistics	Values
Average	0.283
Standard Deviation	0.028
Median	0.28
Minimum	0.248
Maximum	0.359
1st Quartile	0.268
3rd Quartile	0.286
Total images	34
Cloud free Images of a farm	12
Clouded Images of a farm (NaN)	22

Table 3-3: farms selected for Gap-filling

Farm Number	Crop types	Planted	Harvested
1	Pepper	September 2016	From December 2016 to April 2017
2	Groundnuts	July 2016	November 2016
3	Maize Farm	July 2016	November 2016
4	Pepper	August 2016	From December 2016 to April 2017
5	Groundnuts	July 2016	November 2016
6	Rice	February 2016	May 2016
		July 2016	November 2016
7	Rice	February 2016	May 2016
		August 2016	December 2016
8	Groundnuts	July 2016	November 2016
9	Rice	February 2016	May 2016
		August 2016	December 2016
10	Rice	August 2016	December 2016

As shown in Figure 3-8, the crop calendar and satellite observation of study area based on the farmers interview. Helped to understand the duration of crop planting and harvesting. Also helped to understand the phenology curve by relating the image captured and the growth stage that the crop is experiencing.

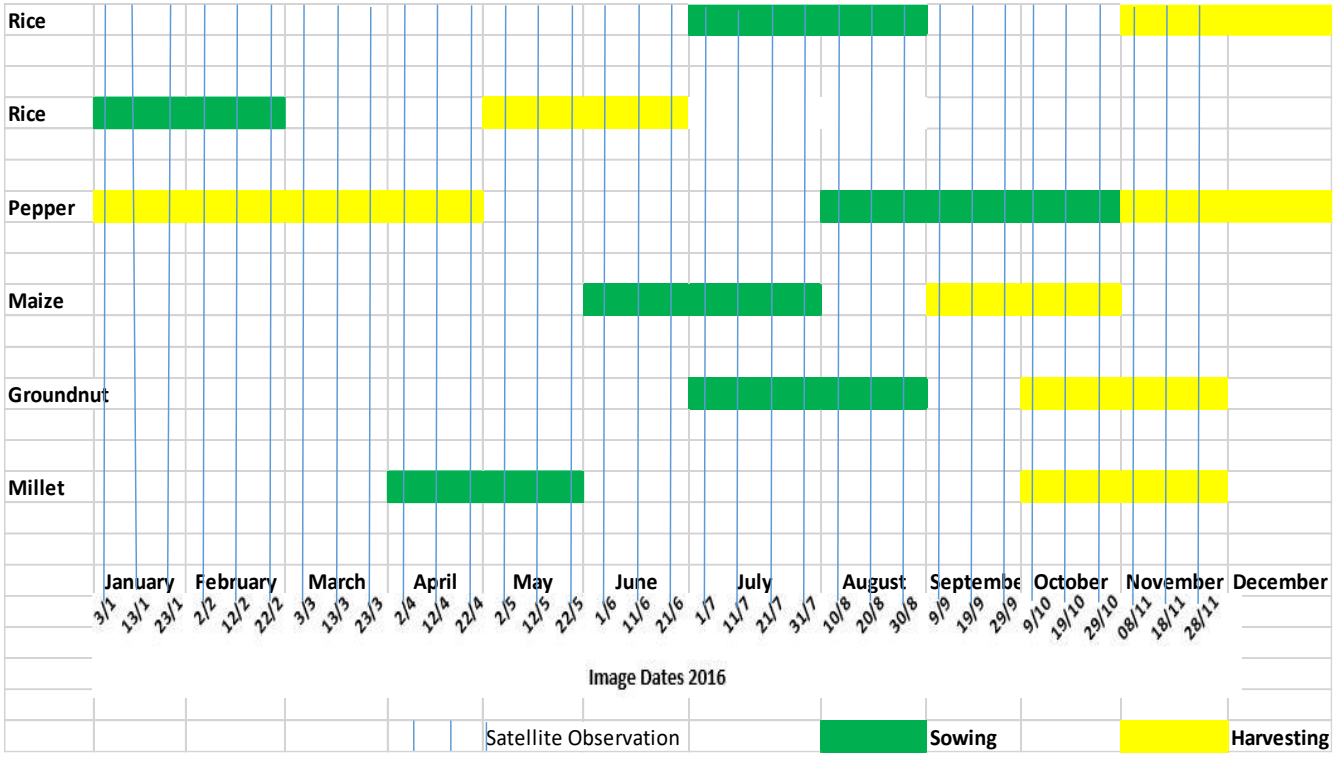


Figure 3-8: Crop calendar of study area with the satellite observation after every 10 days

3.2.9. Gap-filling of missing information

After cloud masking and farms values extraction, the three missing information Gap-filling technique (nearest neighbor, linear and Gaussian interpolator) were applied to fill the gap in Sentinel-2 that was caused by cloud existence. The equations for both nearest neighbor, linear and Gaussian interpolation was written in excel. Followed by arranging the images according to date for all 34 images from January to November. The image arrangement comprised of cloud free farm extracted values together with NaN farm image dates. The image arrangement (Image dates with cloud free and NaN) was inserted as input to the written equations in excel. Therefore, the cloud free extracted farm values used to gap-filling the missing information of NaN farm image dates.

3.2.10. Evaluation of interpolation algorithm used

Leave one out cross validation (LOOCV) was performed for every interpolation technique to get the predicted values. Tables 3-4 illustrates the LOOCV implemented for the Gaussian Interpolation technique applied for Gap-filling information of Farm 1 (pepper farm). The cells mark with asterisk (*) are the value left out during the LOOCV implementation and Gap-filled by the Gaussian Interpolation technique. The value predicted in asterisk cells were compared to the known farm values to test the accuracy of each Gap-filling technique applied. Mean square error (MSE), Root mean square error (RMSE), Mean absolute error (MAE) and Mean absolute percentage error (MAPE) were then performed to assess the accuracy of the Gap-filling techniques.

Table 3-4: LOOCV for Gaussian Interpolation technique

known values	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Fold 11	Fold 12
0.263	0.204*	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263	0.263
0.281	0.281	0.329*	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281
0.359	0.359	0.359	0.284*	0.359	0.359	0.359	0.359	0.359	0.359	0.359	0.359	0.359
0.285	0.285	0.285	0.285	0.304*	0.285	0.285	0.285	0.285	0.285	0.285	0.285	0.285
0.248	0.248	0.248	0.248	0.248	0.271*	0.248	0.248	0.248	0.248	0.248	0.248	0.248
0.266	0.266	0.266	0.266	0.266	0.266	0.265*	0.266	0.266	0.266	0.266	0.266	0.266
0.283	0.283	0.283	0.283	0.283	0.283	0.283	0.280*	0.283	0.283	0.283	0.283	0.283
0.289	0.289	0.289	0.289	0.289	0.289	0.289	0.289	0.276*	0.289	0.289	0.289	0.289
0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.279*	0.268	0.268	0.268
0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.268	0.272*	0.268	0.268
0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.281	0.278*	0.281
0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.294*

* Are the value left out during the LOOCV implementation and Gap-filling using Gaussian Interpolation technique

3.2.11. Evaluation of Gap-filling techniques in different portions of crop growth rates.

To determine which interpolation techniques will be the best in gap-filling crop farms, according to the different crop growth rates. Phenology curve of the crop was plotted. Figure 3-9 shows the phenology curve for the pepper Farm 1 and growth rate portions that forms phenology curve such as linear, stationary and exponential curves. This was used to assess the accuracy of the interpolators in Gap-filling information of dynamic land covers such as crops.

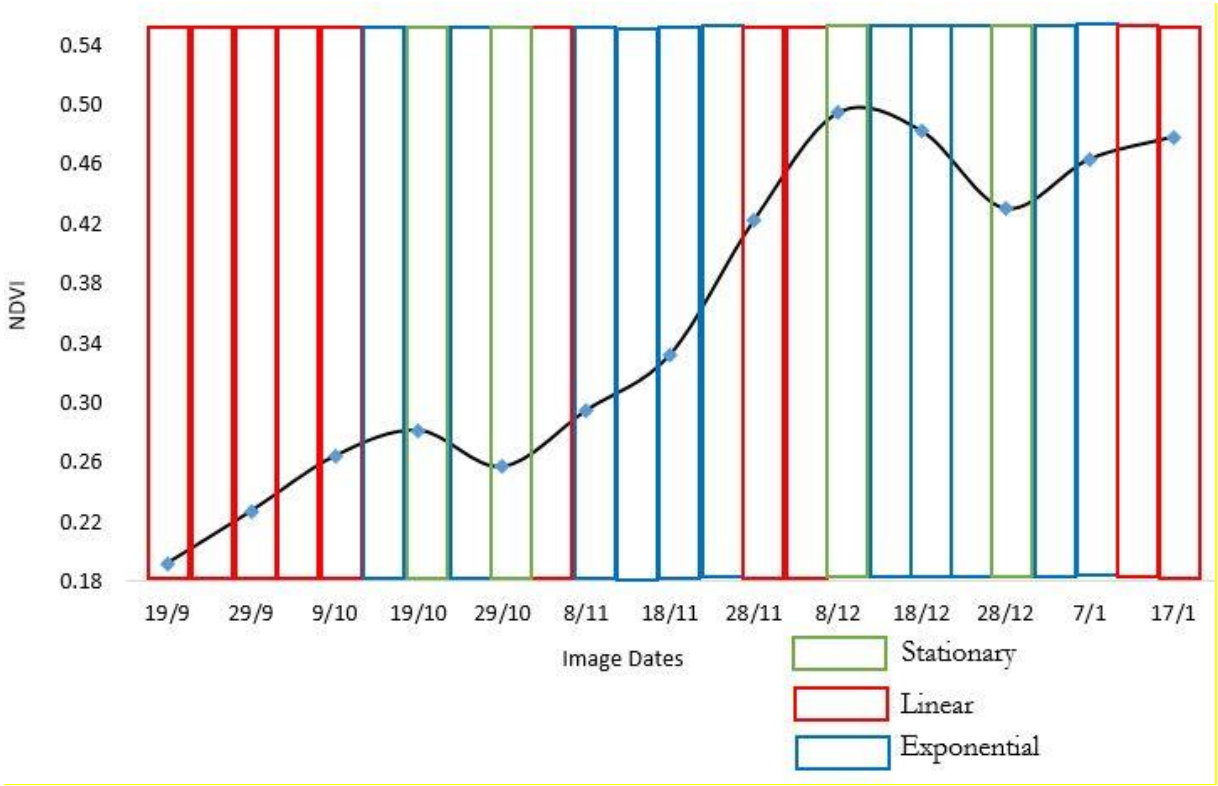


Figure 3-9: Phenology curve with different growth rate portion from pepper Farm 1.

The evaluation of Gap-filling techniques was conducted based on observation of the phenology curve portions. In each segment of the phenology curve, the evaluation was conducted independently for all Gap-filling techniques applied. For example Figure 3-9 portions with red color forms the linear portions of phenology curve, the images that forms this portion alone was taken and all the interpolation techniques was applied to Gap-filling the missing information in LOOCV and the accuracy assessment of the interpolators in Gap-filling the missing information was assessed. Likewise, for portions with green color, that forms stationary portions and segments with blue color that formed exponential portions of phenology curve were all Gap-filled in LOOCV by all interpolator, and the accuracy of each interpolator was assessed.

3.2.12. Evaluation of Gap-filling techniques in temporally static land cover classes.

This was performed to assess which interpolation technique perform well in Gap-filling the missing information in static land cover such as built-up and water bodies. Therefore, cloud free water body and built-up areas was selected and the value from blue band was extracted. Each interpolation technique was applied to test which technique performed better in gap-filling of static land cover. Figure 3-10, shows the reflectance values extracted from cloud free images at mid of Tono dam which representing the static land covers.

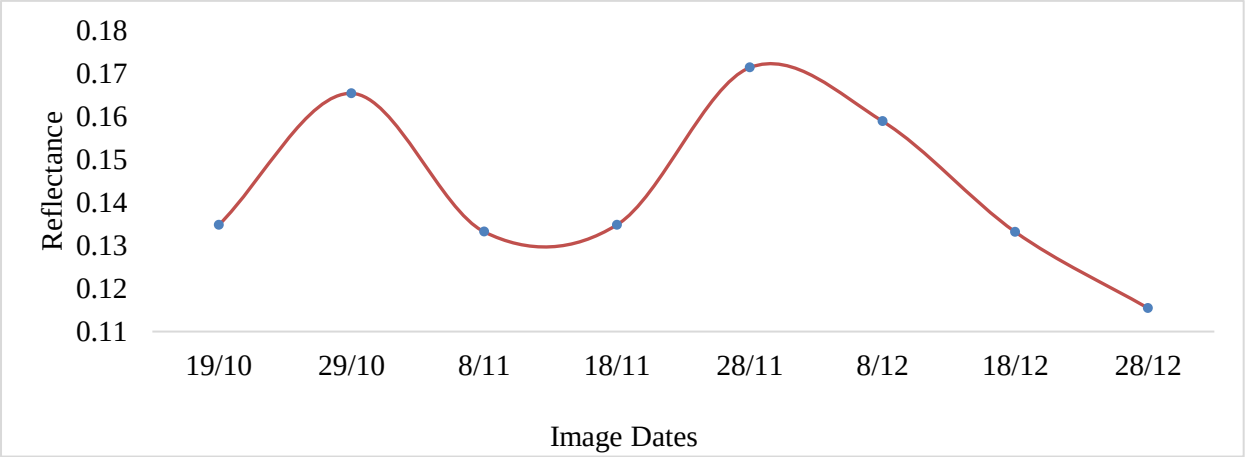


Figure 3-10 : Average known reflectance value of water body generated from the blue band.

Figure 3-11, shows the reflectance values extracted from cloud free images representing the static land covers

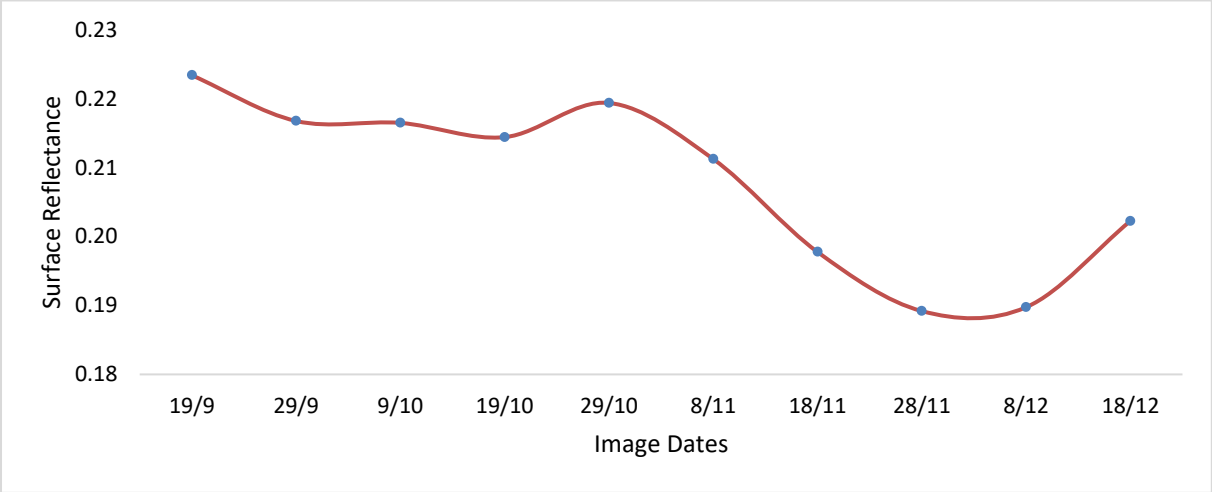


Figure 3-11: Average known reflectance value of built-up generated from the blue band

3.2.13. Significance test of interpolation techniques

After the accuracy test of interpolators, the significant test was conducted to conclude which interpolation technique is best in gap-filling. It started by testing the data normality using the Shapiro-Wilk normality test in SPSS version 20. The purpose for normality test was to understand type of statistical test to run. The data was normally distributed for all interpolators. Due to normality of data, the parametric test one-way ANOVA was selected to test the significance of interpolators if they significantly differ or which interpolator significantly performed better than other.

The repetitive measure one way ANOVA was performed and discovered there was statistical significant different between interpolators. To understand which interpolators were significantly differ, led to the Post-hoc test, and controlling for Bonferroni effect (thus alpha divided by number of comparisons) where significance was made at 0.05

3.2.14. Actual Gap-filling by using the python script

In order to perform actual Gap-filling, the python scripts was developed by using the interpolation algorithms tested. The scripts developed by using Microsoft Visual Studio software. The Script for Nearest Neighbour and Linear interpolation algorithm is available in Appendix IV and V

4. RESULTS

4.1. Objective 1

To gap-filling the missing information in sentinel-2 images by using temporal interpolations algorithm and accessing their accuracy in Gap-filling of missing information.

