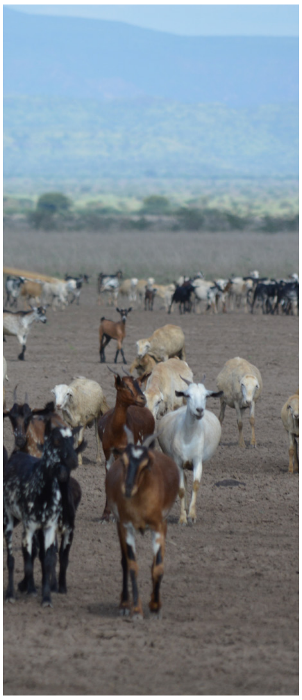


ILRI RESEARCH BRIEF 84
October 2017

Crowdsourcing for rangeland conditions—Process innovation and beyond

Oscar Naibei, Nathan Jensen, Rupsha R Banerjee and Andrew Mude

Rangeland forage conditions are so closely associated with the wellbeing of pastoralists that they are frequently used by humanitarian agencies to monitor for coming disasters (Luseno et al. 2003; Aboud et al. 2006). Information on forage conditions is also necessary for rangeland management programs, and is used by development and government programs working to mitigate the risks that pastoral households face (Kurtz et al. 2010; Kadi et al. 2011, Fernández-Gimenez et al. 2012; Chantarat et al. 2013).

The gold standard for monitoring vegetation conditions requires constant surveillance by teams of ecologists. However, this approach is extremely labour intensive and costly. Remotely sensed data offers a cost effective alternative, but suffers from its own set of drawbacks, many of which relate to questions about the relationship between remotely sensed data and the actual conditions on the ground (Cihlar et al. 1997; Franklin 2001).

The satellite-based Advanced Very High Resolution Radiometer (AVHRR) sensor has provided daily time series of normalized difference vegetation index (NDVI) data from the across the earth for more than a decade (Zhang et al. 2003, White et al. 1997). However, although the NDVI data is of high temporal and spatial resolution, the resolution is insufficient to distinguish between plant species and, thus, the palatability of the vegetation that it is sensing. In the case of forage monitoring efforts, it is important that index readings reflect the conditions of palatable vegetation species and that the aggregation of all vegetation into a single index misses important variations.

Crowdsourcing is an innovative data collection approach that can potentially be exploited to address these challenges by collecting additional data that complement existing methods. Generally, crowdsourcing refers to the act of outsourcing work to an undefined and often large group of people (Howe 2006). In this case, crowdsourcing leverages digital technology and local knowledge to gather low-cost and near real time data on vegetation type, palatability and carrying capacity to improve existing forage models relying on remotely sensed data.

Index-based livestock insurance

The Index-Based Livestock Insurance (IBLI) product seeks to mitigate the effects of drought on the welfare of pastoral households by providing indemnity payments when forage conditions are extremely poor. IBLI relies on an index of remotely-sensed NDVI data, a measure of greenness. The IBLI index is generated by normalizing NDVI data with respect to historic values, so that index values identity relative vegetation conditions.

Although NDVI-based indices have been shown to be quite accurate proxies for forage conditions, the IBLI project could benefit from methods for calibrating the index and verifying actual conditions on the ground—potentially allowing for low-cost auditing to reduce basis risk associated with the index product.

Methods

Crowdsourcing is not a new technique for collecting, processing or disseminating information. It has been used for quite some time over a wide range of areas, such as climate data, photography, traffic congestion and safety (Howe 2006; Schenk and Guittard 2009; Chiu et al. 2014). The Crowdsourcing for rangeland conditions project—implemented through a collaboration between the International Livestock Research Institute (IRLI), Cornell University and the University of Sydney—applied a crowdsourcing approach to collect detailed information on forage conditions.

The impetus for the project was to examine cost-effective ways of providing local high-frequency information that could be used to calibrate the index used by the IBLI contracts (see box). ILRI's IBLI project recruited a team of 112 local pastoralists—leveraging their mobility and expert knowledge of forage—to collect information across the 22,500 km^2 study site. The study participants were trained to collect data using a survey containing close-ended questions constructed specifically to determine the palatability of vegetation affecting the NDVI index. The surveys were loaded onto mobile phones—provided to participants by the project—so that they could collect and submit the surveys as they went about their daily activity of herding livestock.

