DARE Workshop, 20-22 Nov 2017

Data science for high impact weather and floodprediction

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

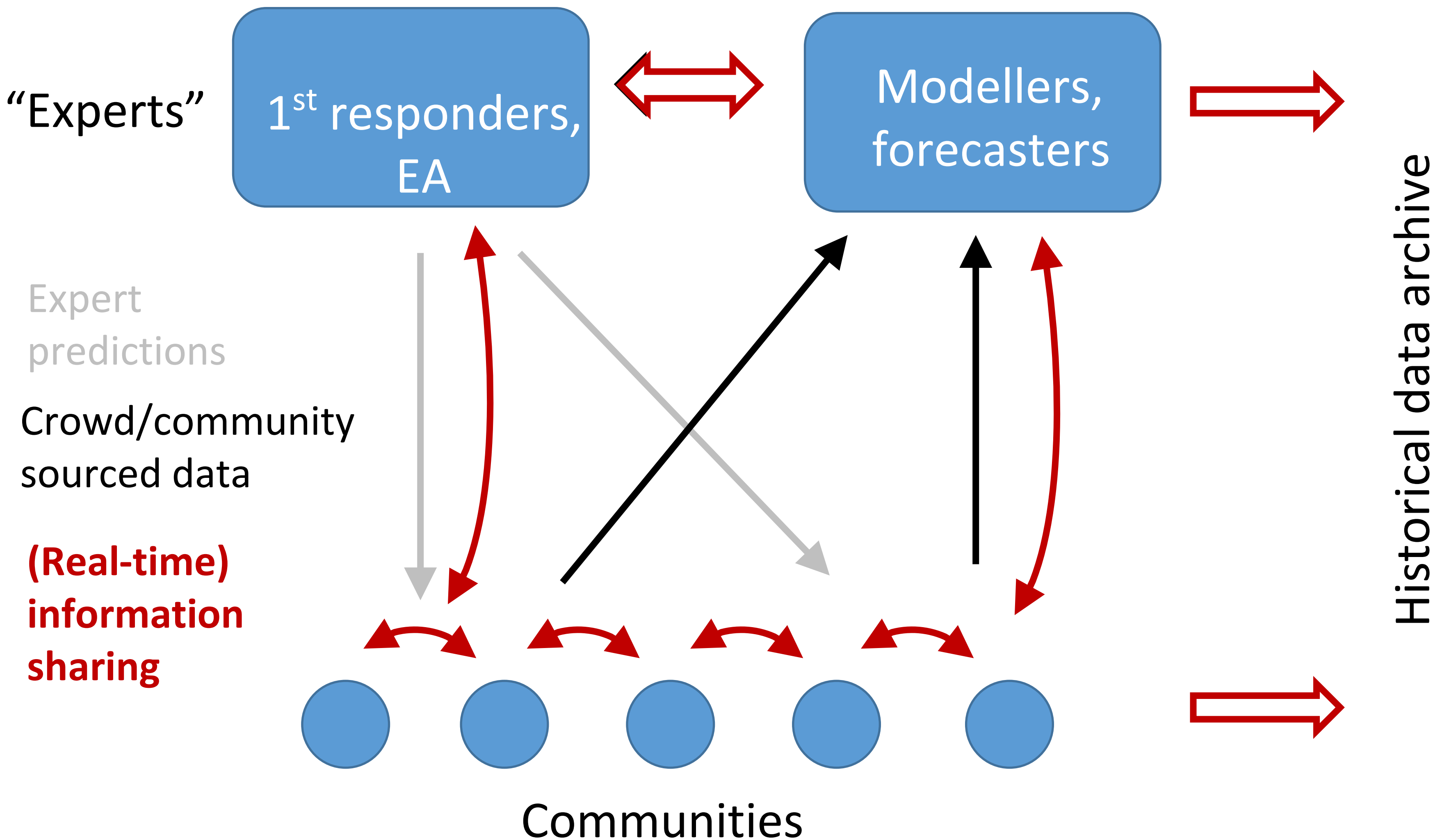
Key questions (multiple issues including flooding):