4.1.1. Application of various Gap-filling techniques

Table 4-1 shows results of filling the gap of missing information (NaN) from Farm 1 using the Nearest Neighbour forward Interpolation (NNFI), Nearest Neighbour Backward Interpolation (NNBI), Linear Interpolation (LI) and Gaussian Interpolation (GI). It can be observed that, Sentinel-2 image downloaded on 9th September 2016, of Farm 1 covered with clouds (NaN). NNFI Gap-filling the NaN with a value of 0.266 and NNBI Gap-filling it with value 0.257. While LI filled it with value 0.261 and finally GI filled it with value 0.263. Likewise, in the image of 9th October 2016 in Farm 1, there was missing information (NaN) due to cloud cover, but in the table as observed the NNFI Gap-filling it with value 0.257 and NNBI Gap-filling it with value 0.226. While LI filled it with value 0.236 and finally GI filled it with value 0.232. Therefore, from the Table 4-1, all NaN values have been gap-filling, as observed the constructed values depending on the type of algorithm used. The corresponding graph for the gap filling of missing information of Farm 1 is found in Appendix VII

Table 4-1: Gap-filling result for various interpolation techniques in Farm 1

	REFLECTANCE					NDVI				
Image Dates	NaN	NNFI	NNBI	LI	GI	NaN	NNFI	NNBI	LI	GI
1-Jul	0.295	0.295	0.295	0.295	0.295	0.498	0.498	0.498	0.498	0.498
11-Jul	NaN	0.295	0.235	0.283	0.288	NaN	0.498	0.263	0.451	0.478
21-Jul	NaN	0.295	0.235	0.271	0.274	NaN	0.498	0.263	0.404	0.427
31-Jul	NaN	0.295	0.235	0.259	0.256	NaN	0.498	0.263	0.357	0.363
10-Aug	NaN	0.295	0.235	0.247	0.241	NaN	0.498	0.263	0.31	0.302
20-Aug	0.235	0.235	0.235	0.235	0.235	0.263	0.263	0.263	0.263	0.263
30-Aug	0.266	0.266	0.266	0.266	0.266	0.246	0.246	0.246	0.246	0.246
9-Sep*	NaN*	0.266*	0.257*	0.261*	0.263*	NaN*	0.246*	0.271*	0.259*	0.258*
19-Sep	0.257	0.257	0.257	0.257	0.257	0.271	0.271	0.271	0.271	0.271
29-Sep	NaN	0.257	0.226	0.246	0.246	NaN	0.271	0.281	0.275	0.276
9-Oct*	NaN*	0.257*	0.226*	0.236*	0.232*	NaN*	0.271*	0.281*	0.278*	0.28*
19-Oct	0.226	0.226	0.226	0.226	0.226	0.281	0.281	0.281	0.281	0.281
29-Oct	0.238	0.238	0.238	0.238	0.238	0.257	0.257	0.257	0.257	0.257
8-Nov	0.245	0.245	0.245	0.245	0.245	0.295	0.295	0.295	0.295	0.295
18-Nov	0.248	0.248	0.248	0.248	0.248	0.332	0.332	0.332	0.332	0.332
28-Nov	0.246	0.246	0.246	0.246	0.246	0.422	0.422	0.422	0.422	0.422

* Are the value taken as an example in explanation of the gap-filled values

Table 4-2, shows the different interpolators how they Gap-filling the missing information according to the algorithm used. All the 7 NaN values has been Gap-filling from 9 cloud free images as observed in Table 4-1. The average value for NNFI is 0.263 while NNBI is 0.244, LI is 0.254 and GI is 0.253. Checking the

Minimum reflectance values from all interpolators NNFI is 0.226 while NNBI is 0.226, LI is 0.226 and GI is 0.226. Checking the Maximum reflectance values from all interpolators NNFI is 0.295 while NNBI is 0.295, LI is 0.295 and GI is 0.295.

Table 4-2: Descriptive statistics after Gap-filling Farm 1 Table: 4-1

	Nearest Forward	Nearest Backward	Linear	Gaussian	NDVI Nearest Forward	NDVI Nearest Backward	Linear NDVI	Gaussian NDVI
AVERAGE:	0.263	0.244	0.254	0.253	0.353	0.297	0.325	0.328
STD DEV:	0.024	0.018	0.018	0.020	0.109	0.068	0.078	0.083
N	16	16	16	16	16	16	16	16
MEDIAN:	0.26	0.24	0.25	0.25	0.29	0.27	0.29	0.29
MIN:	0.226	0.226	0.226	0.226	0.246	0.246	0.246	0.246
MAX:	0.295	0.295	0.295	0.295	0.498	0.498	0.498	0.498
Q 1:	0.246	0.235	0.243	0.240	0.269	0.263	0.269	0.269
Q 3:	0.295	0.250	0.262	0.263	0.498	0.285	0.369	0.378
CLOUD FREE	9	9	9	9	9	9	9	9
CLOUDED	7	7	7	7	7	7	7	7

Table 4-3 shows results of gap-filling missing information (NaN) from farm 6 using the NNFI, NNBI, LI and GI. It can be observed that, Sentinel-2 image downloaded on the 22nd February 2016, of farm 6 covered with clouds (NaN). NNFI Gap-filling the NaN value with value 0.359 and NNBI Gap-filling it with value 0.285. While LI filled it with value 0.320 and finally GI filled it with value 0.323. Likewise, the in the image of 2nd April 2016 in farm 6, there was missing information but the NNFI gap-fill it with value 0.285, NNBI with value 0.248, LI with value 0.257 and GI with value 0.251. The corresponding graph for the gap filling of missing information of farm 6 is found in Appendix VIII

Table 4-3: Gap-filling result for various interpolation techniques in Farm 6

	REFLECTANCE					NDVI				
Image Dates	NaN	NNFI	NNBI	LI	GI	NaN	NNFI	NNBI	LI	GI
3-Jan	NaN	0.263	0.263	0.245	0.245	NaN	0.175	0.175	0.226	0.226
13-Jan	0.263	0.263	0.263	0.263	0.263	0.175	0.175	0.175	0.175	0.175
23-Jan	0.281	0.281	0.281	0.281	0.281	0.123	0.123	0.123	0.123	0.123
2-Feb	NaN	0.281	0.359	0.320	0.326	NaN	0.123	0.141	0.132	0.132
12-Feb	0.359	0.359	0.359	0.359	0.359	0.141	0.141	0.141	0.141	0.141
22-Feb	NaN*	0.359*	0.285*	0.320*	0.323*	NaN*	0.141*	0.140*	0.141*	0.141*
3-Mar	0.285	0.285	0.285	0.285	0.285	0.140	0.140	0.140	0.140	0.140
13-Mar	NaN	0.285	0.248	0.276	0.271	NaN	0.140	0.227	0.162	0.154
23-Mar	NaN	0.285	0.248	0.267	0.259	NaN	0.140	0.227	0.184	0.184
2-Apr	NaN*	0.285*	0.248*	0.257*	0.251*	NaN*	0.140*	0.227*	0.206*	0.214*
12-Apr	0.248	0.248	0.248	0.248	0.248	0.227	0.227	0.227	0.227	0.227
22-Apr	0.266	0.266	0.266	0.266	0.266	0.199	0.199	0.199	0.199	0.199
2-May	NaN	0.266	0.283	0.267	0.268	NaN	0.199	0.597	0.225	0.204
12-May	NaN	0.266	0.283	0.268	0.270	NaN	0.199	0.597	0.252	0.218
22-May	NaN	0.266	0.283	0.269	0.271	NaN	0.199	0.597	0.278	0.240

* Are the value taken as an example in explanation of the gap-filled values

Table 4-4, shows the different interpolators in what way they Gap-filling the missing information according to the algorithm used. All the 9 NaN values were Gap-filling from 6 cloud free images as observed in the Table 4-3. The average value for NNFI is 0.284 while NNBI is 0.280, LI is 0.280 and GI is 0.279. Checking the Minimum reflectance values from all interpolators NNFI is 0.248 while NNBI is 0.248, LI is 0.245 and GI is 0.245. Checking the Maximum reflectance values from all interpolators NNFI is 0.359 while NNBI is 0.359, LI is 0.359 and GI is 0.359. The same procedure of Gap-filling was performed in all 10 farms.

Table 4-4: Descriptive statistics after Gap-filling Farm 6 in Table 4-3

	Nearest Forward	Nearest Backward	Linear	Gaussian	NDVI Nearest Forward	NDVI Nearest Backward	Linear NDVI	Gaussian NDVI
AVERAGE:	0.284	0.280	0.280	0.279	0.164	0.262	0.187	0.181
STD DEV:	0.032	0.035	0.031	0.032	0.033	0.177	0.048	0.040
N	15	15	15	15	`15	15	15	15
MEDIAN:	0.281	0.281	0.268	0.270	0.141	0.199	0.184	0.184
MIN:	0.248	0.248	0.245	0.245	0.123	0.123	0.123	0.123
MAX:	0.359	0.359	0.359	0.359	0.227	0.597	0.278	0.240
Q 1:	0.266	0.256	0.265	0.261	0.140	0.141	0.141	0.141
Q 3:	0.285	0.284	0.283	0.283	0.199	0.227	0.225	0.216
CLOUD FREE	6	6	6	6	6	6	6	6
CLOUDED	9	9	9	9	9	9	9	9

4.1.2. Evaluation of Gap-filling techniques in dynamic land cover (Farms)

This section involves the results of evaluation (accuracy test) of different Gap-filling techniques. Table 4-5 shows actual values of Farms 1 in comparison to the values generated by the interpolators. Example, the actual values for row 5 in the table was 0.313 but the gap-filled values from the NNFI was 0.280, NNBI was 0.295, LI was 0.287 and GI was 0.288. Likewise, in row 9, the actual value was 0.266 but the gap filled to 0.235, 0.257, 0.246 and 0.246, respectively for NNFI, NNBI, LI and GI.

Table 4-5: Actual values and its gap-filled values for Farm 1

Actual existed Farm Values from LOOCV	Gap-filling Values by NNFI	Gap-filling Values by NNBI	Gap-filling Values by LI	Gap-filling Values by GI
0.232	0.294	0.294	0.311	0.311
0.294	0.232	0.277	0.255	0.261
0.277	0.294	0.280	0.287	0.287
0.280	0.277	0.313	0.295	0.295
0.313	0.280	0.295	0.287	0.288
0.295	0.313	0.295	0.303	0.307
0.295	0.295	0.235	0.267	0.262
0.235	0.295	0.266	0.280	0.283
0.266	0.235	0.257	0.246	0.246
0.257	0.266	0.226	0.246	0.246
0.226	0.266	0.238	0.247	0.245
0.238	0.226	0.245	0.236	0.235
0.245	0.238	0.248	0.243	0.244
0.248	0.245	0.246	0.246	0.246
0.246	0.248	0.248	0.250	0.250

Table 4-6, shows the results of statistical test of the interpolation techniques for the Gap-filling of the missing information (NaN) in Farm 1. From the Table 4-6, NNBI recorded MSE of 0.001, RMSE of 0.027,

MAE of 0.019, and MAPE of 7.32% and an accuracy of 92.67%. Accuracy values obtained for LI are 0.001, 0.029, 0.021, 8.07% and 91.93% respectively for MSE, RMSE, MAE MAPE and accuracy. The NNBI and LI statistical tests do not differ much as observed as they both have almost the same MSE, RMSE, MAE, MAPE and Accuracy.

Table 4-6: Farm 1, Statistical test of Gap-filling techniques

	NNFI	NNBI	LI	GI
MSE	0.001	0.001	0.001	0.001
RMSE	0.032	0.027	0.029	0.029
MAE	0.024	0.019	0.021	0.021
MAPE %	9.25	7.32	8.07	8.08
ACCURACY %	90.75	92.68	91.93	91.92

The same procedure was performed to test Gap-filling techniques in all 10 farms.

4.1.3. Evaluation of Gap-filling techniques in static land cover type.

The Table 4-7, represents the actual and gap-filled values of the images from built-up. In row 7 of the table, the actual value was 0.198 but the gap-filled values from the NNFI was 0.211, NNBI was 0.189, LI was 0.200 and GI was 0.198

Table 4-7: Actual known values and its corresponding gap-filled values of Built-up

Actual Farm Values from LOOCV	**Gap-filling Values by NNI**	**Gap-filling Values by NNB**	**Gap-filling Values by LI**	**Gap-filling Values by GI**
0.223	0.217	0.217	0.217	0.217
0.217	0.223	0.216	0.220	0.220
0.216	0.217	0.214	0.216	0.215
0.214	0.216	0.219	0.218	0.218
0.219	0.214	0.211	0.213	0.213
0.211	0.219	0.198	0.209	0.211
0.198	0.211	0.189	0.200	0.198
0.189	0.198	0.190	0.194	0.192
0.190	0.189	0.202	0.196	0.193
0.202	0.190	0.190	0.190	0.190

The Table 4-8, represent the results of statistical test of interpolators in refilling the images that captured over built up. From the Table 4-8, GI recorded MSE of 0.00002, RMSE of 0.005, MAE of 0.003, and MAPE of 1.88% and an accuracy of 98.11%. Accuracy values obtained for LI are 0.00003, 0.0056, 0.004, 2.33% and 97.66% respectively for MSE, RMSE, MAE MAPE and total accuracy The LI and GI statistical tests as observed they both have the same MSE, RMSE, MAE, MAPE and Accuracy

Table 4-8: Statistical test of Gap-filling techniques over built-up

	NNFI	NNBI	LI	GI
MSE	0.000	0.000	0.000	0.000
RMSE	0.007	0.008	0.005	0.005
MAE	0.006	0.006	0.004	0.003
MAPE %	3.10	3.38	2.33	1.88
ACUURACY %	96.89	96.61	97.66	98.11

The same procedure was performed to test Gap-filling techniques in all 10 static land covers. The validation results for all farms and all static land cover found in Table 4-9.

Table 4-9, shows the results of accuracy test of all gap-filling techniques in all static and dynamic land covers selected. The general accuracy results shows that GI has highest mean absolute percentage accuracy of 90.98 followed by LI with 90.95, NNBI 90.29 and NNFI 89.83. Also the GI was best by having lowest Root mean square error of 0.0243, followed by LI 0.0244, NNBI 0.0262 and NNFI 0.0272

Table 4-9: Accuracy test results of gap filling techniques in all land covers

	RMSE				ACCURACY %			
LAND COVER	NNFI	NNBI	LI	GI	NNFI	NNBI	LI	GI
Farm 1	0.032	0.027	0.029	0.029	90.75	92.68	91.93	91.92
Farm 2	0.038	0.038	0.034	0.035	89.97	90.29	91.18	90.98
Farm 3	0.032	0.033	0.028	0.030	90.31	90.31	91.21	90.59
Farm 4	0.024	0.025	0.021	0.021	92.01	91.63	92.72	92.59
Farm 5	0.039	0.038	0.031	0.029	90.18	90.97	92.10	92.82
Farm 6	0.035	0.035	0.031	0.032	90.94	90.98	92.69	92.47
Farm 7	0.015	0.014	0.011	0.011	94.58	94.98	95.98	95.98
Farm 8	0.029	0.028	0.023	0.024	90.86	92.06	93.32	93.18
Farm 9	0.037	0.032	0.028	0.026	85.10	87.01	88.98	89.23
Farm 10	0.058	0.057	0.043	0.042	88.55	90.03	89.87	90.06
Water 1	0.026	0.024	0.028	0.028	83.70	84.67	85.04	85.22
Water 2	0.032	0.031	0.034	0.033	84.96	84.56	83.57	83.72
Water 3	0.027	0.026	0.027	0.027	86.84	87.41	85.88	86.20
Water 4	0.031	0.029	0.030	0.030	81.51	81.24	83.52	83.16
Water 5	0.030	0.029	0.029	0.029	81.98	81.37	84.51	84.20
Built 1	0.008	0.008	0.006	0.005	96.89	96.61	97.67	98.11
Built 2	0.010	0.010	0.014	0.013	94.93	94.88	94.10	94.20
Built 3	0.009	0.009	0.011	0.012	97.41	97.32	96.11	95.81
Built 4	0.010	0.010	0.007	0.007	95.58	95.59	97.34	97.39
Built 5	0.023	0.021	0.023	0.023	89.50	91.21	91.36	91.78
Average	0.0272	0.0262	0.0244	0.0243	89.83	90.29	90.95	90.98

4.1.4. Interpolators Significant test Results

The repeated measure ANOVA found that there was statistical significant different between interpolators with an $F(3) = 9.787$; $p < 0.05$. Led to post-hoc test and control for Bonferroni with the result in Table 4-10, shows that is only NNFI has significant statistical different with other interpolators. Nevertheless, NNBI, LI and GI there is no statistical significant different between them.

