DARE Workshop, 20-22 Nov 2017
Data science for high impact weather and flood prediction

Geoff Parkin
School of Engineering
Newcastle University

Community-sourced flood data
Acknowledgements

Eleanor Starkey, PhD (2017), now at JBA
David Walker, PhD candidate
Paul Quinn, Andy Large, Phil James, John Gowing
Tyne Rivers Trust, International Water Management Institute (IWMI)
Communities at Morpeth, Haltwhistle, Dangila (Ethiopia)

A brief glossary of terms

**Citizen science**
- Overall term for the process in which members of the public engage in activities with professional scientists
- Various levels of engagement possible

**Crowd-sourced data**
- Lowest level – no or little active participation, eg social-media harvesting

**Community-based data**
• High level of engagement including problem definition

**Other terms:** ‘participatory’, ‘volunteered geographical information’ (VGI), ‘co-production of knowledge’ etc

What is the interest for communities?

Flood-related information:
• Flood risk reduction (ie planning)
• Real-time flood warning
• Measurements of rainfall, river levels (also flows), flood extents, for post-flood assessments

Other interests:
• Erosion and sediment movement, water quality, habitats...
• Support improvement of flood models (calibration, validation, working towards real-time data assimilation)

What is the interest for NWP data assimilation?

• Potential for more ground-based observations
• Observers may have motivation for relatively high-quality observations if properly organised
• Observations at key locations (high flood risk)
• River level/flow measurements can help to constrain spatial rainfall estimates in catchments (model validation)

Although...
• Pressure, temperature, humidity are not of main interest
  • Much of the data may not be available in real-time

Key questions (multiple issues including flooding):
• Can communities feasibly monitor their catchment using simple, low-cost methods?
• Are community-based data reliable and meaningful?
• Can community-based data be used to inform modelling of catchment response?
• Is a community-based approach sustainable and scalable?

What motivates communities to carry out monitoring?
Case studies: Morpeth, Ethiopia, Haltwhistle

Overall framework

“Experts”

1\textsuperscript{st} responders, EA

Modellers, forecasters

Expert predictions
Crowd/community sourced data
(Real-time) information sharing

Communities

Historical data archive
Historical perspective: extension of recorded data

Historical evidence from newspaper archives and other sources (e.g., gravestone marks) used to create a chronology of flash flood events.

**Figure 1**  Time series of flash floods by decade from 1800 to 2010 divided by severity for (a) Northeast England.

*Source:* Archer et al. (2016)
Value of community observations: local responses
Evidence for flood response in ungauged catchments (Morpeth)

Source: Archer et al. (2016)
Case study: Morpeth
Crowd-sourcing
Case study: Morpeth

View from my house, 6 Sept 2008
Synoptic situation – 6th Sept 2008, 12.00
Hourly rainfall
River stage in town centre (EA recorder)

Water level rise
approx. 3.7 m (12 ft)

Peak level approx.
28.25 mAOD at 17:30
(problems with logger
at peak)
Crowd-sourcing: people wanting to tell their stories
Observation locations
Interpolated observations (not modelled data)
Additional value – interpretation of flood mechanisms

Surface water flooding and poor drainage early during the flood prevents emergency access to High Stanners.

Surface water - 11 am

River water - Approx. 1 pm
Additional value – interpretation of flood mechanisms

Flood levels were high on streets in Middle Greens from surface water and drainage discharges and possibly through hydraulic connections to the river (leaky joints, ineffective flap valves?) even before the wall is overtopped from the river.
Rainfall: Radar and Mogreps 1.5km ensemble member
Mogreps ensemble member rainfall 30-hour totals

a) font reservoir

b) harwood
c) wallington
d) newbiggin
Spatial rainfall comparison

Rainfall pattern interpolated from daily raingauge totals

Pattern of mean 1km rainfall measured from probabilistic forecast model

1st day (9am 5th to 9am 6th Sept)

1st day (12am 5th to 9am 6th Sept)

2nd day (9am 6th to 9am 7th Sept)

2nd day (9am 6th to 6pm 6th Sept)
Flood modelling comparisons: extents and depths
Case study: Ethiopia

Community-based monitoring
Fig. 1. The global network of World Weather Watch stations colour-coded to show reporting rates (WMO, 2003).
Walker et al. (2016)
Case study: Ethiopia
Reliability: Ethiopia study

Assessments made of community-based observations of rainfall, river flows, groundwater levels

The quality of community-based rainfall observations in the Ethiopia study equals or exceeds that of formal monitoring.
network and gridded rainfall products, taking account of spatial variability

: Walker et al. (2016)
Use of community-based data for modelling groundwater potential mapping.