- 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



Historical perspective: extension of recorded data

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

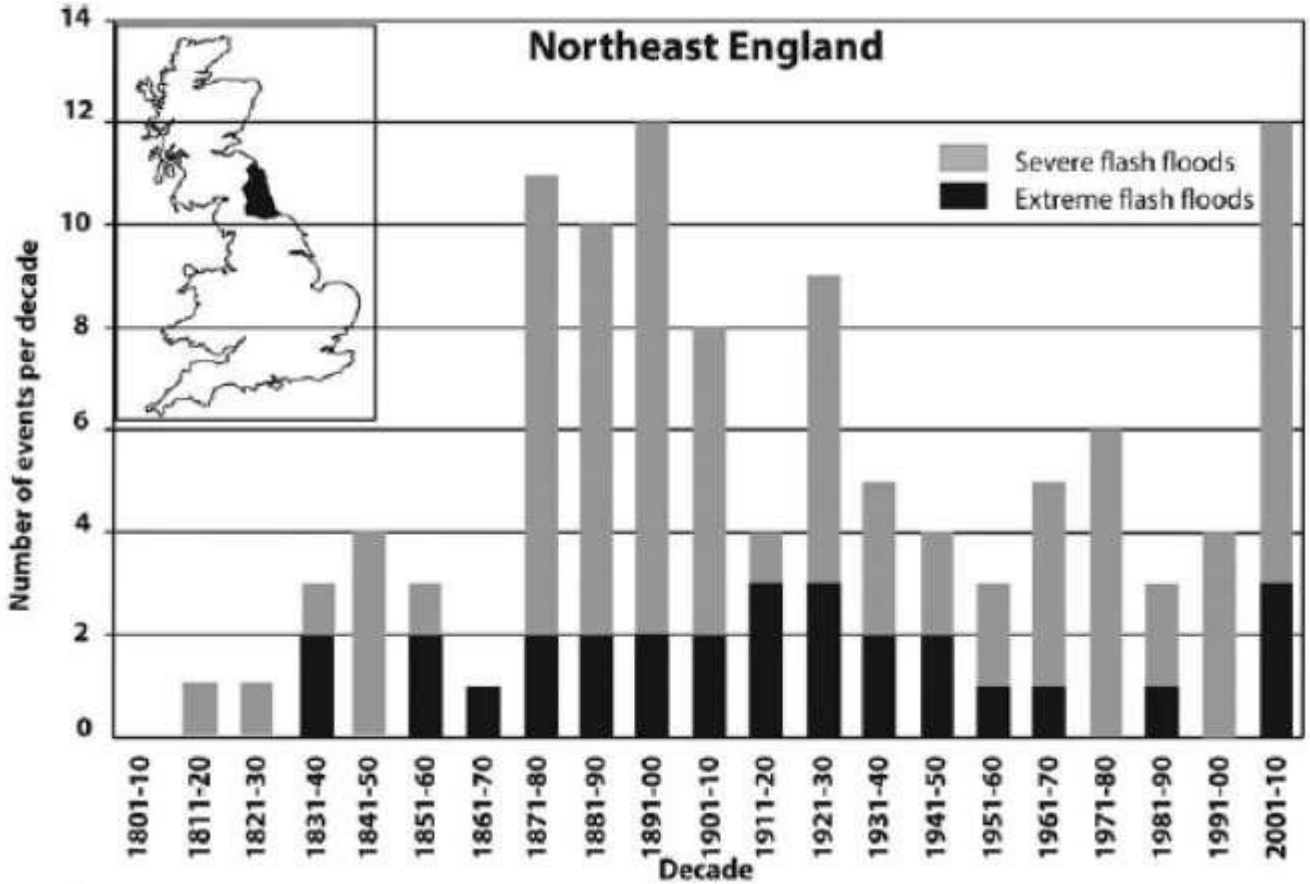


Figure 1 | Time series of flash floods by decade from 1800 to 2010 divided by severity for (a) Northeast England

Source: Archer et al. (2016)

Value of community observations: local responses

ungauged catchments (Morpeth)