Table 4-10: Statistical significance test results of the interpolators used

(I) Accuracy	(J) Accuracy	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1 NNFI	2	.463	.182	.121	-.074	.999
	3	1.127*	.303	.009	.233	2.020
	4	1.153*	.313	.009	.232	2.074
2 NNBI	1	-.463	.182	.121	-.999	.074
	3	.664	.281	.174	-.164	1.492
	4	.690	.272	.120	-.110	1.491
3 LI	1	-1.127*	.303	.009	-2.020	-.233
	2	-.664	.281	.174	-1.492	.164
	4	.027	.072	1.000	-.185	.238
4 GI	1	-1.153*	.313	.009	-2.074	-.232
	2	-.690	.272	.120	-1.491	.110
	3	-.027	.072	1.000	-.238	.185

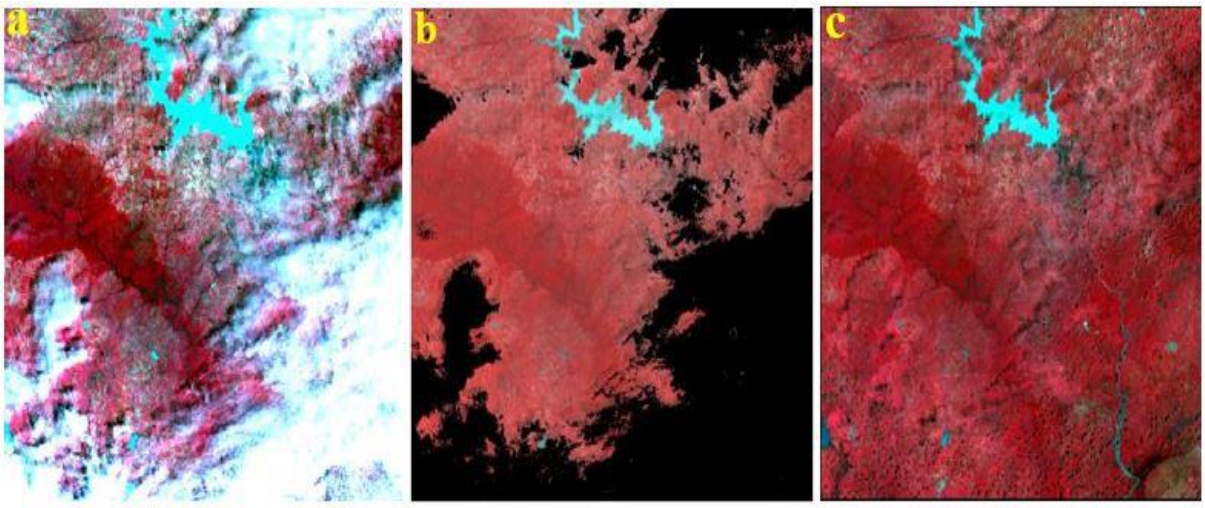
Based on estimated marginal means

*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: Bonferroni.

4.1.5. Actual Gap-filling by using the python script developed

Figure 4-1 shows the result of actual Gap-filling from the python script with the nearest neighbour and linear interpolation algorithms. Figure 4-1 (a) and (d) are TOA reflectance level 1C with cloud cover in images. The Figure 4-1 (b) and (d) are the Level 2A images where all the cloud cover surrounding has been masked. Finally, Figure 4-1 (c) and (f) are level 3A images whereby, the missing information has been gap-filled by using the python script developed with the tested interpolators.



29 September 2016. Level 1C *Level 2A Image* *Level 3A image*

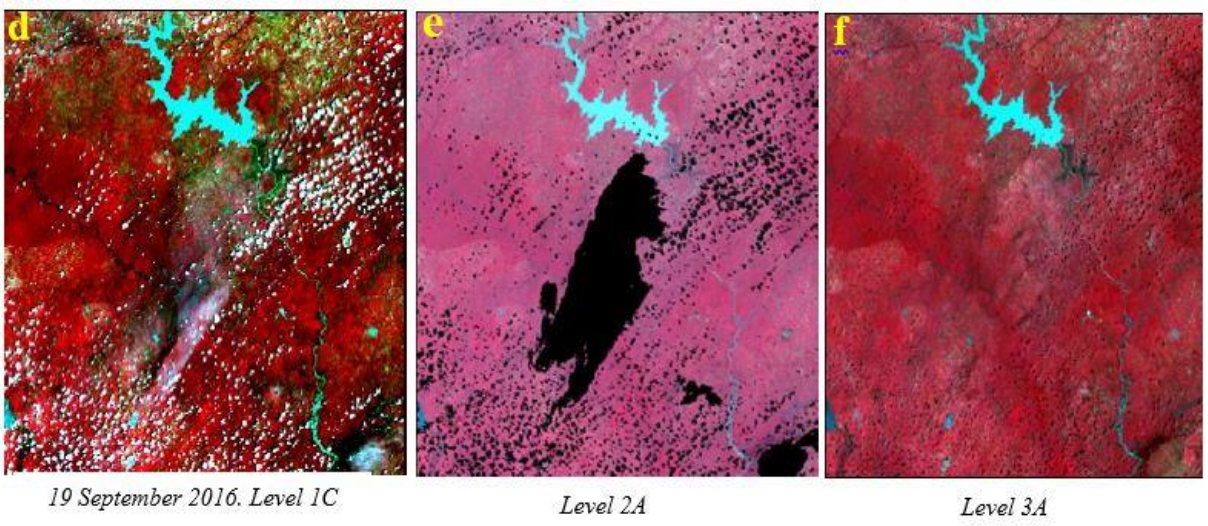


Figure 4-1: Level 3A images produced by python script developed

4.2. Objective 2

To assess the accuracy of interpolators in gap-filling of missing information in different crop growth rates.

.

Table 4-11, shows the result of all Gap-filled portions of crop growth rates of farm 1 from LOOCV. The results comes from all interpolators in each growth rate portion, ready for application of prediction statistical tests. In addition, it shows the actual values and its constructed values. Example from the linear portion the row 1, the actual value was 0.227 but the gap-filling values from the NNFI was 0.192, NNBI was 0.264, LI was 0.228 and GI was 0.233.

Table 4-11: Actual and its gap-filled values of all phenology curve portions of Farm 1

Portion	Actual Farm Values from LOOCV	Gap-filling Values by NNI	Gap-filling Values by NNB	Gap-filling Values by LI	Gap-filling Values by GI
Linear	0.227	0.192	0.264	0.228	0.233
	0.332	0.295	0.422	0.358	0.357
	0.422	0.332	0.494	0.413	0.402
	0.494	0.422	0.422	0.512	0.512
	0.192	0.227	0.227	0.190	0.190
	0.295	0.332	0.332	0.242	0.242
	0.264	0.227	0.281	0.254	0.255
	0.281	0.264	0.264	0.301	0.301
Stationary	0.264	0.281	0.281	0.306	0.306
	0.281	0.264	0.257	0.260	0.260
	0.257	0.281	0.281	0.298	0.298
	0.281	0.257	0.257	0.219	0.219
	0.257	0.281	0.295	0.288	0.288
	0.295	0.257	0.257	0.233	0.233
Exponential	0.422	0.494	0.494	0.506	0.506
	0.494	0.422	0.482	0.452	0.463
	0.482	0.494	0.430	0.462	0.474
	0.430	0.482	0.482	0.470	0.470

	0.430	0.463	0.463	0.448	0.448
	0.463	0.430	0.478	0.454	0.454
	0.478	0.463	0.463	0.496	0.496

The Table 4-12, represent the results of statistical test of interpolators in Gap-filling the linear growth rate portions. Here were the images that formed the linear growth portions, the results shows that LI out performed all other interpolator in gap -filling the linear growth rate portions by obtaining, low RMSE of 0.001 and higher accuracy of 97.22%.

Table 4-12: Statistical test results of interpolators accuracy in linear growth rate

	RMSE				ACCURACY			
Growth rate	NNFI	NNBI	LI	GI	NNFI	NNBI	LI	GI
linear	0.035	0.035	0.002	0.002	81.77	81.77	98.96	98.96
linear	0.035	0.037	0.001	0.006	84.58	83.70	99.56	97.46
linear	0.037	0.017	0.010	0.009	85.98	93.47	96.26	96.75
linear	0.090	0.072	0.009	0.020	78.68	82.92	97.88	95.23
linear	0.015	0.015	0.018	0.018	96.86	96.86	96.23	96.23
linear	0.023	0.023	0.011	0.011	90.01	90.01	95.35	95.35
linear	0.023	0.013	0.005	0.005	90.92	95.14	97.89	98.01
linear	0.013	0.016	0.002	0.003	95.37	93.94	99.28	98.81
linear	0.060	0.117	0.029	0.027	79.79	60.14	90.17	90.82
linear	0.117	0.099	0.009	0.010	71.50	75.93	97.79	97.69
linear	0.099	0.062	0.019	0.013	80.60	87.97	96.32	97.42
linear	0.039	0.044	0.003	0.013	92.70	91.76	99.53	97.55
linear	0.045	0.049	0.002	0.002	85.06	83.55	99.25	99.25
linear	0.013	0.022	0.004	0.003	95.19	91.88	98.68	98.85
linear	0.030	0.020	0.005	0.005	88.21	92.01	98.10	98.19
linear	0.035	0.037	0.001	0.006	84.58	83.70	99.56	97.36
linear	0.072	0.072	0.018	0.018	85.43	85.43	96.36	96.36
linear	0.052	0.052	0.040	0.040	87.91	87.91	90.70	90.70
linear	0.033	0.015	0.009	0.009	92.87	96.76	98.06	98.06
linear	0.018	0.017	0.000	0.001	93.30	93.67	99.85	99.47
linear	0.033	0.033	0.018	0.018	92.33	92.33	95.81	95.81
Average	0.044	0.041	0.010	0.011	87.32	87.66	97.22	96.87

Also in Table, 4-13 elaborates the images that formed the exponential growth portions. The results shows that GI out performed all other interpolator in gap -filling the exponential growth rate portions by obtaining, low RMSE of 0.011 and higher accuracy of 96.22%.

Table 4-13: Statistical test results of interpolators accuracy in exponential growth rate

	RMSE				ACCURACY			
Growth rate	NNFI	NNBI	LI	GI	NNFI	NNBI	LI	GI
Exponential	0.037	0.037	0.053	0.053	87.31	87.31	82.15	82.15
Exponential	0.037	0.090	0.026	0.025	88.74	72.90	92.08	92.53
Exponential	0.012	0.052	0.020	0.008	97.51	89.21	95.85	98.34
Exponential	0.033	0.015	0.009	0.009	92.87	96.76	98.06	98.06
Exponential	0.007	0.007	0.009	0.010	98.75	98.75	98.44	98.28
Exponential	0.012	0.008	0.002	0.003	95.08	96.76	99.16	98.77
Exponential	0.009	0.031	0.011	0.011	96.51	88.06	95.77	95.77
Exponential	0.003	0.033	0.015	0.015	99.04	88.24	94.60	94.60
Exponential	0.030	0.030	0.010	0.010	86.63	86.63	95.70	95.70
Exponential	0.028	0.028	0.016	0.016	90.61	90.61	94.45	94.45
Exponential	0.011	0.028	0.008	0.005	95.75	89.63	96.94	98.12

Exponential	0.008	0.011	0.002	0.002	96.96	95.57	99.30	99.31
Exponential	0.035	0.035	0.002	0.002	81.77	81.77	98.96	98.96
Exponential	0.037	0.017	0.010	0.009	85.98	93.56	96.21	96.59
Exponential	0.037	0.090	0.026	0.025	88.86	72.89	92.17	92.47
Exponential	0.090	0.072	0.009	0.020	78.67	82.94	97.87	95.26
Exponential	0.017	0.006	0.006	0.003	93.95	97.92	97.80	98.96
Exponential	0.009	0.024	0.007	0.004	96.21	89.96	96.87	98.37
Exponential	0.019	0.010	0.004	0.000	91.59	95.44	98.08	99.97
Exponential	0.031	0.035	0.004	0.005	83.29	81.09	97.96	97.48
Exponential	0.003	0.008	0.002	0.000	98.77	96.85	99.04	99.86
Average	0.024	0.032	0.012	0.011	91.66	89.18	96.07	96.38

Finally. in Table 4-14, were the images that formed the stationary growth rate portions, the results shows that Nearest neighbour out performed all other interpolator in gap -filling the stationary growth rate portions by obtaining, low RMSE of 0.024 and higher accuracy of 92.32%.

Table 4-14: Statistical test results of interpolators accuracy in stationary growth rate

	RMSE				ACCURACY			
Growth rate	NNFI	NNBI	LI	GI	NNFI	NNBI	LI	GI
stationary	0.017	0.017	0.020	0.020	93.87	93.87	92.98	92.98
stationary	0.024	0.038	0.031	0.031	90.53	85.34	87.93	87.93
stationary	0.072	0.012	0.042	0.031	85.41	97.57	91.49	93.77
stationary	0.033	0.033	0.018	0.018	92.33	92.33	95.81	95.81
stationary	0.016	0.052	0.034	0.041	94.28	81.82	88.05	85.71
stationary	0.052	0.060	0.056	0.057	77.78	74.68	76.23	75.86
stationary	0.062	0.039	0.050	0.050	89.26	93.19	91.23	91.23
stationary	0.044	0.044	0.005	0.005	91.02	91.02	98.98	98.98
stationary	0.007	0.055	0.008	0.009	98.76	90.36	98.50	98.38
stationary	0.002	0.001	0.002	0.002	99.07	99.46	99.27	99.33
stationary	0.001	0.001	0.004	0.004	99.45	99.45	98.52	98.52
stationary	0.000	0.004	0.002	0.002	99.81	98.68	99.24	99.24
stationary	0.030	0.000	0.015	0.015	88.63	99.81	94.41	94.41
stationary	0.002	0.002	0.014	0.014	99.33	99.33	95.63	95.63
stationary	0.002	0.016	0.007	0.010	99.33	94.98	97.83	96.76
stationary	0.020	0.002	0.011	0.009	92.60	99.09	95.85	96.54
stationary	0.005	0.040	0.023	0.022	97.62	82.02	89.82	90.06
stationary	0.085	0.004	0.045	0.050	62.68	98.35	80.51	78.14
stationary	0.009	0.046	0.027	0.023	96.79	83.22	90.37	91.65
stationary	0.017	0.017	0.042	0.042	93.56	93.56	84.09	84.09
stationary	0.024	0.024	0.041	0.041	90.66	90.66	84.05	84.05
Average	0.025	0.024	0.024	0.024	92.04	92.32	91.94	91.86

Interpolators significant test results for linear growth rate portions

The repeated measure ANOVA found that there was statistical significant different between interpolators with an $F(3) = 1.595$; $p < 0.05$. Led to post-hoc test and control for Bonferroni with the result in Table 4-15, which shows that is only NN has significant statistical different with other interpolators in gap-filling the linear portions. Nevertheless, between LI and GI there is no statistical significant different exists between them.

Table 4-15: Statistical significance test results of the interpolators used for linear growth rate portion

(I) Linear	(J) Linear	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
NNFI	NNBI	-.343	1.192	1.000	-3.833	3.146
	LI	-9.902*	1.435	.000	-14.104	-5.701
	GI	-9.557*	1.398	.000	-13.648	-5.466
NNBI	NNFI	.343	1.192	1.000	-3.146	3.833
	LI	-9.559*	1.687	.000	-14.497	-4.621
	GI	-9.213*	1.660	.000	-14.073	-4.354
LI	NNFI	9.902*	1.435	.000	5.701	14.104
	NNBI	9.559*	1.687	.000	4.621	14.497
	GI	.346	.218	.772	-.293	.984
GI	NNFI	9.557*	1.398	.000	5.466	13.648
	NNBI	9.213*	1.660	.000	4.354	14.073
	LI	-.346	.218	.772	-.984	.293

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Interpolators significant test results for stationary growth rate portion

In Table 4-16, the repeated measure ANOVA found that there was no statistical significant different between interpolators with an $F(3) = 1.418$; $p > 0.05$. Therefore, in gap-filling the stationary portions there is no significant different of interpolators accuracy.