Groundwater potential zones:
- High
- Medium
- Low
- Very low

High potential areas match the darker more agricultural areas.

Duration of groundwater recession indicates potential.

Based on high-resolution hydrological modelling.

Gendered focus groups and participatory mapping.
Case study: Haltwhistle

Community-based monitoring
Case study: Haltwhistle Burn

PhD study, part of multi-purpose Catchment Restoration Fund (CRF) project
Case study: Haltwhistle Burn

Simple, low-cost monitoring

Standardised data collection methods, minimal training, guidance documents
Case study: Haltwhistle Burn

Use of digital technologies
Case study: Haltwhistle Burn

Web host for data (feedback to communities)
http://research.ncl.ac.uk/haltwhistleburn/

Background guidance document produced on approaches and methods (>8,000 purchases and downloads)
http://www.fwr.org/Catchment/frr0021.pdf
Case study: Haltwhistle Burn

Knowledge exchange ⇒ Collect data ⇒ Submit ⇒ Share ⇒ Feedback

River level and rainfall data for Townfoot (Haltwhistle)

Observations made by the Citizen Scientists throughout 2014

Move your mouse over the interactive graph below...
: Starkey (2017b)

Source: Starkey (2017a)
Data use: qualitative understanding
Community-based timeline for 30th April 2014 event

Source: Starkey (2017a)
Data use in hydrological modelling

Source: Starkey (2017b)
Shetran catchment model setup with river gauge locations

Source: Starkey (2017a)
Data use in hydrological modelling

Source: Starkey (2017a)
Source: Starkey (2017a)
Modelled scenarios (one-at-a-time excluded data)

Source: Starkey (2017a)
Potential: use in hydrological modelling

Source: Starkey (2017a)
April 2014 event hydrographs – rapid rise is only reproduced for scenarios which include community-based data

Data use in hydrological modelling

Relative model performance for each scenario

Source: Starkey (2017a)
Peak discharge underestimated and time of peak delayed in models which do not include community data (B, C, H)

Source: Starkey (2017a)
Motivation: Haltwhistle Burn community
Some survey findings - community willingness to participate:
• hotspot areas identified for monitoring
• Some people prefer routine: rely on them for ongoing observations
• Some people prefer the ‘exciting events’: rely on them for detail during peaks
• Some people stick to their usual routes, others travel to the middle/upper regions of the catchment
• Photos and videos most common form of observation
• Rainfall monitoring very popular
• Technology: people still put off by it
• Training cards have helped significantly: standardise approaches, ensures date, time, location

Some evidence of sustainability
Another Northumberland community, Acomb, initiated and are running their own monitoring system independently
Community-led monitoring: manual & automatic methods (includes rain gauge at the school)

Community-led monitoring: live data streaming from automatic WLRs every 15-min (x3 locations)

"This early warning system has been invaluable [...] At times of heavy rain, I can see, particularly on the Birkley Burn, what is happening 1 mile upstream [...] and hence take the necessary action" Flood Warden

Source: Starkey (2017b)
Potential for use of community-based observations
Catchment partnerships (evidence- & participatory-based)

Movement of ground-based knowledge (identifying what really matters on a local level)
Source: Starkey (2017b)

Who may be interested in community-based observations?

- The communities themselves:
  - Better understanding of their own environment
  - Confidence in their own information

- Rivers Trusts and local community groups:
  - Identify issues and activities for catchment improvements

- Lead local flood authorities:
  - Information on flood events and locations on smaller watercourses
• Environment Agency / SEPA:
  • Peak river levels and flood extents, for updating models
  • Inform emergency responses

Who may be interested in community-based observations?

• Flood action groups and partnerships:
  • Informing flood protection and emergency response planning

• Schools:
  • Improve learning, geographic sciences, maths ...

• Public / media:
• General understanding of climate and flooding
• National science organisations (CEH, Met Office, universities ...):
  • Supplement national monitoring networks
  • Project-specific studies, eg understanding impacts of interventions

Next steps

Key issues include:
• Developing new low-cost technologies (eg camera-based direct river flow measurements)
• Data collection with a purpose
• Standards for data collection
• Data quality control and quality assurance
• Open data sharing (eg WOW)
• Development of appropriate interpretative models
• Visualisation of data and interpretations

New PhD student studying data quality and ‘value’ of community-based data
Visualisation: Geovisionary software
Realistic visualizations of data and model scenarios

Scalable overlays of multiple datasets including near-real-time data streaming
Thank you for listening
Any questions?
DARE Workshop, 20-22 Nov 2017
Data science for high impact weather and flood prediction

References