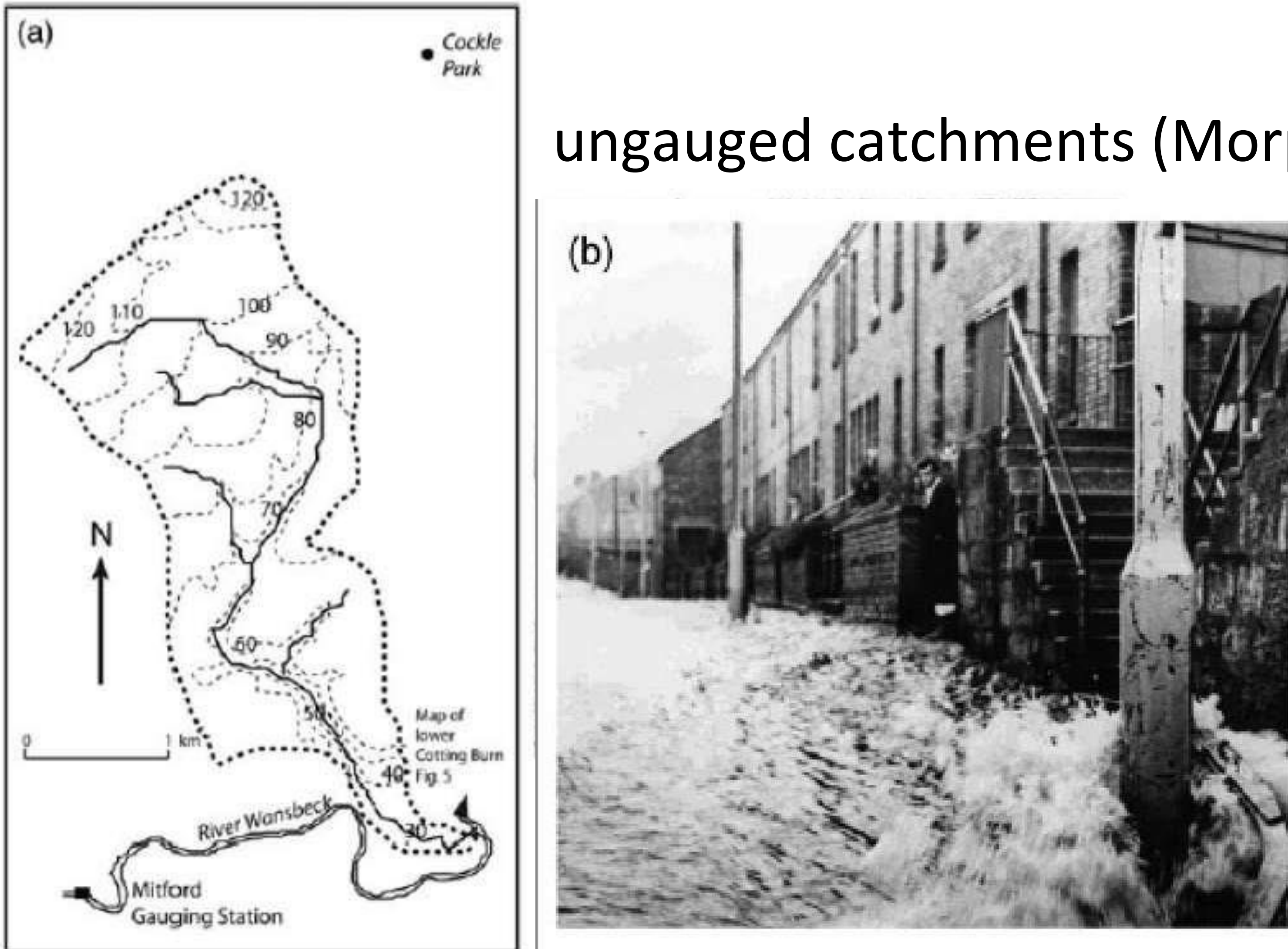


Figure 4 | (a) Cotting Burn catchment and (b) September 1968 flash flood in Dacre Street adjacent to Swinney's Iron Works, Morpeth.

Evidence for flood response in

Source: Archer et al. (2016)