Table 4-16: Statistical significance test results of the interpolators used for stationary growth rate portion

SOURCE		TYPE III SUM OF SQUARES	DF	MEAN SQUARE	F	SIG.	PARTIAL ETA SQUARED
STATIONARY	SPHERICITY ASSUMED	2.563	3	.854	.036	.991	.002
	GREENHOUSE-GEISSER	2.563	1.418	1.807	.036	.919	.002
	HUYNH-FELDT	2.563	1.495	1.714	.036	.928	.002
	LOWER-BOUND	2.563	1.000	2.563	.036	.851	.002
ERROR(STATIONARY)	SPHERICITY ASSUMED	1408.381	60	23.473			
	GREENHOUSE-GEISSER	1408.381	28.367	49.649			
	HUYNH-FELDT	1408.381	29.906	47.094			
	LOWER-BOUND	1408.381	20.000	70.419			

Interpolators significant test results for exponential growth rate portion

The repeated measure ANOVA found that there was statistical significant different between interpolators with an $F(3) = 1.932$; $p < 0.05$. Led to post-hoc test and control for Bonferroni with the result in Table

4-17, which shows that, is only NN has significant statistical different with other interpolators in gap-filling the exponential portions. Nevertheless, between LI and GI there is no statistical significant different exists between them.

Table 4-17: Statistical significance test results of the interpolators used for exponential growth rate portion

(I) EXPONENTIAL	(J) EXPONENTIAL	MEAN DIFFERENCE (I-J)	STD. ERROR	SIG.[B]	95% CONFIDENCE INTERVAL FOR DIFFERENCE[B]	
					LOWER BOUND	UPPER BOUND
	NNBI	2.476	1.419	.578	-1.676	6.629
NNFI	LI	-4.410*	1.425	.034	-8.580	-.240
	GI	-4.721*	1.329	.012	-8.611	-.832
	NNFI	-2.476	1.419	.578	-6.629	1.676
NNBI	LI	-6.886*	1.509	.001	-11.302	-2.470
	GI	-7.198*	1.469	.001	-11.497	-2.898
	NNFI	4.410*	1.425	.034	.240	8.580
LI	NNBI	6.886*	1.509	.001	2.470	11.302
	GI	-.311	.224	1.000	-.967	.344
	NNFI	4.721*	1.329	.012	.832	8.611
GI	NNBI	7.198*	1.469	.001	2.898	11.497
	LI	.311	.224	1.000	-.344	.967

BASED ON ESTIMATED MARGINAL MEANS
*. THE MEAN DIFFERENCE IS SIGNIFICANT AT THE .05 LEVEL.
B. ADJUSTMENT FOR MULTIPLE COMPARISONS: BONFERRONI.

5. DISSCUSSIONS

Results from the study reveals success in the Gap-filling of missing information in Sentinel-2 satellite imageries due to cloud cover. This was achieved because Sentinel-2 images have a temporal resolution of 10 days, thus providing continues observations in the time series space. Temporal gap-filling techniques have been used over the years in the gap-filling of missing information in time series data sets e.g. MODIS and AVHRR (Aleixandre et al., 2011; Kandasamy et al., 2013; Moody et al., 2005; Moreno et al., 2014; Yin et al., 2016). Jönsson & Eklundh (2004), developed TIMESAT software which uses temporal approach to gap-filling time-series images by built-in temporal filters such as Savitzky-Golay and asymmetric Gaussian filters. Also, Inglada et al., (2015) used temporal linear interpolation to gap-fill missing value in a time series data, which were used to calculate spectral indices such as brightness, NDVI and NDWI.

The study reveals a minimum accuracy of 89.83% for Nearest neighbour forward interpolation and a maximum accuracy of 90.98% for Gaussian interpolation technique. The high accuracies in gap-filling were attained because, all image that were used in the process were atmospherically corrected, clouds were removed and the images were converted to surface reflectance. This helped to bring similarities to images captured on different dates and hence increased the accuracy in gap-filling of missing information process (Agapiou et al., 2011; Hadjimitsis et al., 2010). In addition, the Sentinel-2 has a temporal resolution of 10 days indicating a relatively little land cover changes (more especially for the temporary static land cover classes e.g. built-up) between image dates.

The presence to absence ratio between cloud free and clouded pixels of a farm in time series images did not affect the accuracy of gap-filling process. Because from farm 1 to farm 5, the number of cloud free pixels of farms was higher than the number of clouded pixels, the maximum accuracy observed was 92.98%, and the minimum was 89.97%. Also for farm 6 to farm 10 the clouded pixels of a farms was higher than cloud free pixels and the maximum accuracy observed was 95.98% and the minimum was 85.10%.

Despite the high accuracies (>80%) observed for all interpolation algorithms, Linear interpolation and Gaussian interpolation techniques recorded statistically significant accuracies compared to nearest neighbour. A repeated measure ANOVA indicated a statistically significant difference at the $p < .05$ level in accuracy scores for the four accuracy measures: $F(3,) = 9.787$: $p = .02$. A Post-hoc comparisons using the Bonferroni test indicated that the mean score for Nearest neighbour forward interpolation (M= 10.17, SD= 4.71) was significantly different from Linear interpolation (Mean = 9.05, SD = 4.47) and Gaussian interpolation (Mean = 9.01, SD = 4.50). Nearest neighbour backward interpolation, linear interpolation and Gaussian interpolation did not differ significantly from each other. Also, Nearest neighbour forward interpolation did not significantly differ from Nearest neighbour backward interpolation (Mean= 9.71, SD= 4.69).

The most accurate interpolator in general gap-filling was Gaussian interpolation. The Gaussian interpolator performed well in gap-filling of missing information by an accuracy of 90.98% and RMSE of 0.0243. Gaussian interpolation performed better due to the nature of rise and fall in reflectance values of temporarily dynamic and static land cover classes. According to Jönsson & Eklundh (2002), the Gaussian function is very flexible and have the ability to follow the narrow peaks of time series data. The rise and fall in reflectance values across the different land cover classes is a replica of narrow peaks that Gaussian algorithm models. Also, USRA (2016) elaborated that, the reflectivity and the absorbance of light by a substance is highly dependent on its molecular makeup. Therefore, everyday, over time in every substance, there are changes that occur in the substance molecular make up; due to both physical and chemical earth processes

that causes small variations of reflectance values. This variation in the molecular makeup of substances results in relative fluctuations in reflectance levels, hence Gaussian algorithm can best model these variations.

Also, Linear interpolation performed better by an accuracy of 90.95% and RMSE of 0.024 because, it is estimating the values between two points in straight line and the polynomial degree is fixed to one without considering nature and distribution of data (Kandasamy et al., 2013). Aleixandre et al., (2011) proposed that, linear interpolation could be used in time series gap-filling, because, it is efficient in gap-filling of time series data. They used linear interpolation to fill the gap and smoothed time series to improve the reliability and continuity of Leaf Area Index (LAI) data. According to Hocke & Kampfer (2009) and Jönsson & Eklundh (2002), the performance of temporal interpolators in gap filling and noise reduction depends to the nature of data i.e. whether it is time series data or small data set. Hocke & Kampfer (2009) suggested using of temporal linear interpolation in, gap-filling, smoothing and noise reduction of time series data or combining linear interpolation, filtering, spectral approaches in gap-filling of time series data. Though linear interpolation performed well in time series, a plethora of research reveals some challenges of poor performance when number of images in the time series dataset are more than 128 (Aleixandre et al., 2011; Kandasamy et al., 2013; Verger et al., 2013).

However, to gap-fill temporarily, dynamic land covers (i.e. crop farms), the growth stages (rates) of the crop as depicted by the phenology curve, dictates the choice of interpolation algorithm. From the results, Gaussian interpolation out performed others by obtaining higher accuracy of 96.38% in the exponential segment of the phenology curve compared to Nearest neighbour forward that obtained accuracy of 91.66% Nearest neighbour backward that obtained accuracy of 89.18%. In addition, linear interpolation obtained accuracy of 96.07%. This is because the Gaussian function has been designed with a truncated exponential function that helped to detect the exponential direction in the portion of phenology curve. This makes it flexible to detect the narrow peaks and capture the rise and fall of data (Jönsson & Eklundh, 2002, 2004; Lan & Jorgenson, 2001; Revolvy, 2016). The designed truncated exponential function within the Gaussian algorithm helped to detect the exponential direction in the portion of phenology curve and cause the Gaussian interpolation to perform better in gap-filling in this portion than other interpolators used in the study.

In the linear segment of the phenology curve, linear interpolation out performed others by obtaining higher accuracy of 97.21%, compared to Nearest neighbour forward that obtained accuracy of 87.31%, Nearest neighbour backward that obtained accuracy of 87.65% and Gaussian interpolation obtained accuracy of 96.87%. The linear interpolation performed well in linear segment of phenology curve because of the linear format of data point. This is because, the linear equation developed within the algorithm fits two points in straight line and the polynomial degree is fixed to one without considering nature and distribution of data (Kandasamy et al., 2013; Kendall & Han, 2003; Paulbourke, 1999). Therefore, the linear interpolation fits well the data points that forms linear relationships the (planar straight line). In addition, good performance of linear interpolation to linear portion was because of the regression best fit line characteristics (linear equation) (Campbell, 2008; Nau, 2014).

Finally, in the stationary portions of crop growth rates, all the interpolation techniques were evaluated in gap-filling the portion. The Nearest neighbour out performed other interpolators by having the higher accuracy of 92.32% for Nearest backward and 92.03% for Nearest forward; compared to the linear interpolation that obtained the accuracy of 91.94% and Gaussian, interpolation obtained the accuracy of 91.86%. According to OTexts (2017) the stationary curves involves data series that they have almost similar values and the time plot will form the horizontal line characteristics. Since the stationary portion involves

the similarities of data points in time series, causes the Nearest neighbour to perform well in the stationary portion. Because in gap-filling Nearest neighbour it takes the actual values as it is to predict the new data point. Therefore, the actual values and predicted values will not have different; the error will be very low and cause the nearest neighbour to perform well in stationary portion of phenology curve than other interpolation technique.

Despite the high accuracy performance of Gaussian interpolation observed in gap-filling the exponential growth rate, a repeated measure ANOVA indicated a statistically significant difference at $p < 0.05$ level in accuracy scores for the four accuracy measures: $F(3,) = 1.932$: $p = 0.00$. A Post-hoc comparisons using the Bonferroni test indicated that, the Nearest neighbour interpolation for both forward and backward, were statistically significant different from linear and Gaussian interpolation. Therefore, linear interpolation can also be used to gap-fill images that form exponential growth rate. Since its accuracy statistically did not differ from the Gaussian interpolation.

Regardless of the higher accuracy performance of linear interpolation observed in gap-filling the linear growth rate, a repeated measure ANOVA indicated a statistically significant difference at $p < 0.05$ level in accuracy scores for the four accuracy measures: $F(3,) = 1.595$: $p = 0.00$. A Post-hoc comparisons using the Bonferroni test indicated that, the Nearest neighbour interpolation for both forward and backward, were statistically significant different from linear and Gaussian interpolation. Therefore, Gaussian interpolation can also be used to gap-fill images that form linear growth rate. Since its accuracy statistically did not differ from the linear interpolation.

Irrespective of the higher accuracy performance of Nearest neighbour interpolation observed in gap-filling the stationary growth rate, a repeated measure ANOVA indicated no statistically significant difference at $p > 0.05$ level in accuracy scores for the four accuracy measures: $F(3,) = 1.418$: $p = 0.919$. Therefore, both Nearest neighbour, linear and Gaussian interpolation can also be used to gap-fill images that form stationary growth rate. Since their accuracies statistically did not differ from each other.

6. CONCLUSION AND RECOMMENDATIONS

6.1. Conclusion

The act of cloud cover formation in the atmosphere is permanent natural phenomena in our planet earth. It causes a lot of information missing and information deterioration in optical satellite imageries. These cloud effects have been observed more to affect Sentinel-2 due to its altitude. The advantageous characteristics of Sentinel-2, such as open source download availability for all individuals and entities globally, 10m spatial resolution. 5 days' revisit time, richness in spectral bands especially four red-edge bands for monitoring of vegetation and the global capturing of important land ocean processes. These characteristics will continue make Sentinel-2 the most dependent and used satellite imagery. Since its launch in June 2015 and March 2017 the number of downloads from ESA has reach up to 52600 in 24 hours (Copernicus Open Access Hub, 2017).

In line with the benefits that comes with Sentinel-2 and given the accompanying challenges of cloud cover in the tropics, makes the gap-filling of missing information in Sentinel-2 images a very important scientific action to be taken. This project has demonstrated success in cloud removal in Sentinel-2 images. Similarly, success in temporal gap-filling of Sentinel-2 time series images using Gaussian and Linear temporal interpolation techniques obtained an accuracy of 90.98% and 90.95% respectively. In addition, the Nearest neighbour also acquired good accuracy of 90.29% for Nearest forward and 89.83% for Nearest backward. This answers the first question stated that *"which interpolation technique performs better in gap-filling of missing information in sentinel-2 images?"* Whereby, the Gaussian interpolation proved to perform better than others by acquiring higher accuracy of 90.98% and the lowest root mean square error of 0.0243

However, the statistical significant different test found that, there was significant different between the interpolators used. Where by Nearest neighbour forward interpolation was significantly different from all interpolators used, such as linear interpolation and Gaussian interpolation. Nevertheless, Nearest neighbour backward interpolation, linear interpolation and Gaussian interpolation did not differ significantly from each other. The test for significant different between the interpolators led to an *acceptance* of first alternative hypothesis stated that "*there is significant different in accuracies of interpolators in gap-filling Sentinel-2 images"* thereby, in gap-filling of the missing information Gaussian, linear and Nearest neighbour backward can be used since there is no significant different in gap-filling accuracy that exists between them.

For temporarily dynamic land cover classes i.e. crop farms, the interpolation algorithm that should be applied, must depend on the crop growth rates such as linear, stationary and exponential growth rates.

In gap-filling linear growth rates, the project demonstrate that, the linear interpolation should be used to gap-fill the missing information, since linear interpolator out performed other interpolators by acquired higher percentage accuracy of 97.21% and lower root mean square error of 0.010. This answers the question *"Which interpolation technique outperform others in gap-filling the missing information in terms of linear crop growth rates?"* Therefore, for linear growth rates the linear interpolation should be used to gap-fill the missing information.

The significant test of interpolators accuracies in gap-filling the missing information in linear growth rates, led to an *acceptance* of the alternative hypothesis which stated that "*There is significant different in accuracies of interpolators in gap-filling the linear crop growth rates."* The results showed that, there is significant different between interpolators accuracies in gap-filling the linear growth rates, where, the Nearest neighbour interpolation is significantly different from the linear and Gaussian interpolation. Therefore, in gap-filling

the linear growth rates portions, both Gaussian and linear interpolation can be used, since there is no significant different exists between them in gap-filling this linear crop growth rates.

In gap-filling stationary growth rates, the project demonstrate the Nearest neighbour interpolation should be used to gap-fill the missing information, since Nearest neighbour interpolator out performed other interpolators by acquired higher percentage accuracy of 92.32% and 92.03% and lower root mean square error of 0.024 and 0.025. This answers the question *"Which interpolation technique outperform others in gap-filling the missing information in terms of stationary crop growth rates?"* Therefore, for stationary growth rates, the Nearest neighbour interpolation should be used to gap-fill the missing information.

The significant test of interpolators accuracies in gap-filling the missing information in stationary growth rates led to an *acceptance* of the *null* hypothesis which stated that "*There is no significant different in accuracies of interpolators in gap-filling the stationary crop growth rates."* Subsequently, the results showed that, there is no significant different between interpolators accuracies in gap-filling the stationary growth rates. Therefore, in gap-filling the stationary growth rates portions, both Nearest neighbour, Gaussian and linear interpolation can be used, because, there is no significant different exists between them in gap-filling the stationary crop growth rate.

Finally, in gap-filling exponential growth rates, the project shows the Gaussian interpolation should be used to gap-fill the missing information, since Gaussian interpolator out performed other interpolators by acquired higher percentage accuracy of 96.38% and lower root mean square error of 0.011. This answers the research question *"Which interpolation technique outperform others in gap-filling the missing information in terms of exponential crop growth rates?"* Therefore, for exponential growth rates the Gaussian interpolation should be used to gap-fill the missing information.

The significant test of interpolators accuracies in gap-filling the missing information in exponential growth rates led to an *acceptance* of the *alternative* hypothesis which stated that "*There is significant different in accuracies of interpolators in gap-filling the exponential crop growth rates."* Afterward, the results showed that, there is significant different between interpolators accuracies in gap-filling the exponential growth rates whereby, the nearest neighbour interpolation is significantly different from the linear and Gaussian interpolation. Therefore, in gap-filling the exponential growth rate portions, both Gaussian and linear interpolation can be used, because, there is no significant different exists between them in gap-filling exponential crop growth rate.