Rational and study design process

Site selection and pre-testing

The project was piloted in Isiolo county, northern Kenya. Isiolo was selected as the study site because the local topography of the county is arid and semi-arid low plains, and pastoralism is the main livelihood activity. In addition, Isiolo has relatively good mobile network coverage, when compared to the other pastoral regions, making it ideal for the iterations of pre-testing and piloting that were needed.

Researchers pre-tested the survey approach to rangeland monitoring and vegetation classification during focus group discussions with pastoralists in the region. During the discussions, they asked groups of participants to describe the characteristics of the existing vegetation conditions. Researchers then discussed the characteristics that they had identified as important for the project. Specifically, researchers interested in distinguishing between vegetation types (grass, shrubs or trees), the abundance of each type, how green each were, and the palatability of each for goats, cattle and camels. They also tested a set of icons that would be used in the survey to indicate vegetation types, abundance, and leaf colour. Photos of nearby and distant rangelands in various conditions were used to provoke discussion and help researchers improve the survey questions and icons (Figure 1).

The structure of crowdsourcing tool

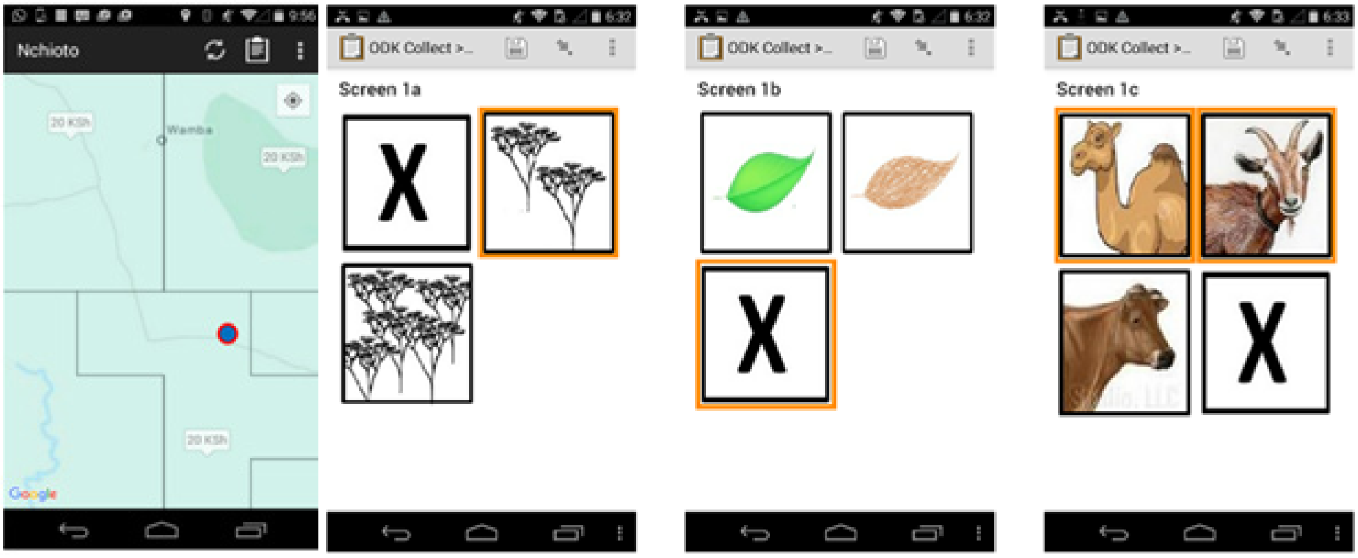
Mobile phone-based platforms were the core technology enabling the project. The mobile tool had two main components. The first was a custom-made application, Nchioto (the one that gives direction), displaying a map of the research area, the phone's location, and the rewards associated with making a submission from any of the nearby regions. The rewards could be dynamically and remotely updated by the researchers in order to adjust the spatial distribution of submissions. The incentive for each submission was between KES 5 and 40 (USD 0.05–0.4).