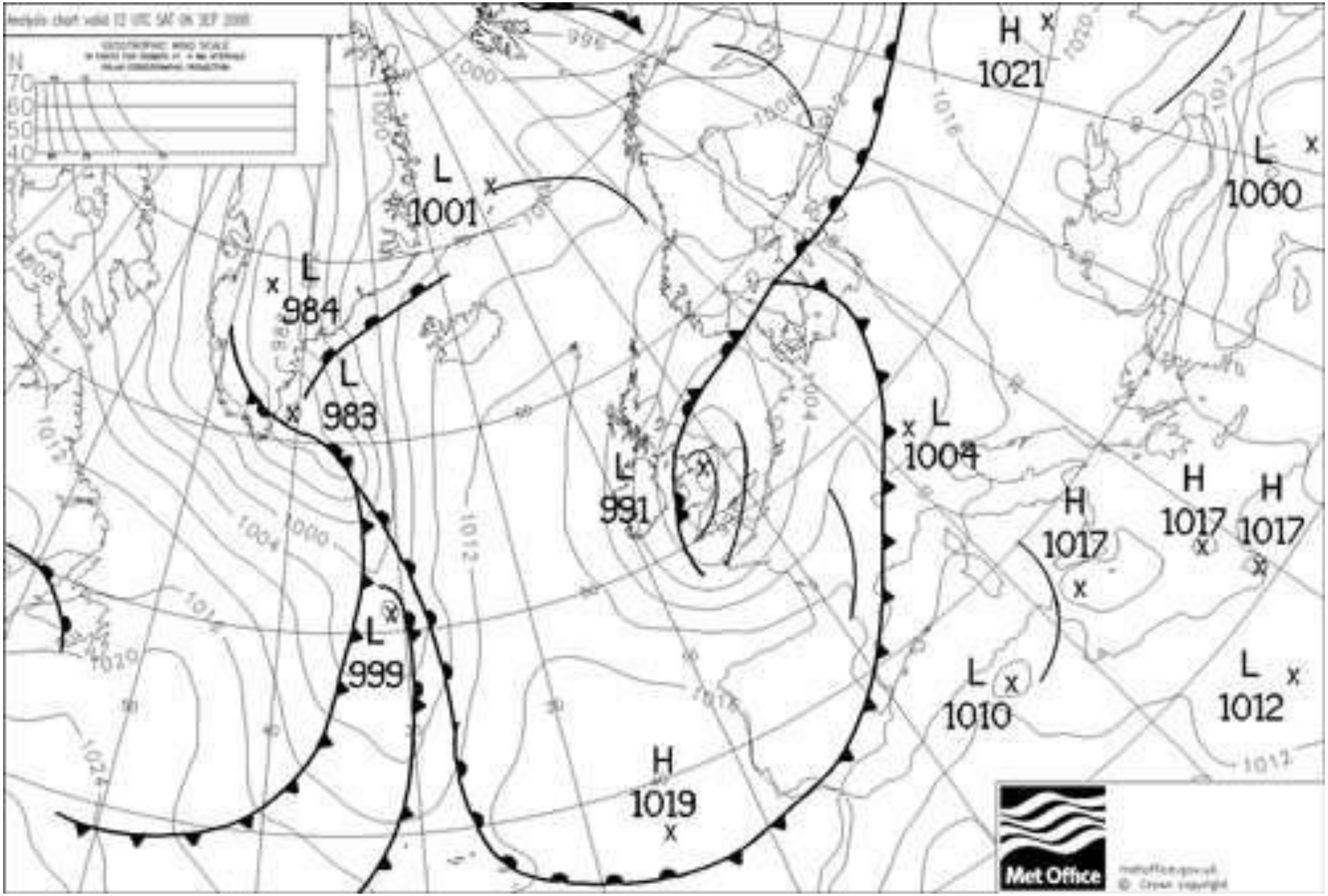
Case study: Morpeth

Crowd-sourcing

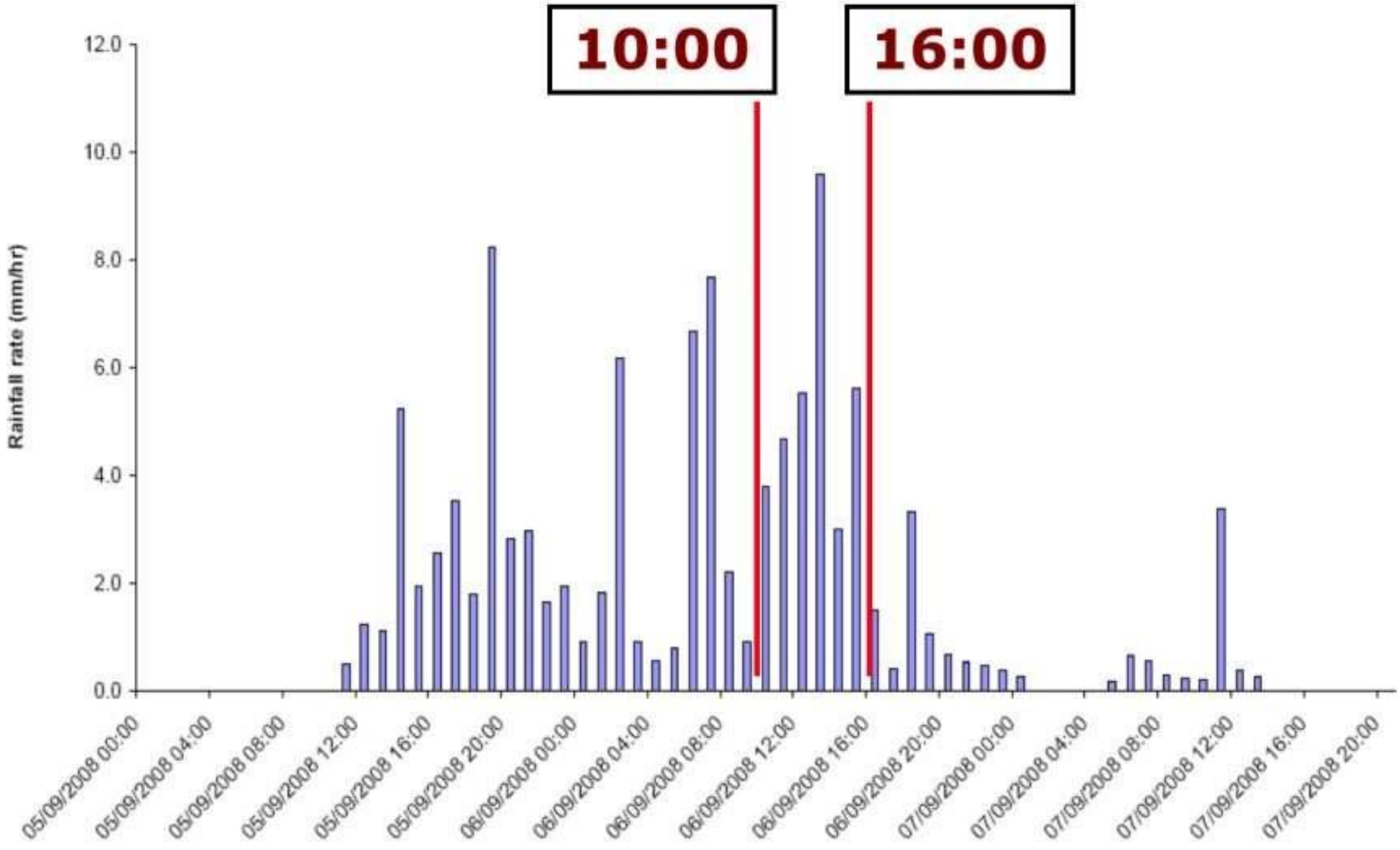