Since the clouds are contributing to information loss and data deterioration, brings challenges in implementation of ESA (European space agency) operational applications goals such as continuous monitoring of vegetation (crops and forests) and land monitoring. With the presence of clouds in Sentinel-2 images, it is not possible to analyse the vegetation changes, quantification and monitoring of natural vegetation and global agriculture trend. Therefore, when the clouds are well removed and the lost information (gap) has been filled, it helps to revitalize the mission and goals of satellite launching. It provides again the continues cloud free images which are well corrected, and the signal to noise ratio has been secured. Now the important operational applications such as continuous monitoring of crops and forest can take place. Example accurate land cover quantification and land cover change analysis, identification of forest degradation and forest fragmentation and quantify the causatives of degradation and fragmentation. In addition, the gap-filled images can be used in planning, management and inventory mapping of forests. This is by analysing forest health, forest productivity, ecosystem and species diversity.

Furthermore, the gap-filled images resumes analysis of important agricultural application that froze due to cloud existence in Sentinel-2 images. Now the gap-filled images can be used in agriculture application such

as mapping of food source continuously, crop classification, mapping crop types or biophysical parameters, change detection and crop growth stage mapping. Also can be applied in monitoring of crop growth, yields estimation, calculate water shortage, erosion risk, nutrients deficiencies and generally assessing the crop health such as, disease, pest and insect invasion

6.2. Recommendations

The project recommending that, more study should be conducted to find a methodology to detect cloud shadows in Sentinel-2 images, because to some extent, dense shadows causes loss of information in Sentinel-2 images. Also in this study, after successful removing of clouds and gap-filling of missing information, still there were dark cloud shadow spots in images that brought noise in images. It is very difficult to detect cloud shadows in satellite images, but more difficult in Sentinel-2 due to lack of thermal infrared band. Cloud shadow detection needs thermal band to detect cloud height, in order to estimate and relate the distance between cloud and its cloud shadow. The difficulties of cloud shadow removal has been revealed in this study by applying several methodologies such as F-mask, ESA SNAP, ATCOR to detect cloud shadows in Sentinel-2 without success. F-mask in detecting cloud shadows, used the geometrical relationship between the cloud and cloud shadow but still it was difficult to detect. Therefore, to attain better clear gap-filled Sentinel-2 images, the cloud shadow must be masked.

Also, rather than depending only on the information of the nearest cloud free pixels in a time series to gap fill the missing information in Sentinel-2, the study recommend the application of other gap-filling information sources such as using of other optical or radar images that captured information at the same date and the same area. Also a stationary camera at the farm area that will take photos at steady intervals, the spectrometer that will record reflectance at the farm at steady intervals, expert knowledge from the agricultural extension officers can be used as the source of gap-filling information and finally, a developed App that will remind farmers to take observations at steady intervals.

7. LIST OF REFERENCES

Abdulla, F. (2014). A New (Proposed) Formula for Interpolation and Comparison with Existing Formula of Interpolation. *Mathematical Theory and Modeling*, *4*(4), 33–48.

Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C., & McGill, M. (2008). Cloud detection with MODIS. Part II: Validation. *Journal of Atmospheric and Oceanic Technology*, *25*(7), 1073–1086. https://doi.org/10.1175/2007JTECHA1053.1

Addink, E. A. (1999). A comparison of conventional and geostatistical methods to replace clouded pixels in NOAA-AVHRR images. *International Journal of Remote Sensing*, *20*(5), 961–977. https://doi.org/10.1080/014311699213028

Agapiou, A., Hadjimitsis, D. G., Papoutsa, C., Alexakis, D. D., & Papadavid, G. (2011). The Importance of accounting for atmospheric effects in the application of NDVI and interpretation of satellite imagery supporting archaeological research: The case studies of Palaepaphos and Nea Paphos sites in Cyprus. *Remote Sensing*, *3*(12), 2605–2629. https://doi.org/10.3390/rs3122605

Aleixandre Verger, Frederic, B., & Weiss, M. (2011). A multisensor fusion approach to improve LAI time series. *Remote Sensing of Environment*, *115*(10), 2460–2470. https://doi.org/10.1016/j.rse.2011.05.006

Amin, R., Gould, R., Hou, W., Lee, Z., & Arnone, R. (2011). Automated detection and removal of cloud shadows on HICO images. *Proceedings of SPIE - The International Society for Optical Engineering*, *8030*(803001–1), 1–10. https://doi.org/10.1117/12.887761

Arellano, P. (2003). Missing information in Remote Sensing: Wavelet approach to detect and remove clouds and their shadows, University of Twente, Faculty of geo-Information science and Earth observation(Master's Thesis) Retrieved from https://www.itc.nl/Pub/Home/library/Academic_output/AcademicOutput.html?p=2&y=4&l=20

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, *4*(0), 40–79. https://doi.org/10.1214/09-SS054

Baillarin, S., Gigord, P., & Hagolle, O. (2008). Automatic registration of optical images, a stake for future missions: Application to ortho-rectification, time series and mosaic products. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2*(1), 1112–1115. https://doi.org/10.1109/IGARSS.2008.4779194

Baillarin, S. J., Meygret, A., Dechoz, C., Petrucci, B., Lacherade, S., Tremas, T., … Spoto, F. (2012). Sentinel-2 level 1 products and image processing performances. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *XXXIX*(September), 7003–7006. https://doi.org/10.1109/IGARSS.2012.6351959

Barnston, A. G. (1992). Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score. *Weather and Forecasting*. https://doi.org/10.1175/1520-0434(1992)007<0699:CATCRA>2.0.CO;2

Blue leaf. (2016). Linear Interpolation with Excel. Retrieved January 9, 2017, from http://www.blueleafsoftware.com/Products/Dagra/LinearInterpolationExcel.php

Bryce Christensen. (2015). Cross-Validation for Genomic Prediction in SVS - Our 2 SNPs...®. Retrieved January 22, 2017, from http://blog.goldenhelix.com/bchristensen/cross-validation-for-genomic-prediction-in-svs/

Campbell, D., & Campbell, S. (2008). Statlab Workshop Introduction to Regression and Data Analysis. Retrieved from https://*statlab.stat.yale.edu/workshops/IntroRegression/StatLab-IntroRegressionFa08.pdf*

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Cihlar, J., & Howarth, J. (1994). Detection and removal of cloud contamination from AVHRR images. *IEEE Transactions on Geoscience and Remote Sensing*, *32*(3), 583–589. https://doi.org/10.1109/36.297976

CMMAP. (2016). Why are clouds white 1–3. Retrieved from https:// *littleshop.physics.colostate.edu/tenthings/Why%20are%20clouds%20white.pdf*

Copernicus Open Access Hub. (2017). Open Access Hub. Retrieved February 19, 2017, from https://scihub.copernicus.eu/

Daniel, Z. (2015). the Impact of Irrigation Schemes on Farmers'Income and Livelihoods in the Upper East Region of Ghana. Kwame Nkurumah University of science and Technology (Master's thesis) Retrieved from http://ir.knust.edu.gh/handle/123456789/8196

Eckardt, R., Berger, C., Thiel, C., & Schmullius, C. (2013). Removal of optically thick clouds from multi-spectral satellite images using multi-frequency SAR data. *Remote Sensing*, *5*(6), 2973–3006. https://doi.org/10.3390/rs5062973

EOX. (2016). Understanding Sentinel-2 Satellite Data – EOX. Retrieved January 2, 2017, from https://eox.at/2015/12/understanding-sentinel-2-satellite-data/

ESA. (2015). SENTINEL-2 User Handbook, (1), 64. https://doi.org/GMES-S1OP-EOPG-TN-13-0001

ESA. (2016). Level-1C Cloud Masks - Sentinel-2 MSI Technical Guide - Sentinel Online. Retrieved January 6, 2017, from https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks

Gan, K. K. (2013). Lecture 3: Gaussian Probability Distribution, 7. Retrieved from. *https://www.physics.ohio-state.edu/~gan/teaching/spring04/Chapter3.pdf*

Griggin, M., Burke, H., Mandl, D., & Miller, J. (2003). Cloud cover detection algorithm for EO-1 Hyperion imagery. *IGARSS 2003 2003 IEEE International Geoscience and Remote Sensing Symposium Proceedings IEEE Cat No03CH37477*, *1*(C), 483–494. https://doi.org/10.1109/IGARSS.2003.1293687

GSU. (2016). Blue Sky and Rayleigh Scattering. Retrieved January 6, 2017, from http://hyperphysics.phy-astr.gsu.edu/hbase/atmos/blusky.html

Hadjimitsis, D. G., Papadavid, G., Agapiou, a., Themistocleous, K., Hadjimitsis, M. G., Retalis, a., … Clayton, C. R. I. (2010). Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. *Natural Hazards and Earth System Science*, *10*(1984), 89–95. https://doi.org/10.5194/nhess-10-89-2010

Hagolle, O., Huc, M., Dedieu, G., & Petrucci, B. (2016). MACCS : a Multisensor Atmospheric Correction and Cloud Screening processor High resolution time series with constant viewing angles. Retrieved from. *esaconferencebureau.com/docs/12c04_doc/3-sentinel2_symposium_hagolle.pdf*

Hagolle, O., Huc, M., Pascual, D. V., & Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VEN??S, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment*, *114*(8), 1747–1755. https://doi.org/10.1016/j.rse.2010.03.002

Hagolle, O., Sylvander, S., Huc, M., Claverie, M., Clesse, D., Dechoz, C., … Poulain, V. (2015). SPOT-4 (Take 5): Simulation of sentinel-2 time series on 45 large sites. *Remote Sensing*, *7*(9), 12242–12264. https://doi.org/10.3390/rs70912242

Hamilton. (2010). The Normal or Gaussian Distribution. Retrieved from. http://*www.hamilton.ie/ollie/EE304/Normal.pdf*

HARRIS. (2016). Digital Number, Radiance, and Reflectance. Retrieved from http://www.harrisgeospatial.com/Home/NewsUpdates/TabId/170/ArtMID/735/ArticleID/13592/Digital-Number-Radiance-and-Reflectance.aspx

Harvey. (1989). Evaluating predictive accuracy, *1*(1). Retrieved from. *homepage.univie.ac.at/robert.kunst/prognos7.pdf*

Hocke, K., & Kampfer, N. (2009). and Physics Gap filling and noise reduction of unevenly sampled data by means of the Lomb-Scargle periodogram, (2002), 4197–4206. Retrieved from. https://*www.atmos-chem-phys.net/9/4197/2009/acp-9-4197-2009.pdf*

Hyndman, R. J., & Koehler, A. B. (2005). and Business Statistics Another Look at Measures of Forecast Accuracy Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(November), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., … Koetz, B. (2015). Assessment of an operational system for crop type map production using high temporal and spatial resolution satellite optical imagery. *Remote Sensing*, *7*(9), 12356–12379. https://doi.org/10.3390/rs70912356

Irea. (2016). Optical Remote Sensing. Retrieved from http://www.irea.cnr.it/en/index.php?option=com_k2&view=itemlist&task=category&id=1&Itemid=161

Jojy, C. (2015). A Survey of Cloud Detection Techniques For Satellite Images, 2485–2490. Retrieved from. *https://www.irjet.net/archives/V2/i9/IRJET-V2I9305.pdf*

Jonathan Landy. (2015). Leave-one-out cross-validation – EFavDB. Retrieved January 22, 2017, from http://efavdb.com/leave-one-out-cross-validation/

Jönsson, P., & Eklundh, L. (2002). Seasonality Extraction by Function Fitting to Time-Series of Satellite Sensor Data, *40*(8), 1824–1832. Retrieved from. https://*ieeexplore.ieee.org/abstract/document/1036010/*

Jönsson, P., & Eklundh, L. (2004). TIMESAT - A program for analyzing time-series of satellite sensor data. *Computers and Geosciences*, *30*(8), 833–845. https://doi.org/10.1016/j.cageo.2004.05.006

kaggle. (2016). Mean Absolute Error | Kaggle. Retrieved January 11, 2017, from https://www.kaggle.com/wiki/MeanAbsoluteError

Kandasamy, S., Baret, F., Verger, a., Neveux, P., & Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products. *Biogeosciences*, *10*(6), 4055–4071. https://doi.org/10.5194/bg-10-4055-2013

Kang, S., Running, S. W., Zhao, M., Kimball, J. S., & Glass, J. (2005). Improving continuity of MODIS terrestrial photosynthesis products using an interpolation scheme for cloudy pixels. *International Journal of Remote Sensing*, *26*(February 2016), 1659–1676. https://doi.org/10.1080/01431160512331326693

Kendall, E. A., & Han, W. (2003). Teaching Numerical Analysis Using Elementary Numerical Analysis, Section 4-1. Retrieved from http://homepage.math.uiowa.edu/~atkinson/ftp/ENA_Materials/Overheads/sec_4-1.pdf

Kiyoshi, H. (2014). DN to Reflectance, 1–9. retrievd from https://*www.rsgis.ait.ac.th/~honda/textbooks/advrs/DN2Reflectance_2p.pdf*

Koren, H. (August 2009). cloud detection in MODIS images,. Retrieved from https://*publications.nr.no/5121/Koren_-_Cloud_detection_in_MODIS_images.pdf*

Kusi, K. A. (2013). Assessing the Impacts of Climate Change on the Availability of Stored Water for Irrigation Purposes in the Semi-Arid Areas of Ghana. *MSc Thesis*, 109. Retrieved from https://*ir.knust.edu.gh/bitstream/123456789/5412/1/Kusi_MSc_Final_Thesis.pdf*

Lan, K., & Jorgenson, J. W. (2001). A hybrid of exponential and gaussian functions as a simple model of asymmetric chromatographic peaks, *915*, 1–13.

Levy, D. (2010). Introduction to Numerical Analysis, 121. https://doi.org/10.1137/1024026

Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling and Software*, *53*, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008

Mandanici, E., & Bitelli, G. (2016). Preliminary Comparison of Sentinel-2 and Landsat 8 Imagery for a Combined Use. *Remote Sensing*, *8*(12), 1014. https://doi.org/10.3390/rs8121014

Martins, J. V. (2002). MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters*, *29*(12), 8009. https://doi.org/10.1029/2001GL013252

MathWorks. (2016). Gaussian forward Interpolation formula - File Exchange - MATLAB Central. Retrieved January 10, 2017, from https://www.mathworks.com/matlabcentral/fileexchange/42741-gaussian-forward-interpolation-formula?requestedDomain=www.mathworks.com

Moody, E. G., King, M. D., Platnick, S., Schaaf, C. B., & Gao, F. (2005). Spatially Complete Global Spectral Surface Albedos: Value-Added Datasets Derived From Terra MODIS Land Products. *Ieee Transactions on Geoscience and Remote Sensing*, *43*(1), 144–158. https://doi.org/10.1109/TGRS.2004.838359

Moreno, Á., García-Haro, F. J., Martínez, B., & Gilabert, M. A. (2014). Noise reduction and gap filling of fAPAR time series using an adapted local regression filter. *Remote Sensing*, *6*(9), 8238–8260. https://doi.org/10.3390/rs6098238

Movebank. (2016). Env-DATA interpolation methods | Movebank. Retrieved January 9, 2017, from https://www.movebank.org/node/6400

Nau, R. (2014). Notes on linear regression analysis, (c), 1–18. Retrieved from. https://www.*people.duke.edu/~rnau/notes_on_linear_regression_analysis--robert_nau.pdf*

OSSIM. (2016). orthorectification – OSSIM. Retrieved from https://trac.osgeo.org/ossim/wiki/orthorectification

OTexts. (2017). 8.1 Stationarity and differencing | OTexts. Retrieved February 28, 2017, from https://www.otexts.org/fpp/8/1

Paulbourke. (1999). Interpolation methods. Retrieved January 9, 2017, from http://paulbourke.net/miscellaneous/interpolation/

PennState Astro 801. (2016). Radio Waves to Gamma-rays | Astronomy 801: Planets, Stars, Galaxies, and the Universe. Retrieved from https://www.e-education.psu.edu/astro801/content/l3_p4.html

Piazza, A., Conti, F., Viola, F., Eccel, E., & Noto, L. (2015). Comparative Analysis of Spatial Interpolation Methods in the Mediterranean Area: Application to Temperature in Sicily. *Water*, *7*(5), 1866–1888. https://doi.org/10.3390/w7051866

Piwowar, J. M. (2014). *Remote Sensing Imagery*. Retrieved from.https://doi.org/10.1002/9781118899106.oth1

Revolvy. (2016). Gaussian function. Retrieved February 19, 2017, from https://www.revolvy.com/topic/Gaussian function&uid=1575

Ribeiro, M. (2004). Gaussian probability density functions: Properties and error characterization. *Institute for Systems and Robotics, Technical Report*, (February), 1–30. Retrieved from http://hans.fugal.net/comps/papers/ribeiro_2004.pdf

