Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems

School of Information Systems

10-2017

Spatiotemporal identification of anomalies in a wildlife preserve

Bharadwaj KISHAN
Singapore Management University, kishanbs.2016@mitb.smu.edu.sg

Jason Guan Jie ONG
Singapore Management University, jason.ong.2016@mitb.smu.edu.sg

Yanrong ZHANG
Singapore Management University, yrzhang.2016@mitb.smu.edu.sg

Tin Seong KAM
Singapore Management University, tskam@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

Part of the Databases and Information Systems Commons, and the Geographic Information Sciences Commons

2.1 Raster image data

The size of the dataset for the raster images is too big to manipulate by software such as MS Excel. Therefore, all data are prepared by using R scripts by joining data, adding derived variables and so on.

2.2 Identifying new variables for analysis

2.2.1 Satellite Images

Figure 1: Table 1: Derived fields for raster image analysis.

Variables	Description	Formula
NDVI	Normalized Difference Vegetation Index is associated with vegetation content.	(B4-B3)/(B4+B3)
NDWI	Normalized Difference Water Index is correlated with the Water content.	(B2-B4)/(B2+B4)
NDMI	Normalized Difference Moisture Index reflects moisture in vegetated areas.	(B4-B5)/(B4+B5)
NDSI	Normalized Difference Snow Index highlights snow cover over land areas.	(B2-B5)/(B2+B5)
BSI	Bare Soil Index quantify the soil mineral composition.	[(B3+B5) +(B1+B4)]/[(B3+B5) - (B1+B4)]
AVI	Advanced Vegetation Index discriminate different type of vegetation	[(B4×(1-B3) × (B4-B3)] ^1/3

NDVI value is used to determine the regions of plant life, NDWI indicates the presence of a water region (e.g. the Boonsong Lake), BSI for Soil Region and Paths, which directly distinguish key features in the satellite image of the preserve. Low AVI value can indirectly tell the cloud region as well. For the analysis of Climate change, the difference of NDMI reflects the change of the moisture in the preserve and the difference of NDSI reflects the change of snow coverage. The most important indicator is the difference of NDVI value, which reflects the change of the plant health. Refer to the below section 3.2 to interpret the formulas for each variable.

2.2.2 Air Plume Model

The given wind direction in the original dataset is rounded to the nearest tenth to simplify the visualisation since the wind direction varies by a fixed interval of 10 degrees. Each row of the data was replicated three times at each time instance and each of this new data point is given a new label to identify it, i.e. point 1, 2 and 3. A length variable is created to assign a distance to each point and an additional angle parameter is also created to assign a different value for each point. Finally, the current rows are multiplied by 9 to represent the streams of wind captured at monitors 1 to 9. The additional variables generated are used primarily to construct the polygon representing the wind direction at each time instance. Each polygon has 3 points with point 1 having a length of 0 and points 2 and 3 having a length of 1. The angle parameter is used to

aid in the construction of the polygon with point 1 having a default angle of 0, point 2 having the same angle as the wind direction and point 3 having 10 degrees more. A total of 9 polygons are created to represent the wind direction captured at each monitor hence necessitating these steps.

3 VISUALIZATION TOOLS

3.1 Tableau

3.1.1 Developing the timestamp animation

Timestamp information is condensed into weeks, months and days to develop our calendar plots in Tableau. Each instance is added on the 'Pages' pane in Tableau. The path ID variable is created with the first timestamp instance for a car-ID appended with 1 and the last timestamp instance for a car-ID appended with the number of gates it has passed through, helping us trace the definitive chronological path of a vehicle. The coordinates of the vehicle movement can be annotated chronologically with the help of the gate name field. The 'pages' feature helps to animate for each timestamp instance, yielding an animated GIF feature involving both spatial and temporal aspects of anomalous vehicle movement.

3.1.2 Developing the air plume model

The default spread of the wind is assumed to be 10 degrees. However, this can be adjusted by an angle parameter. Setting the angle parameter to 5 for instance increases both ends of the spread by 5 degrees making the effective spread 20 degrees. A cropped image of the original image provided with MC2 which measures 350 by 200 pixels is used as the background image for the air plume model. Polygons are represented by the Cartesian coordinates of each point in the map, which must match the locations of the monitors given in the original image. Calculated fields are used in Tableau to assign the coordinates to the streams of wind captured by each monitor. There will be 1 calculated field in for the x-axis but 10 calculated fields for y-axis. 9 of them represents the streams of wind while the last one is used to align the background image to the height required.