Case study: Morpeth

View from my house, 6 Sept 2008

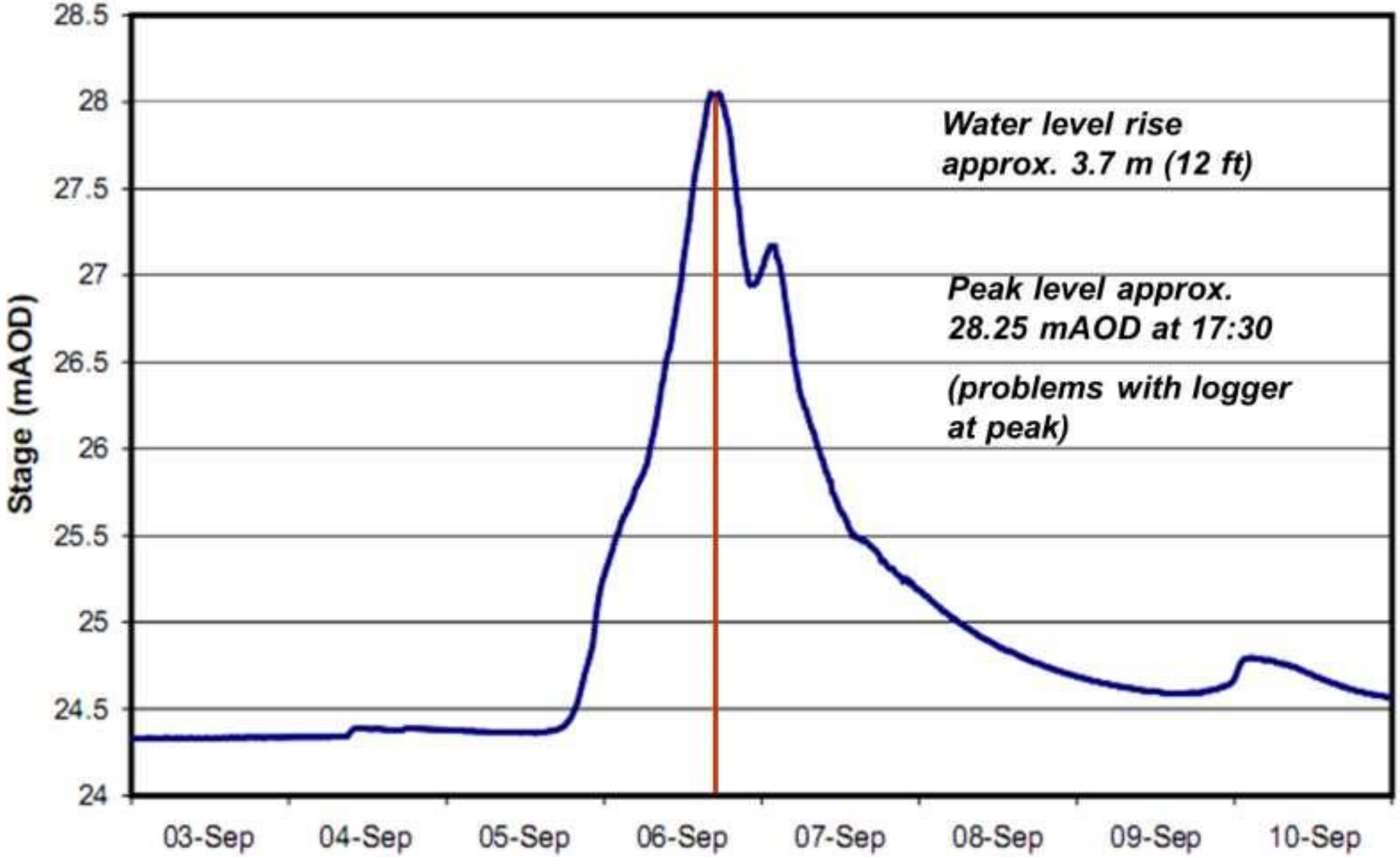
Synoptic situation – 6th Sept 2008, 12.00