Robinson, T. P., & Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*, *50*(2), 97–108. https://doi.org/10.1016/j.compag.2005.07.003

Roy, K., Das, R. N., Ambure, P., & Aher, R. B. (2016). *Be aware of error measures. Further studies on validation of predictive QSAR models*. *Chemometrics and Intelligent Laboratory Systems* (Vol. 152). Elsevier B.V. https://doi.org/10.1016/j.chemolab.2016.01.008

Sammut, C., & Webb, G. I. (2010). Encyclopedia of Machine Learning. *Media*, *33*, 439-447–447. https://doi.org/10.1007/978-0-387-30164-8

Satpalda. (2016). Orthorectification of Satellite Imagery | SATPALDA. Retrieved January 18, 2017, from http://www.satpalda.com/service/orthorectification/

Segl, K., Richter, R., Kuster, T., & Kaufmann, H. (2012). End-to-end sensor simulation for spectral band selection and optimization with application to the Sentinel-2 mission. *Appl Opt*, *51*(4), 439–449. https://doi.org/10.1364/AO.51.000439

Seidu, I. (2011). The Effects of Agricultural Modernisation on Poverty Reduction, (August). University of Development studies (Master'sThesis)Retrieved from, https://*www.udsspace.uds.edu.gh/.../the%20effects%20of%20agricultural%20m...*

Spring. (2016). Gaussian Interpolation. Retrieved January 10, 2017, from http://spring.delta-h.de/download/SPRING4_Webhilfe_E/Ber_interpol11.htm

Statistics how to. (2016). RMSE: Root Mean Square Error. Retrieved January 11, 2017, from http://www.statisticshowto.com/rmse/

Surya, S. R., & Simon, P. (2014). Automatic Cloud Removal from Multitemporal Satellite Images. *Journal of the Indian Society of Remote Sensing*, *43*(1), 57–68. https://doi.org/10.1007/s12524-014-0396-2

Tadd Wood. (n.d.). Using Mean Absolute Error to Forecast Accuracy - Contemporary Analysis. Retrieved January 11, 2017, from http://canworksmart.com/using-mean-absolute-error-forecast-accuracy/

Tian, B., Shaikh, M. A., Azimi-Sadjadi, M. R., Vonder Haar, T. H., & Reinke, D. L. (1999). A study of cloud classification with neural networks using spectral and textural features. *IEEE Transactions on Neural Networks*, *10*(1), 138–151. https://doi.org/10.1109/72.737500

Tobergte, D. R., & Curtis, S. (2013). No Title No Title. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Tono irrigation scheme. (2017). Tono irrigation scheme | Ministry of Food & Agriculture. Retrieved January 25, 2017, from http://mofa.gov.gh/site/?page_id=3032

Turks, C. (1990). 3 : radiometric correction of satellite images : when and why radiometric correction is necessary Aim of Lesson The Bilko for Windows image processing software Image data. *Time*, *1*(November), 79–102.

USRA. (2016). Reflectivity in Remote Sensing. Retrieved from. *sti.usra.edu/.../Reflectivity%20in%20Remote%20Sensing%20-%20student%20version...*

VEA irrigation scheme. (2017). VEA IRRIGATION SCHEME | Ministry of Food & Agriculture. Retrieved January 25, 2017, from http://mofa.gov.gh/site/?page_id=3034

Verger, A., Weiss, M., Kandasamy, S., & Vermote, E. (2013). The CACAO Method for Smoothing , Gap Filling , and Characterizing Seasonal Anomalies in Satellite Time Series, *51*(4), 1963–1972.

Vernier. (2016). What are Mean Squared Error and Root Mean Squared Error? > Vernier Software & Technology. Retrieved January 11, 2017, from https://www.vernier.com/til/1014/

Wang, K., Franklin, S. E., Guo, X. L., He, Y. H., & McDermid, G. J. (2009). Problems in remote sensing of landscapes and habitats. *Progress in Physical Geography*, *33*(6), 747–768. https://doi.org/10.1177/0309133309350121

Weiss, D. J., Atkinson, P. M., Bhatt, S., Mappin, B., Hay, S. I., & Gething, P. W. (2014). ISPRS Journal of Photogrammetry and Remote Sensing An effective approach for gap-filling continental scale remotely sensed. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 106–118. https://doi.org/10.1016/j.isprsjprs.2014.10.001

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

Yin, G., Li, J., Liu, Q., Zhong, B., & Li, A. (2016). Improving LAI spatio-temporal continuity using a combination of MODIS and MERSI data. *Remote Sensing Letters*, *7*(8), 771–780. https://doi.org/10.1080/2150704X.2016.1182657

Zhang, C., Li, W., & Travis, D. (2007). Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach. *International Journal of Remote Sensing*, *28*(22), 5103–5122. https://doi.org/10.1080/01431160701250416

Zhang, C., Li, W., & Travis, D. J. (2009). Restoration of clouded pixels in multispectral remotely sensed imagery with cokriging. *International Journal of Remote Sensing*, *30*(9), 2173–2195. https://doi.org/10.1080/01431160802549294

Zhang, Y., Rossow, W. B., Lacis, A. A., Oinas, V., & Mishchenko, M. I. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *Journal of Geophysical Research*, *109*(D19), D19105. https://doi.org/10.1029/2003JD004457

Zhu, X., Liu, D., & Chen, J. (2012). A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment*, *124*, 49–60. https://doi.org/10.1016/j.rse.2012.04.019

Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269–277. https://doi.org/10.1016/j.rse.2014.12.014

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

8. APPENDICES

Appendix I: Time series images downloaded and used in study

Numb	Identification	Date
1	L1C_PDMC_20160103T225936	03/01/2016
2	L1C_PDMC_20160113T201047	13/01/2016
3	L1C_PDMC_20160127T145447	23/01/2016
4	L1C_PDMC_20160202T201055	02/02/2016
5	L1C_PDMC_20160212T211458	12/02/2016
6	L1C_PDMC_20160222T221621	22/02/2016
7	L1C_PDMC_20160305T122702	03/03/2016
8	L1C_PDMC_20160313T213611	13/03/2016
9	L1C_PDMC_20160323T215208	23/03/2016
10	L1C_PDMC_20160403T125043	02/04/2016
11	L1C_PDMC_20160413T151100	12/04/2016
12	L1C_PDMC_20160422T195342	22/04/2016
13	L1C_PDMC_20160502T213926	02/05/2016
14	L1C_PDMC_20160513T172247	12/05/2016
15	L1C_PDMC_20160522T182359	22/05/2016
16	L1C_PDMC_20160601T211123	01/06/2016
17	L1C_PDMC_20160611T211239	11/06/2016
18	L1C_PDMC_20160621T205324	21/06/2016
19	L1C_PDMC_20160702T064003	02/07/2016
20	L1C_PDMC_20160711T222818	12/07/2016
21	L1C_PDMC_20160722T165539	22/07/2016
22	L1C_PDMC_20160802T042744	31/07/2016
23	L1C_PDMC_20160811T221914	10/08/2016
24	L1C_PDMC_20160821T164814	20/08/2016
25	L1C_PDMC_20160831T214125	30/08/2016
26	L1C_PDMC_20160912T191535	09/09/2016
27	L1C_PDMC_20161005T183225	19/09/2016
28	L1C_PDMC_20160929T194758	29/09/2016
29	L1C_PDMC_20161010T011211	09/10/2016
30	L1C_PDMC_20161118T121843	19/10/2016
31	L1C_PDMC_20161029T204506	29/10/2016
32	L1C_PDMC_20161109T233920	08/11/2016
33	L1C_PDMC_20161118T210906	18/11/2016
34	L1C_PDMC_20161129T002019	28/11/2016

Appendix II: Granule Metadata developed for image for granule

```
<n1:Level-1C_User_Product xmlns:n1="https://psd-13.sentinel2.eo.esa.int/PSD/User_Product_Level-1C.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="https://psd-
13.sentinel2.eo.esa.int/PSD/User_Product_Level-1C.xsd">
    <n1:General_Info>
        <Product_Info>
            <PRODUCT_START_TIME>2016-02-12T10:28:30.356Z</PRODUCT_START_TIME>
            <PRODUCT_STOP_TIME>2016-02-12T10:28:30.356Z</PRODUCT_STOP_TIME>
            <PRODUCT_URI>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_N02.01</PRODUCT_URI>
            <PROCESSING_LEVEL>Level-1C</PROCESSING_LEVEL>
            <PRODUCT_TYPE>S2MSI1C</PRODUCT_TYPE>
            <PROCESSING_BASELINE>02.01</PROCESSING_BASELINE>
            <GENERATION_TIME>2016-02-12T21:14:58.000732Z</GENERATION_TIME>
<PREVIEW_IMAGE_URL>https://pdmcdam2.sentinel2.eo.esa.int/s2pdgs_geoserver/geo_service.php?service=WMS&amp;version=1.1
.0&amp;request=GetMap&amp;layers=S2A_A003351_N0201:S2A_A003351_N0201&amp;styles=&amp;bbox=-
2.085367790894454,10.744423290630301,2.174994947309441,13.56705921472256&amp;width=1580&amp;height=1047&amp;srs=EPSG:
4326&amp;format=image/png</PREVIEW_IMAGE_URL>
            <PREVIEW_GEO_INFO>BrowseImageFootprint</PREVIEW_GEO_INFO>
            <Datatake datatakeIdentifier="GS2A_20160212T102132_003351_N02.01">
      <SPACECRAFT_NAME>Sentinel-2A</SPACECRAFT_NAME>
      <DATATAKE_TYPE>INS-NOBS</DATATAKE_TYPE>
      <DATATAKE_SENSING_START>2016-02-12T10:21:32.031Z</DATATAKE_SENSING_START>
      <SENSING_ORBIT_NUMBER>65</SENSING_ORBIT_NUMBER>
      <SENSING_ORBIT_DIRECTION>DESCENDING</SENSING_ORBIT_DIRECTION>
    </Datatake>
<Query_Options>
<Area_Of_Interest>
                    <Bbox>
                        <LOWER_CORNER>11.251 -180.0000</LOWER_CORNER>
                        <UPPER_CORNER>13.051 180.0000</UPPER_CORNER>
                    </Bbox>
                </Area_Of_Interest>
<FULL_SWATH_DATATAKE>false</FULL_SWATH_DATATAKE>
<PREVIEW_IMAGE>true</PREVIEW_IMAGE>
<Band_List>
<BAND_NAME>B1</BAND_NAME>
<BAND_NAME>B2</BAND_NAME>
<BAND_NAME>B3</BAND_NAME>
<BAND_NAME>B4</BAND_NAME>
<BAND_NAME>B5</BAND_NAME>
<BAND_NAME>B6</BAND_NAME>
<BAND_NAME>B7</BAND_NAME>
<BAND_NAME>B8</BAND_NAME>
<BAND_NAME>B9</BAND_NAME>
<BAND_NAME>B10</BAND_NAME>
<BAND_NAME>B11</BAND_NAME>
<BAND_NAME>B12</BAND_NAME>
<BAND_NAME>B8A</BAND_NAME>
</Band_List>
<METADATA_LEVEL>Expertise</METADATA_LEVEL>
<Aux_List productLevel="Level-1C">
<aux>
                        <GIPP>NO</GIPP>
                    </aux>
<aux>
                        <IERS>NO</IERS>
                    </aux>
</Aux_List>
<PRODUCT_FORMAT>SAFE</PRODUCT_FORMAT>
<AGGREGATION_FLAG>false</AGGREGATION_FLAG>
</Query_Options>
<Product_Organisation>
                <Granule_List>
                    <Granules datastripIdentifier="S2A_OPER_MSI_L1C_DS_SGS__20160212T174440_S20160212T102830_N02.01"
granuleIdentifier="S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_N02.01" imageFormat="JPEG2000">
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B8A</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B11</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B10</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B03</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B04</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B05</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B12</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B09</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B07</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B06</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B08</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B01</IMAGE_ID>
                        <IMAGE_ID>S2A_OPER_MSI_L1C_TL_SGS__20160212T174440_A003351_T30PYT_B02</IMAGE_ID>
                    </Granules>
                </Granule_List>
</Product_Organisation>
        </Product_Info>
```

```
        <Product_Image_Characteristics>
            <Special_Values>
                <SPECIAL_VALUE_TEXT>NODATA</SPECIAL_VALUE_TEXT>
                <SPECIAL_VALUE_INDEX>0</SPECIAL_VALUE_INDEX>
            </Special_Values>
            <Special_Values>
                <SPECIAL_VALUE_TEXT>SATURATED</SPECIAL_VALUE_TEXT>
                <SPECIAL_VALUE_INDEX>65535</SPECIAL_VALUE_INDEX>
            </Special_Values>
            <Image_Display_Order>
                <RED_CHANNEL>3</RED_CHANNEL>
                <GREEN_CHANNEL>2</GREEN_CHANNEL>
                <BLUE_CHANNEL>1</BLUE_CHANNEL>
            </Image_Display_Order>
            <QUANTIFICATION_VALUE unit="none">10000</QUANTIFICATION_VALUE>
            <Reflectance_Conversion>
        <U>1.02780302783143</U>
        <Solar_Irradiance_List>
          <SOLAR_IRRADIANCE bandId="0" unit="W/m²/µm">1913.57</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="1" unit="W/m²/µm">1941.63</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="2" unit="W/m²/µm">1822.61</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="3" unit="W/m²/µm">1512.79</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="4" unit="W/m²/µm">1425.56</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="5" unit="W/m²/µm">1288.32</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="6" unit="W/m²/µm">1163.19</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="7" unit="W/m²/µm">1036.39</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="8" unit="W/m²/µm">955.19</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="9" unit="W/m²/µm">813.04</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="10" unit="W/m²/µm">367.15</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="11" unit="W/m²/µm">245.59</SOLAR_IRRADIANCE>
          <SOLAR_IRRADIANCE bandId="12" unit="W/m²/µm">85.25</SOLAR_IRRADIANCE>
        </Solar_Irradiance_List>
      </Reflectance_Conversion>
            <PHYSICAL_GAINS bandId="0">4.06483122812114</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="1">3.80370314635676</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="2">4.19666838204671</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="3">4.52398456222777</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="4">5.21306949428705</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="5">4.88125648438436</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="6">4.54889907882714</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="7">6.21210795022227</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="8">5.15037808260271</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="9">8.55863365550801</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="10">55.413020614716</PHYSICAL_GAINS>
            <PHYSICAL_GAINS bandId="11">35.4481358454542</PHYSICAL_GAINS>
            <REFERENCE_BAND>0</REFERENCE_BAND>
        </Product_Image_Characteristics>
    </n1:General_Info>
    <n1:Geometric_Info>
        <Product_Footprint>
            <Product_Footprint>
                <Global_Footprint>
                    <EXT_POS_LIST> 11.7585678866461 -2.082226174230733 11.714106179564578 -2.0823668565966535
11.670133195757675 -2.082517993526374 11.67013313641358 -2.082505992769178 10.765681259024202 -2.085367790894454
10.761301638429813 -1.1263473024653257 10.754437756847084 -0.2125546823238774 10.744423290630301 0.745365438051929
10.758432211228888 0.7455468295161269 10.761828635748657 1.2156644283862414 10.76593259155724 2.174994947309441
11.670406143404394 2.1724135019557718 11.670406193549555 2.172424327119259 11.71438022589423 2.1722879961400476
11.758843009943073 2.172161095528018 11.758842958862113 2.1721501500613343 12.574279797627552 2.1696220830002413
12.574279852763716 2.169633143046981 12.61850257875225 2.1694849808348784 12.663253352763395 2.169346241753741
12.663253296580942 2.169335049658126 13.56705921472256 2.166306971195076 13.562094950288943 1.2420785294434757
13.562262153833206 1.2420773483626609 13.561852198871556 1.1968839659339254 13.561609653141844 1.151727696596643
13.561442591822127 1.1517289373037412 13.553063730961096 0.228044389895723 13.547075212573146 0.2281110769465797
13.551882768943523 -0.1381343006559648 13.55204944786505 -0.13813238439255 13.552474906645363 -0.1832440569428614
13.553067829846247 -0.228413653483674 13.552900929246757 -0.2284155118109735 13.560755110486602 -1.0611994820514379
13.560923067766707 -1.061198173615988 13.561183558380115 -1.1066280922281617 13.56161238835793 -1.152097215052814
13.561444280943768 -1.1520984632793774 13.566740327869518 -2.075736277534433 12.662956348688711 -2.0790932123168933
12.662956415177522 -2.079105619833539 12.618206732855702 -2.0792594263075728 12.573985067652204 -2.079423679325417
12.573985002401576 -2.079411418189613 11.758567826194573 -2.082214040112981 11.7585678866461 -
2.082226174230733</EXT_POS_LIST>
                </Global_Footprint>
            </Product_Footprint>
            <RASTER_CS_TYPE>POINT</RASTER_CS_TYPE>
            <PIXEL_ORIGIN>1</PIXEL_ORIGIN>
        </Product_Footprint>
        <Coordinate_Reference_System>
            <GEO_TABLES version="1">EPSG</GEO_TABLES>
            <HORIZONTAL_CS_TYPE>GEOGRAPHIC</HORIZONTAL_CS_TYPE>
        </Coordinate_Reference_System>
    </n1:Geometric_Info>
    <n1:Auxiliary_Data_Info>
        <GIPP_List>
      <GIPP_FILENAME type="GIP_BLINDP"
version="0003">S2A_OPER_GIP_BLINDP_MPC__20150605T094736_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
```

```
      <GIPP_FILENAME type="GIP_CLOINV"
version="0004">S2A_OPER_GIP_CLOINV_MPC__20151021T225159_V20150701T225159_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_DATATI"
version="0007">S2A_OPER_GIP_DATATI_MPC__20151117T131048_V20150703T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_DECOMP"
version="0000">S2A_OPER_GIP_DECOMP_MPC__20121031T075922_V19830101T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_EARMOD"
version="0001">S2__OPER_GIP_EARMOD_MPC__20150605T094736_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_ECMWFP"
version="0002">S2A_OPER_GIP_ECMWFP_MPC__20121031T075922_V19830101T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_G2PARA"
version="0001">S2A_OPER_GIP_G2PARA_MPC__20150605T094736_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_G2PARE"
version="0002">S2A_OPER_GIP_G2PARE_MPC__20150605T094736_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_GEOPAR"
version="0003">S2A_OPER_GIP_GEOPAR_MPC__20150605T094741_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_INTDET"
version="0003">S2A_OPER_GIP_INTDET_MPC__20150731T092205_V20150703T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_JP2KPA"
version="0004">S2A_OPER_GIP_JP2KPA_MPC__20151126T214920_V20150701T214920_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_MASPAR"
version="0007">S2A_OPER_GIP_MASPAR_MPC__20160203T150846_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_OLQCPA"
version="0000">S2A_OPER_GIP_OLQCPA_MPC__20151127T150000_V20000506T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_PRDLOC"
version="0008">S2A_OPER_GIP_PRDLOC_MPC__20151125T152205_V20150703T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_PROBAS"
version="999">S2__OPER_GIP_PROBAS_MPC__20160121T111510_V20151215T111510_20200101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2ABCA"
version="0088">S2A_OPER_GIP_R2ABCA_MPC__20160203T150845_V20160128T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2BINN"
version="0003">S2A_OPER_GIP_R2BINN_MPC__20150605T094803_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2CRCO"
version="0003">S2A_OPER_GIP_R2CRCO_MPC__20151023T224715_V20150622T224715_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094741_V20150622T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094742_V20150622T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094742_V20150622T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094742_V20150622T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DECT"
version="0003">S2A_OPER_GIP_R2DECT_MPC__20150605T094742_V20150622T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B10</GIPP_FILENAME>
```

```
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEFI"
version="0003">S2A_OPER_GIP_R2DEFI_MPC__20150605T094741_V20150622T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094741_V20150622T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DENT"
version="0003">S2A_OPER_GIP_R2DENT_MPC__20150605T094742_V20150622T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2DEPI"
version="0016">S2A_OPER_GIP_R2DEPI_MPC__20160203T150846_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EOB2"
version="0009">S2A_OPER_GIP_R2EOB2_MPC__20150827T165847_V20150819T000000_99999999T999999_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150839_V20160128T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150839_V20160128T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150840_V20160128T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150841_V20160128T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150842_V20160128T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150842_V20160128T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150843_V20160128T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150843_V20160128T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150844_V20160128T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150844_V20160128T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150844_V20160128T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150845_V20160128T000000_21000101T000000_B11</GIPP_FILENAME>
```

```
      <GIPP_FILENAME type="GIP_R2EQOG"
version="0054">S2A_OPER_GIP_R2EQOG_MPC__20160203T150845_V20160128T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094741_V20150622T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2L2NC"
version="0003">S2A_OPER_GIP_R2L2NC_MPC__20150605T094742_V20150622T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2NOMO"
version="0004">S2A_OPER_GIP_R2NOMO_MPC__20150605T094803_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2PARA"
version="0003">S2A_OPER_GIP_R2PARA_MPC__20151023T224307_V20150622T224307_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2SWIR"
version="0017">S2A_OPER_GIP_R2SWIR_MPC__20160203T150845_V20151013T134305_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B08</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_R2WAFI"
version="0003">S2A_OPER_GIP_R2WAFI_MPC__20150605T094742_V20150622T000000_21000101T000000_B12</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_RESPAR"
version="0001">S2A_OPER_GIP_RESPAR_MPC__20150605T094736_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_SPAMOD"
version="0011">S2A_OPER_GIP_SPAMOD_MPC__20151117T131048_V20150901T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_TILPAR"
version="0007">S2A_OPER_GIP_TILPAR_MPC__20151209T095117_V20150622T000000_21000101T000000_B00</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131048_V20150703T000000_21000101T000000_B01</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131049_V20150703T000000_21000101T000000_B02</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131049_V20150703T000000_21000101T000000_B03</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131049_V20150703T000000_21000101T000000_B04</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B05</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B06</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B07</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B08</GIPP_FILENAME>
```

```
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B8A</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B09</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B10</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131050_V20150703T000000_21000101T000000_B11</GIPP_FILENAME>
      <GIPP_FILENAME type="GIP_VIEDIR"
version="0005">S2A_OPER_GIP_VIEDIR_MPC__20151117T131051_V20150703T000000_21000101T000000_B12</GIPP_FILENAME>
</GIPP_List>
<PRODUCTION_DEM_TYPE>S2__OPER_DEM_GLOBEF_PDMC_19800101T000000_S19800101T000000</PRODUCTION_DEM_TYPE
<IERS_BULLETIN_FILENAME>S2__OPER_AUX_UT1UTC_PDMC_20160211T000000_V20160212T000000_20170211T000000</IERS_BULLETIN_FILE
NAME> <GRI_FILENAME>S2A_OPER_AUX_GRI065_PDMC_20130621T120000_S20130101T000000</GRI_FILENAME>
<ECMWF_DATA_REF>S2__OPER_AUX_ECMWFD_FAKE_19800101T000000_V19800101T000000_19800101T000000</ECMWF_DATA_REF>
 </n1:Auxiliary_Data_Info>
    <n1:Quality_Indicators_Info>
        <Cloud_Coverage_Assessment>0.06650666666666667</Cloud_Coverage_Assessment>
        <Technical_Quality_Assessment>
            <DEGRADED_ANC_DATA_PERCENTAGE>0</DEGRADED_ANC_DATA_PERCENTAGE>
            <DEGRADED_MSI_DATA_PERCENTAGE>0</DEGRADED_MSI_DATA_PERCENTAGE>
        </Technical_Quality_Assessment>
        <Quality_Control_Checks>
            <Quality_Inspections>
                <SENSOR_QUALITY_FLAG>PASSED</SENSOR_QUALITY_FLAG>
                <GEOMETRIC_QUALITY_FLAG>PASSED</GEOMETRIC_QUALITY_FLAG>
                <GENERAL_QUALITY_FLAG>PASSED</GENERAL_QUALITY_FLAG>
                <FORMAT_CORRECTNESS_FLAG>PASSED</FORMAT_CORRECTNESS_FLAG>
                <RADIOMETRIC_QUALITY_FLAG>PASSED</RADIOMETRIC_QUALITY_FLAG>
            </Quality_Inspections>
            <Failed_Inspections/>
        </Quality_Control_Checks>
    </n1:Quality_Indicators_Info></n1:Level-1C_User_Product>
```