The second component—a survey tool, launched using the open data kit (ODK)—relied on icons and audio to ensure that low literacy levels did not impair data collection (Figure 2). Each submitted survey included a photo and a set of responses to the survey questions, and was identified with a specific participant, time stamp, and geo-location. In places with no or poor cellular network connectivity, the surveys were stored on the phone until the participant entered a region with strong network coverage; at which time they were automatically submitted. Participants were instructed that all surveys should be completed during daylight hours and at a minimum of 60-minute intervals. They were paid a small reward for each submission in line with those represented on Ntioto at the time of completion.

The project was launched with a two-day training workshop for the participants. For supervision and training purposes, researchers subdivided the areas into five operational sites, each comprising about 20 participants. The training was split into two sessions. The first day focused on discussing the motivation for crowdsourcing conditions, learning the rangeland classification process, and the survey questions. The second day focused on the technical aspects, such as how to open and use the two applications, how to update the Nchioto map, and general trouble shooting.

Figure 1. Examples of photographs used to discuss rangeland conditions.

Figure 2 – Screen shots of the Ntioto application (left) and questions on trees from the vegetation survey (right).



Participants submitted over 110,000 submissions during the 148-day pilot. During this time, researchers ran a randomized control trial in which rewards were varied across groups of participants, across space and across time. This was undertaken to determine the extent to which researchers could affect the spatial distribution of submissions. Data analysis showed that the dynamic rewards successfully reduced the clustering of submissions, both increasing information from regions otherwise under-sampled and reducing costly redundant submissions (Jensen et al. 2017a).

Data validation

Data validation was undertaken by ensuring that the vegetation conditions in each survey matched its accompanying photo. The validation tasks were performed by volunteers through Cornell's UDiscoverIt website (http://www.udiscover.it)[1] and by workers on Amazon's Mechanical Turk (https://www.mturk.com)[2]. The validators were responsible for flagging low-quality photos and identifying obvious features (e.g. green grass, trees) in the submissions, which could then be used to cross validate the accompanying survey data[3]. Each studied submission was validated multiple times, and submissions were only flagged if they failed multiple validations. These validations were then fed into a randomized control trial in which a sub-set of participants were periodically provided with feedback on the quality and quantity of their recent submissions. Analysis of submission quality shows that such feedback does improve data quality (Jensen et al. 2017b).

Discussion

As the stakeholders in pastoral areas continue to grapple with the challenge of poor access to information on reliable rangeland conditions, the need for improved methods for collecting such information is of great importance. This project demonstrated that crowdsourcing can successfully be used to collect accurate, low-cost and real time data on rangeland conditions. Furthermore, the feedback mechanisms and a dynamic reward system helped to mitigate some data quality and sampling issues known to plague crowdsourcing efforts.

This approach has the potential to revolutionize and improve the process of rangeland monitoring. For instance, crowdsourcing approaches to rangeland monitoring can also be used by the National Drought Management Authority to validate and expand their existing monitoring systems. With the growing interest in long-term rangeland management, such methods could be used by local communities in partnership with other relevant stakeholders as a reporting and monitoring platform on the state of rangelands.

More generally, this undertaking has demonstrated that local data can be collected in a cost-effective manner in remote areas that have difficult terrain, sparse mobile phone coverage, and low levels of literacy by community members. These methods hold considerable potential for improving information for a host of other data-challenged fields in arid and the semi-arid land areas—including market prices, disease surveillance and migration. In response to these findings, the IBLI team has recently expanded its agenda to develop a flexible platform for crowdsourcing the collection other data types.

Acknowledgements

The authors would like to thank Chris Barrett, Rich Bernstein, Eddy Chebelyon, Carla Gomes, Russell Toth and Yexing Xue for their invaluable contributions to the research design and implementation, along with the pastoralists who agreed to be the crowdsource agents and participated in this study. The authors also acknowledge

1 The authors would like to thank all the students, friends, family and strangers who volunteered their time for this effort.