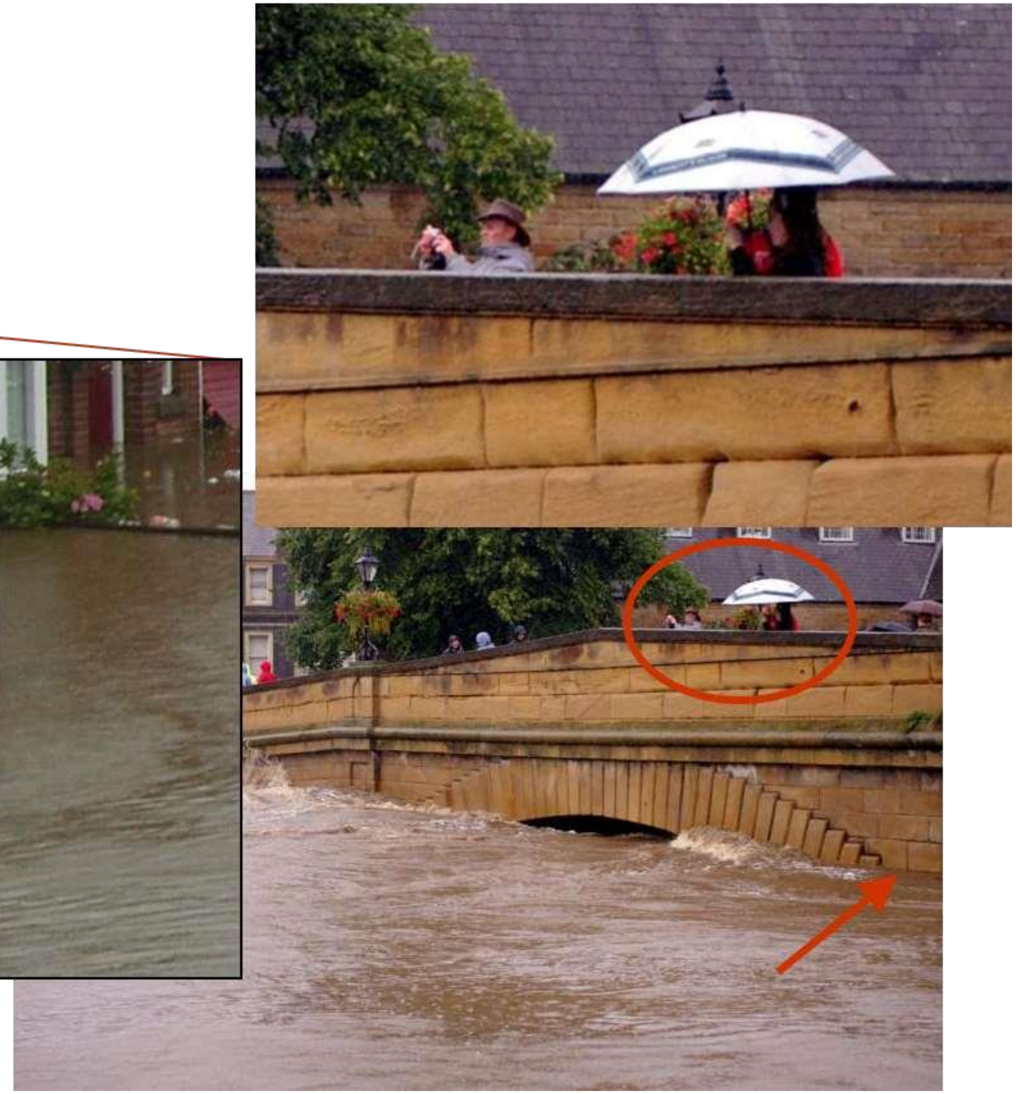
Hourly rainfall