Appendix III: Image values extracted from dynamic and static land covers

Farm	3-Jan	13-Jan	23-Jan	2-Feb	12-Feb	22-Feb	3-Mar	13-Mar	23-Mar	2-Apr	12-Apr	22-Apr	2-May	12-May	22-May
Cloud free	NaN	0.139	0.188	NaN	0.174	NaN	0.147	NaN	NaN	NaN	0.123	0.134	NaN	NaN	NaN
Farm 1															
Cloud free	NaN	0.15	0.152	NaN	0.185	NaN	0.165	NaN	NaN	NaN	0.152	0.156	NaN	NaN	NaN
Farm 2															
Cloud free	NaN	0.15	0.153	NaN	0.183	NaN	0.163	NaN	NaN	NaN	0.152	0.156	NaN	NaN	NaN
Farm 3															
Cloud free	NaN	0.138	0.144	NaN	0.174	NaN	0.156	NaN	NaN	NaN	0.152	0.149	NaN	NaN	NaN
Farm 4															
Cloud free	NaN	0.156	0.159	NaN	0.188	NaN	0.168	NaN	NaN	NaN	0.159	0.162	NaN	NaN	NaN
Farm 5															
Cloud free	NaN	0.137	0.168	NaN	0.171	NaN	0.148	NaN	NaN	0.145	0.133	0.137	NaN	NaN	NaN
Farm 6															
Cloud free	NaN	0.162	0.172	NaN	0.193	NaN	0.173	NaN	NaN	NaN	0.164	0.162	NaN	NaN	NaN
Farm 7															
Cloud free	NaN	0.147	0.179	NaN	0.185	NaN	0.166	NaN	NaN	0.147	0.149	0.152	NaN	NaN	NaN
Farm 8															
Cloud free	NaN	0.154	0.18	NaN	0.185	NaN	0.164	NaN	NaN	0.154	0.162	0.163	NaN	NaN	NaN
Farm 9															
Cloud free	NaN	0.14	NaN	NaN	0.172	NaN	0.153	NaN	NaN	0.158	0.14	0.145	NaN	NaN	NaN
Farm 10															
Cloud free	NaN	0.165	0.126	NaN	0.188	NaN	0.245	NaN	NaN	NaN	0.473	0.315	NaN	NaN	NaN
Farm 11															
Cloud free	NaN	0.141	0.123	NaN	0.116	NaN	0.115	NaN	NaN	NaN	0.128	0.133	NaN	NaN	NaN
Farm 12															
Cloud free	NaN	0.112	0.114	NaN	0.173	NaN	0.997	NaN	NaN	NaN	0.115	0.119	NaN	NaN	NaN
Farm 13															
Cloud free	NaN	0.282	0.246	NaN	0.182	NaN	0.152	NaN	NaN	NaN	0.146	0.159	NaN	NaN	NaN
Farm 14															
Cloud free	NaN	0.124	0.113	NaN	0.115	NaN	0.143	NaN	NaN	NaN	0.78	0.858	NaN	NaN	NaN
Farm 15															
Cloud free	NaN	0.14	0.12	NaN	0.138	NaN	0.148	NaN	NaN	0.264	0.272	0.234	NaN	NaN	NaN
Built-up															
Cloud free	NaN	0.132	0.126	NaN	0.127	NaN	0.122	NaN	NaN	NaN	0.118	0.117	NaN	NaN	NaN
Water body															

Appendix IV: Summary of the information collected from the farms

Farmer's Name	Village	Y	X	Crops 2016	Planting Date	Harvesting	Cropping Pattern	Farm Size (Ha)
Adibora Taanke	Yiguania	10.8794	-1.1117	Peper	18-Oct	February to April 2017	Mono	4
		10.8811	-1.1119	Maize	June	September	Mono	7
		10.8799	-1.1114					
		10.8799	-1.1108					
		10.8792	-1.1107					
		10.8793	-1.1114					
		10.8797	-1.1114					
Kaba Averitica	Wuru	10.8869	-1.1281	Rice	July	November	Mono	2
		10.8871	-1.128					
		10.8873	-1.1278					
		10.8876	-1.1279					
		10.8876	-1.1259					
		10.8875	-1.1286					
Bakwam Bafuga	Gia	10.8893	-1.1475	Peper	August	November to February	Mono	2
		10.8891	-1.1474					
		10.8889	-1.1472					
		10.8892	-1.1469					
		10.8896	-1.1471					
Thomas Kusola	Gia	10.8893	-1.1476	Peper	September	December to February	Mono	1
		10.889	-1.148					
		10.8897	-1.1484					
		10.8891	-1.1479					
		10.8894	-1.1481					
Batoro Atawoze	Gia	10.8889	-1.1479	Maize	September	December	Mono	2
		10.8886	-1.1487					2
		10.8883	-1.1486					
		10.8885	-1.1481					
		10.8887	-1.1478					
Kwamovae Bayagebopan	Gia	10.8882	-1.1485	Peper	19-Jul	October to March	Mono	2
		10.888	-1.1488					
		10.8879	-1.1494					
		10.8885	-1.1487					
		10.8882	-1.1485					

Appendix V: Python Script for nearest neighbour Interpolation

```
from __future__ import print_function
from collections import namedtuple
import os

input = raw_input
ImageInfo = namedtuple("ImageInfo", "path, month, day, yday, band")

def enumerate_images(path):
    # enumerate subdirectories
    for dir in (d for d in (os.path.join(path, d1) for d1 in os.listdir(path)) \
                 if os.path.isdir(d)):
        for imageFile in enumerate_images(dir):
            yield imageFile

    # enumerate *.img files
    for img in (f for f in (os.path.join(path, f1) for f1 in os.listdir(path) \
                if f1.endswith(".img")) if os.path.isfile(f)):
        yield img

def get_image_info(path):
    from datetime import datetime
    parts = str(path).split("\\")[-2:]
    month = parts[0].split("_")[0]
    day = parts[0].split("_")[1]
    dayofyr = datetime(2015, month, day).timetuple().tm_yday
    band = parts[1].split("_")[1].split(".")[0]

    return ImageInfo(path=path, month=month, day=day, yday=dayofyr, band=band)

def nn_fill(imagelist):
    """Fills the images in the list with the nearest neighbour algorithm."""
    import arcpy
    if arcpy.CheckExtension("Spatial") == "Available":
        arcpy.CheckOutExtension("Spatial")
        arcpy.env.overwriteOutput = True
    else:
        print("You do not have the Spatial Analyst Extension\n" +
              "The images will not be processed")
        return

    # internal function to perform the nearest neighbour fill
    def do_nn(raster, fillwith):
        """
        Fills raster with the image loaded from the path
        specified by the fillwith argument.
        """
        if len(fillwith) == 0:
            return raster
        else:
            fillPath = str(fillwith[0])
            fillraster = arcpy.Raster(fillPath)
            print("        with " + "\\".join(fillPath.split("\\")[-2:]), end="")
            result = arcpy.sa.Con(raster == 0, fillraster, raster)
            print(" | Min Value: " + str(result.minimum))
            return result


    for i in range(len(imagelist)):
        r = arcpy.Raster(imagelist[i])
        print("Filling " + imagelist[i])
        step = 1

        while r.minimum == 0:
            older = list(reversed(imagelist[:max(0, i-step+1)]))
            newer = imagelist[i+step:]
            r = do_nn(r, older) # use older images to remove clouds

            if r.minimum == 0:  # retry with newer images if there are clouds
                r = do_nn(r, newer)
```

```
            if r.minimum == 0:  # still has clouds, go to next set
                step += 1

                if step >= len(imagelist):  # exceeded images in set
                    break   # stop iterations and save the image


        # save the final image
        complete = r.minimum != 0
        outputfile = str(imagelist[i]).replace(".img", ("_NN_FF" if complete \
                                                        else "_NN_IF")+".img")
        r.save(outputfile)
        print("Output saved to: " + outputfile + "\n")

def main():
    imgDir = input("Enter the base directory of the satellite images:\n").strip("\"")

    imagePaths = list(enumerate_images(imgDir))
    bandNames = {"B": "BLUE", "G": "GREEN", "R": "RED", "N": "NIR"}

    for band in list("BGRN"):
        print("Processing " + bandNames[band] + " images . . .")
        nn_fill([path for path in imagePaths if str(path).upper().endswith( \
                "_"+ band + ".IMG")])

    pass

# start the main method if the script was started directly
if __name__ == "__main__":
    main()
    input("Press any key to exit . . .")
```

