Can the sentiment expressed in trail users' tweets help to assess the effectiveness of Environmental Stewardship Agreements? An exploratory analysis of the Pennine Way National Trail, England

• Tom Wilson1 • Robin Lovelace1 • Andrew Evans1

## Abstract

## Keywords

• big data analysis • sentiment analysis • Environmental Stewardship Scheme • Volunteered Geographic Information

Tom Wilson () [gy10tlw@leeds.ac.uk](mailto:gy10tlw@leeds.ac.uk)

1 School of Geography, University of Leeds, Leeds, United Kingdom

## Introduction

* Agri-environmental schemes
  + Environmental Stewardship Scheme
* Natural England and their role in manage both ESS and National Trails
* Study area

The focus of this research is the 431 km (268 mile) long Pennine Way National Trail (PWNT), the oldest of England’s 15 National Trails which officially opened on 24th April 1965 (Figure 1). Along its route the PWNT crosses agricultural land managed under ESS. For this analysis a 5km spatial buffer was drawn around the PWNT which will hereafter referred to as the PWNT corridor.

This research aims to:

* devise a process to extract the sentiment conveyed within trail users' tweets and perform exploratory analyses to determine whether this information can be used by Natural England to assess the effectiveness of ESS agreements

Through the following set of objectives:

* development of a process to select tweets relavent to the scope of the project and extract the sentiment conveyed within the selected tweets.
* cspatial analyses of tweet locations and tweet viewshed, to explore patterns that may exist.
* analysis of the finding and presentation of reccomendations to Natural England.

## Background

* Environmental Stewardship Scheme
  + what it is (larggest agri-environmental scheme in England). Broad and shallow approach. 5 main objectives:
    - Conservation of wildlife and their habitats
    - Maintenance and enhancement of landscape quality and character
    - Protection of the historic environment
    - Protection of soils and reduction of water pollution
    - Providing opportunities for people to visit and learn about the countryside.
* ESS aims to provide the funding and guidance to enable farmers and land managers to fulfil these objectives
  + how it works (payments to landowners for managing their land in an environmentaly concious way)
  + histroy of the scheme
    - Government’s response to increasing levels of agricultural intensification and its negative impacts on wildlife and landscape character. ESS follows on from Countryside Stewardship Scheme, is more flexible, and a broad and shallow approach (open to all).
  + significance of ESS in the context of the National Trail system
    - ESS agreements within the corridor of a National Trail have a particularly important role to play in providing positive experiences to users of the trail. One of the primary objectives of ESS is the maintenance and enhancement of landscape quality and character. Furthermore, Natural England identifies enhancement of the landscape, natural, and historic features within the trail corridor as one of its quality standards of the National Trail system.
    - Current methods to determine trail users' opinions
      * large scale qualitative surveys such as Monitor of Engagement with the Natural Environment (MENE): annual representative quota-sampled survey of ~50k people. Very broad definition of natural environment.
* Sentiment Analysis
  + what is it?
    - Seeking the sentiment conveyed by others is fundamental to the decision-making of individuals and organisations alike. As a consequence, sentiment is seen as a key influencer of action, and central to human behaviour (Liu 2012)
  + why conduct sentiment analysis (current uses and purpose)
    - In short there is a human desire to know what other people think.
  + approaches to sentiment analysis (lexical, machine learning)
  + sentiment analysis of social web challenges (abbreviations, slang, poor grammar)
  + SentiStrength tool used in this project (advantages and disadvantages, suitability for this research)
    - SentiStrength was developed by Thelwall et al., University of Wolverhampton <http://sentistrength.wlv.ac.uk>
* Social web data (big data) - Do we need to provide background about this?

## Data

* 60,434 geocoded tweets collected between 2014-06-03 and 2014-07-25
* GPX file of the route of the Pennine Way National Trail (PWNT) (Walk Unlimited 2014)
  + with 5km and 25km buffers derived
* Shapefile of ESS agreement boundaries in England (Natural England 2014c)
  + clipped to 5km and 25km PWNT buffers
* Land Cover Map 2007
  + clipped to 5km and 25km PWNT buffers
* Digital Elevation Model
  + clipped to 5km and 25km PWNT buffers
  + low resolution (90m) SRTM
* PWNT trail counter data (provided by Natural England)

## Methods

* Spatial selection
  + in R the Twitter dataset was reduced to only the tweets which originated from within the PWNT 5km corridor
* Lexical selection
  + in R the tweets relevant to hiking the PWNT were selected using natural language processing
* The text of each tweet was processed and ‘cleaned’ of spurious characters
* Each tweet and its spatial information was imported into ArcMap 10 (ESRI, 2011)
* The TweetText and TweetID were input into SentiStrength for sentiment analysis
* Sentiment analysis outputs were combined with tweet spatial information in ArcMap 10/QGIS
* Viewshed analyses were conducted for each overall positive and overall negative tweet. Viewshed analyses were conducted in ArcMap 10 (ESRI, 2011) and included:
  + Calculation of the viewshed
  + Determination of majority land cover class within the viewshed
  + Determination of the ruggedness within the viewshed
  + Determination of the presence of ESS agreements within the viewshed.

## Results

* The sentiment expressed in the twitter messages along the PWNT trail corridor:
  + 47 tweets expressed positive sentiment,
  + 22 negative sentiment
  + 102 expressed no sentiment (i.e. neutral).
  + 10 tweets expressed both positive and negative sentiment.
  + Overall tweet sentiment consisted 40 positive, 16 negative and 105 neutral.
  + 105 tweets did not convey any sentiment and were classed as neutral. It was discovered that of these 105 tweets, 94 contained a URL within the TweetText.
    - Missing sentiment in images?
  + [Link to interactive map of trail user tweets](http://tom-wilson.info.s3-website-us-west-2.amazonaws.com/PennineTweets.html)

## Discussion and Reccomnedations

* limitations
  + issues with social web data
    - limitations of twitter API - opacity, accessibility, representation and digital divide, ethics
  + mobile phone coverage on remote National Trails
* reccomendations
  + social campaign
  + explore image-sharing and the sentiment of images

## Conclusions

## References

Liu, B. 2012. “Sentiment Analysis and Opinion Mining.” *Synthesis Lectures on Human Language Technologies* 5 (1): 1–167.

Natural England. 2014c. “Environment Stewardship Scheme Agreements.” *[Shapefile] Created by Natural England, August 2014*. <http://www.geostore.com/environment-agency/WebStore?xml=environment-agency/xml/ogcDataDownload.xml>.

Walk Unlimited. 2014. “The Pennine Way.” <[online] Available at: http://nationaltrail.co.uk/pennine-way. Accessed 7 January 2015>.