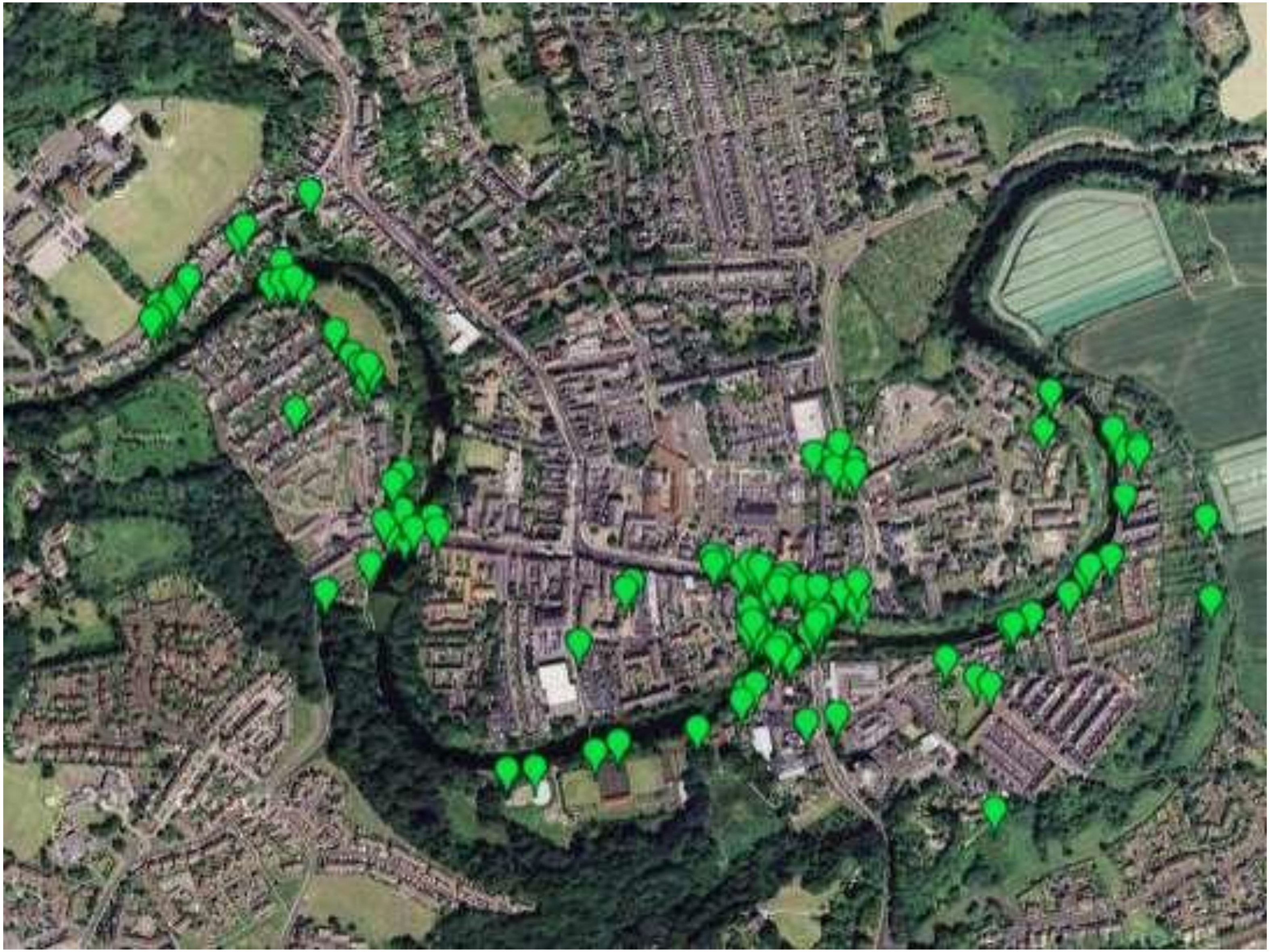
River stage in town centre (EA recorder)