2 Amazon's Mechanical Turk, an online labour market, facilitates the matching supply and demand of labour for small tasks easily completed over the internet. In this case, the tasks were to make observations on the quality and content of photos submitted with each survey.

3 For more details, see: econthatmatters.com/2015/06/from-pastoralists-to-mechanical-turks-using-the-crowd-to-validating-crowdsourced-data.

funding and resource support from the Atkinson Center for a Sustainable Future's Academic Venture Fund; ILRI's IBLI project; and National Science Foundation's Expeditions CompSustNet: Expanding the Horizons of Computational Sustainability, Award CCF-1522054.

References

Aboud, A. et al. 2006. *Natural Resource Management and Biodiversity Conservation in the drylands of eastern and central Africa*. ASARECA.

Chantarat, S. et al. 2013. Designing index-based livestock insurance for managing asset risk in northern Kenya. *Journal of Risk and Insurance* 80:(1):205–237.

Chiu, C.M. et al. 2014. What can crowdsourcing do for decision support? *Decision Support Systems* 65:40–49.

Cihlar, J. et al. 1997. Multitemporal, multichannel AVHRR data sets for land biosphere studies—artefacts and corrections. *Remote Sensing of Environment* 60:35–57.

Fernandez-Gimenez, M.E. et al. 2012. *Lessons from the Dzud: adaptation and resilience in Mongolian pastoral social-ecological systems*. Washington DC, USA: World Bank.

Franklin, S.E. 2001. *Remote sensing for sustainable forest management*. CRC Press.

Howe, J. 2006. *Crowdsourcing: A definition. Crowdsourcing: Tracking the rise of the amateur*.

Huete, A. et al. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment* 83:195–213.

Jensen, N. et al. 2017a. *Don't follow the crowd: Incentives for directed spatial sampling*. Conference Paper Agricultural and Applied Economics Association. Annual Meeting, 30 July–1 August 2017, Chicago, Illinois, USA.

Jensen, N. et al. 2017b. *Monitoring A, while hoping for A and B: Field experimental evidence on multidimensional task monitoring*. Working paper.

Kadi, M. et al. 2011. *The state of climate information services for agriculture and food security in East African countries*. CCAFS Working Paper 5. Copenhagen, Denmark: CCAFS.

Kurtz, D.B. et al. 2010. Ground and satellite based assessment of rangeland management in sub-tropical Argentina. *Applied Geography* 30:210–220.

Luseno, W.K. et al. 2003. Assessing the Value of Climate Forecast Information for Pastoralists: Evidence from Southern Ethiopia and Northern Kenya. *World Development* 31:1477–1494.

Sakamoto, T. et al. 2005. A crop phenology detection method using time-series MODIS data. *Remote sensing of environment* 96:366–374.

Schenk, E. et al. 2009. *Crowdsourcing: What can be outsourced to the crowd, and why*. Workshop on open source innovation. Strasbourg, France.

White, M.A. et al. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* 11:217-234.

Zhang, X. et al. 2003. Monitoring vegetation phenology using MODIS. *Remote sensing of environment* 84:471–475.

Photo credit
ILRI/Nathan Jensen

Oscar Naibei, Nathan Jensen, Rupsha R Banerjee and Andrew Mude all work for the International Livestock Research Institute.

Contact

Oscar Naibei
ILRI Kenya,
o.naibei@cgiar.org
https://ibli.ilri.org.

ILRI thanks all donors that globally support its work through their contributions to the CGIAR system

Patron: Professor Peter C Doherty AC, FAA, FRS
Animal scientist, Nobel Prize Laureate for Physiology or Medicine–1996

Box 30709, Nairobi 00100 Kenya
Phone +254 20 422 3000
Fax +254 20 422 3001
Email ilri-kenya@cgiar.org

ilri.org
better lives through livestock

ILRI is a CGIAR research centre

Box 5689, Addis Ababa, Ethiopia
Phone +251 11 617 2000
Fax +251 11 667 6923
Email ilri-ethiopia@cgiar.org

ILRI has offices in East Africa • South Asia • Southeast and East Asia • Southern Africa • West Africa