Appendix VI: Python Script for Linear interpolation

```
from __future__ import print_function
from collections import namedtuple
import os
import sys

input = raw_input
ImageInfo = namedtuple("ImageInfo", "path, month, day, yday, band")

def all_combos(seq, min_length=None, max_length=None, shorter_first=False):
    """Returns all combinations of elements in the specified sequence."""
    from itertools import combinations

    min_length = 0 if min_length == None else min_length
    max_length = len(seq) if max_length == None else max_length

    lengths = range(min_length, max_length+1)
    if not shorter_first:
        lengths = reversed(lengths)

    for length in lengths:
        for subset in combinations(seq, length):
            yield subset

def enumerate_images(path):
    # enumerate subdirectories
    for dir in (d for d in (os.path.join(path, d1) for d1 in os.listdir(path)) \
                 if os.path.isdir(d)):
        for imageFile in enumerate_images(dir):
            yield imageFile

    # enumerate *.img files
    for img in (f for f in (os.path.join(path, f1) for f1 in os.listdir(path) \
                            if f1.endswith(".img"))\
                            if os.path.isfile(f)):
        yield img

def get_image_info(path):
    from datetime import datetime
    name = os.path.basename(path)
    month = int(name.split("_")[0])
    day = int(name.split("_")[1])
    dayofyr = datetime(2016, month, day).timetuple().tm_yday
    band = os.path.splitext(name)[0][-1]

    return ImageInfo(path=path, month=month, day=day, yday=dayofyr, band=band)

def li_fill(imagelist, daystouse, scratch=None, extent=None):
    """Fills the images in the list with the linear interpolation algorithm."""
    import arcpy
    from math import ceil, floor
    if arcpy.CheckExtension("Spatial") == "Available":
        arcpy.CheckOutExtension("Spatial")
        arcpy.env.overwriteOutput = True

        # set scratch workspace; allows parallel runs
        import tempfile
        if scratch == None or not os.path.exists(scratch):
            scratch = tempfile.gettempdir()

        arcpy.env.scratchWorkspace = tempfile.mkdtemp(prefix="li_fill_", dir=scratch)
        print("Scratch Workspace: " + arcpy.env.scratchWorkspace)

        # set the procesing extent
        if extent != None:
            arcpy.env.extent = extent
    else:
        print("You do not have the Spatial Analyst Extension\n" +
```

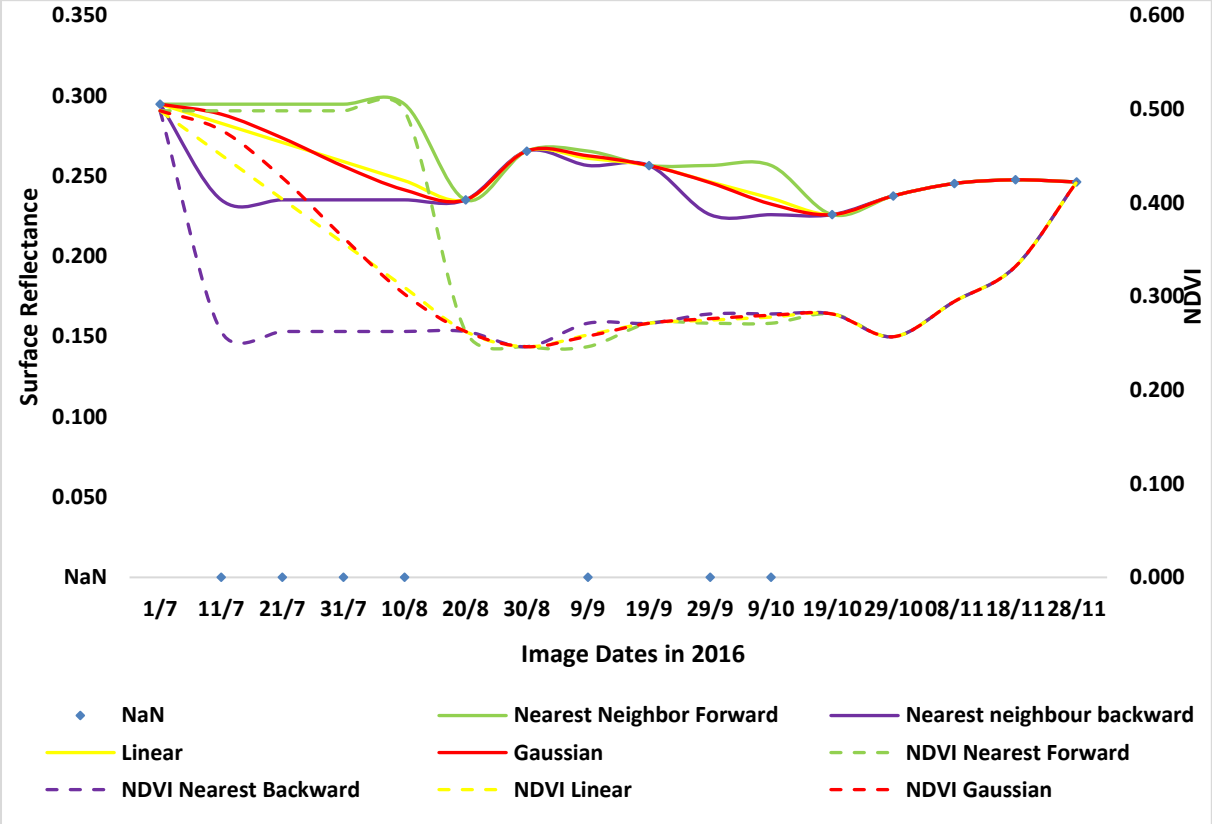
```
          "The images will not be processed")
    return

# internal function to perform the linear interpolation fill
def do_li(raster, yday, fillimagepaths):
    """Fills raster with the images loaded from the paths specified."""
    n = len(fillimagepaths)
    print("Interpolating with " + str(n) + " image(s)")
    print("\n".join("  > " + os.path.basename(p) for p in fillimagepaths))

    xvalues = [get_image_info(img).yday for img in fillimagepaths]
    yvalues = [arcpy.Raster(path) for path in fillimagepaths]

    if any(r.maximum == 0 for r in yvalues):
        # skip processing if any fill image  is completely covered with clouds
        print(" >> No valid pixels were found in the input images to process.")
        return raster
    else:
        # fill the pixels where a value exists in each fill image
        fillarea = yvalues[0] > 0
        for i in range(1, len(yvalues)):
            fillarea &= yvalues[i]

        # optimization: precalculate intermediate rasters
        ysum = sum_rasters(yvalues)
        yxsum = sum_rasters([xvalues[x] * yvalues[x] for x in range(n)])
        denom = ((n * sum(x**2 for x in xvalues)) - sum(xvalues)**2)

        # calculate a and b values for use in the equation
        a = ((ysum * sum(x**2 for x in xvalues)) - (sum(xvalues) * yxsum)) / denom
        b = ((n * yxsum) - (sum(xvalues) * ysum)) / denom

        # fill empty pixels where all fill images have values
        result = arcpy.sa.Con((raster <= 0) & fillarea, a + (b * yday), raster)
        if result.minimum < 0 or result.maximum > 1:
            print("    Replacing out-of-range values with zero.")
            result = arcpy.sa.Con((raster > 1) | (raster < 0), 0, raster)
        print(" >> Min Value: " + str(result.minimum) + " Max Value: " + str(result.maximum))
        return result


backdays = int(ceil(daystouse / 2.0))
fwddays = int(floor(daystouse/ 2.0))
log = Logger()

#for i in range(len(imagelist)):
for i in range(backdays, len(imagelist)-fwddays):
    yday = get_image_info(imagelist[i]).yday
    # skip images after March 31
    if yday > (31+29+31):
        continue

    log.attach()    # prepare to log outputs
    extra = 0
    print("Filling " + imagelist[i] + " with up to " + str(daystouse+extra) + " images.")

    # check for and skip generation
    prevoutput = [output for output in \
                  (imagelist[i].replace(".img", "_LI_"+rslt+".img") for rslt in ("FF", "IC"))
                  if os.path.exists(output)]
    if len(prevoutput) > 0:
        log.detatch()
        log.clearcache()
        prev = arcpy.Raster(prevoutput[0])
        print("\n".join(("    The image has already been processed.",
                         " >> Previous Result:",
                         "    File name: " + prevoutput[0],
                         "    Min value: " + str(prev.minimum),
                         "    Max Value: " + str(prev.maximum),
                         "")
                        ))

        continue    # skip generation
```

```
        r = arcpy.Raster(imagelist[i])
        usedcombinations = []   # used to avoid repeating combinations when extra images are added


        while r.minimum <= 0 or r.maximum >= 1:     # image has clouds; do linear interpolation
            older = imagelist[max(0, i-(backdays+extra)):i]
            newer = imagelist[i+1:i+(fwddays+extra)+1]

            # iterate through all combinations in the fill images
            # which have not been evaluated yet and
            # all rasters have max pixel value greater than 0 (no clouds)
            for imageset in (ss for ss in all_combos(older + newer, min_length=2) \
                            if not ss in usedcombinations and \
                            all(arcpy.Raster(r1).maximum > 0 for r1 in ss)):

                r = do_li(r, yday, imageset)
                usedcombinations.append(imageset)
                if r.minimum > 0:
                    break

            if r.minimum <= 0 or r.maximum >= 1:    # increase the number of extra images used
                extra += 1
                print("Failed to fill all pixels.\n" + \
                      "Increasing maximum number of images to " + str(daystouse + (extra * 2)))

            # stop when images in set have been exceeded
            if (backdays + extra >= len(imagelist)) or \
               (fwddays + extra >= len(imagelist)) or \
               (daystouse + (extra * 2) > 10):
                print("Exceeded limits of images to use for filling.")
                break

        # save the final image
        complete = r.minimum > 0 and r.maximum <= 1
        outputfile = str(imagelist[i]).replace(".img", ("_LI_FF" if complete else "_LI_IC")+".img")
        r.save(outputfile)
        print("Output saved to: " + outputfile + "\n")
        log.detatch()
        with open(outputfile.replace(".img", ".log"), 'w') as logfile:
            logfile.write(''.join(log.cache))
            log.clearcache()

def sum_rasters(imagelist):
    """Returns a raster that is the sum of all coincident pixels in the specified raster sequence."""
    if len(imagelist) == 0:
        return None

    result = imagelist[0]
    for i in range(1, len(imagelist)):
        result = result + imagelist[i]

    return result


def main():
    imgDir = input("Enter the base directory of the satellite images:\n").strip("\"")
    numDays = int(input("What is the maximum number of images to use for interpolation? "))
    inbands = input("Enter the bands to process [BGRN]: ").strip().upper()
    extent = input("Enter the extent to process: ").strip()

    scratch = input("\nEnter the scratch workspace directory " + \
                    "(Leave blank to use the default):\n").strip("\"")

    imagePaths = list(enumerate_images(imgDir))
    bandNames = {"B": "BLUE", "G": "GREEN", "R": "RED", "N": "NIR"}

    for band in (b for b in list(inbands) if b in list("BGRN")):
        print("Processing " + bandNames[band] + " images . . .")
        li_fill([path for path in imagePaths if str(path).upper().endswith("_"+ band + \
                 ".IMG")], numDays, scratch, extent)

class Logger(object):
```

```
    def __init__(self, filename=None):
        import sys
        self.terminal = sys.stdout
        self.cache = []
        self.log = None
        self.writetolog = False
        self.logtofile(filename)

    def attach(self):
        import sys
        sys.stdout = self

    def clearcache(self):
        self.cache = []

    def detatch(self):
        import sys
        sys.stdout = self.terminal

    def flush(self):
        self.log.flush()

    def logtofile(self, filename, clearcache=True):
        self.writetolog = filename != None
        if self.writetolog:
            self.log = open(filename, "a", 1)
            self.log.write(''.join(self.cache))
            self.log.flush()

            if clearcache:
                self.cache = []

    def write(self, message):
        self.terminal.write(message)
        if self.writetolog:
            self.log.write(message)
        else:
            self.cache.append(message)

# start the main method if the script was started directly
if __name__ == "__main__":
    main()
    input("Press any key to exit . . .")
```

Appendix VII: Refilled results for various interpolation techniques for Pepper Farm one.

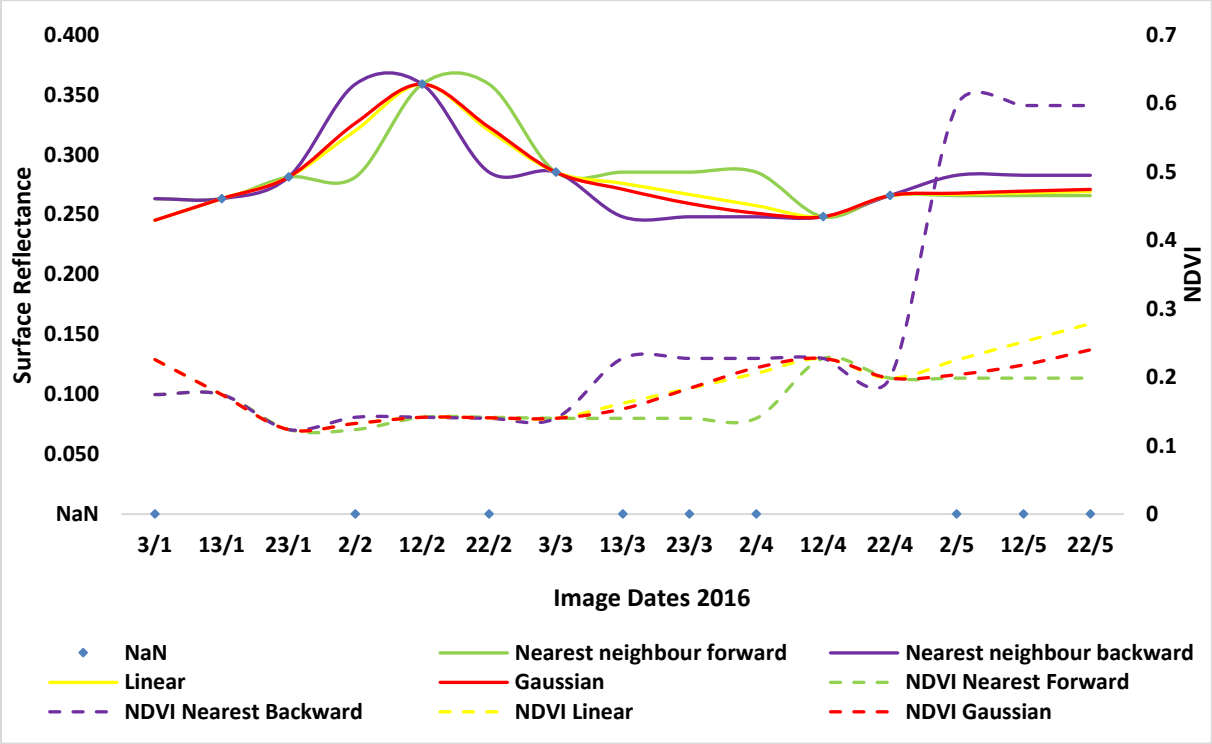
Planted in August and Harvested from November to April 2017



	Nearest Forward	Nearest Backward	Linear	Gaussian	NDVI Nearest Forward	NDVI Nearest Backward	Linear NDVI	Gaussian NDVI
AVERAGE:	0.263	0.244	0.254	0.253	0.353	0.297	0.325	0.328
STD DEV:	0.024	0.018	0.018	0.020	0.109	0.068	0.078	0.083
N	16	16	16	16	16	16	16	16
MEDIAN:	0.26	0.24	0.25	0.25	0.29	0.27	0.29	0.29
MIN:	0.226	0.226	0.226	0.226	0.246	0.246	0.246	0.246
MAX:	0.295	0.295	0.295	0.295	0.498	0.498	0.498	0.498
Q 1:	0.246	0.235	0.243	0.240	0.269	0.263	0.269	0.269
Q 3:	0.295	0.250	0.262	0.263	0.498	0.285	0.369	0.378
CLOUD FREE	9	9	9	9	9	9	9	9
CLOUDED	7	7	7	7	7	7	7	7

Appendix VIII: Refilled results for various interpolation techniques for farm six.

Season 1.Planted in February and harvested in May 2016 Season 2:Planted in July and Harvested in November 2016



	Nearest Forward	Nearest Backward	Linear	Gaussian	NDVI Nearest Forward	NDVI Nearest Backward	Linear NDVI	Gaussian NDVI
AVERAGE:	0.284	0.280	0.280	0.279	0.164	0.262	0.187	0.181
STD DEV:	0.032	0.035	0.031	0.032	0.033	0.177	0.048	0.040
N	15	15	15	15	`15	15	15	15
MEDIAN:	0.281	0.281	0.268	0.270	0.141	0.199	0.184	0.184
MIN:	0.248	0.248	0.245	0.245	0.123	0.123	0.123	0.123
MAX:	0.359	0.359	0.359	0.359	0.227	0.597	0.278	0.240
Q 1:	0.266	0.256	0.265	0.261	0.140	0.141	0.141	0.141
Q 3:	0.285	0.284	0.283	0.283	0.199	0.227	0.225	0.216
CLOUD FREE	6	6	6	6	6	6	6	6
CLOUDED	9	9	9	9	9	9	9	9