3.1.3 Processing satellite imagery

Raster images are plotted by using diverging color for every pixel with its relevant measure value. For e.g., if measure value is NDVI and diverging color is from Red to Green, the graph would reflect the plant health of the preserve. The greener the image is, the healthier the plant is in that region. The graphs plotted below with different measure value to reflect performance in the preserve:
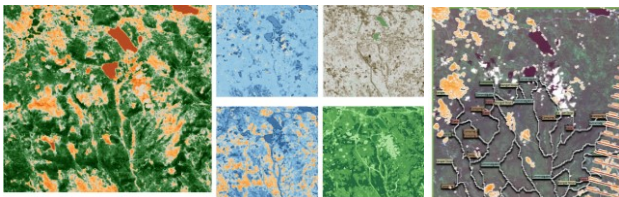

Figure 1: (Left most: Colored by NDVI, Middle (Clockwise): NDVI, NDWI, BSI, NDMI and AVI; Right most: Possible plant degradation area overlapped with Preserve map).

To identify the degradation, difference of NDVI value between two different Timestamps with the same month has to be taken. (Because there could be unknown seasonal factors affecting the band values). Since plants thrive in summer, we compare the difference of NDVI on June between 2015 and 2016. We utilised

the same approach to identify possible burning that has occurred in the preserve. We set our map as the background image and trace where there is deterioration or changes in surface profile. Cloudy satellite imagery impeded our analysis in a few cases.

3.2 QGIS

QGIS is a Geographic Information System' (GIS) software which allows us to process raster images. False color method uses three different wavelength bands as RGB values to visualize raster images to highlight certain features such as ice, cloud, burned land etc. In this case, we used some common false color combinations:
- Combination 1: B4(red), B2 (blue), B3(green) combination is useful in seeing changes in plant health.
- Combination 2: B5(red), B4(blue), B2(green) combination is useful in showing floods or newly burned land.
- Combination 3: B1(red), B5(blue), B6(green) combination is useful in differentiating between snow, ice, and clouds.
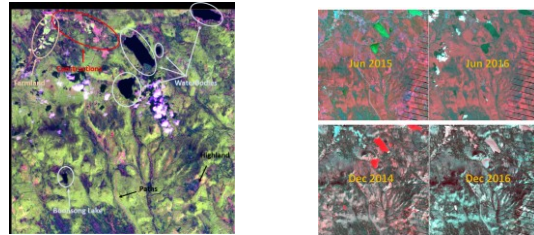

Figure 2: Images using False color method (Left: [B5, B4, B2]; Top right: ([B4, B2, B3]; Bottom right: [B1, B5, B6])

Combination 2 as RGB compares the difference between Sep 2014 and Sep 2016. There were several obvious clear-cut areas in pink color in Sep 2016 which were not in Sep 2014. Similarly, we use Combination 1 to show the overall change of plant health and Combination 3 to show the overall change of snow coverage in preserve before and after.

4 CONCLUSION

The paper illustrates how choosing the relevant visual analytics tools help in transforming and visualizing the data to compile evidence, both minor and major to add weight to a hypothesis. The timestamp animation yield anomalous vehicle patterns as a GIF which are superimposed on satellite imagery of plant health. Coinciding timeframes of suspicious patterns helped us piece the first building block of our hypothesis. Subsequently, the Air Plume model can attribute a chemical emission to one of the factories located near the preserve. Reinforced with evidence from the newsletters, we piece together the second block of our hypothesis. This is mainly made possible due to the visualizations in Tableau and the Raster images in QGIS that help to see patterns of interest varying on both temporal and spatial scales.

REFERENCES

[1] Tableau software. Make your data make an impact. [Online] https://www.tableau.com
[2] JMP software. Statistical discovery from SAS. [Online] https://www.jmp.com/en_us/home.html
[3] Wood, J. (2016). Visual Analytic Design for Contextualizing Sensor Data. Paper presented at the VIS 2016, 23-28 Oct 2016, Baltimore, USA.