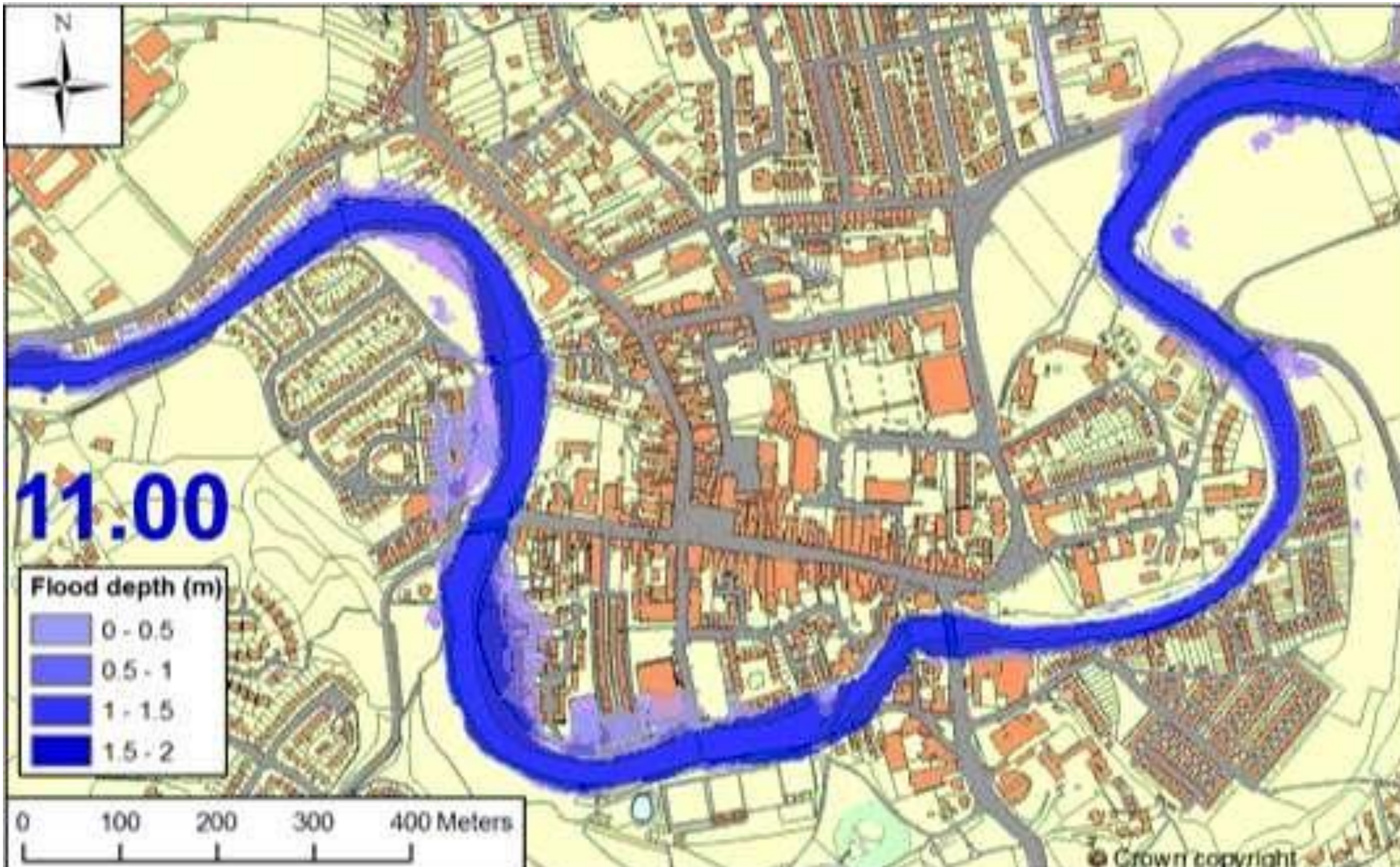
Crowd-sourcing: people wanting to tell their stories



Observation locations

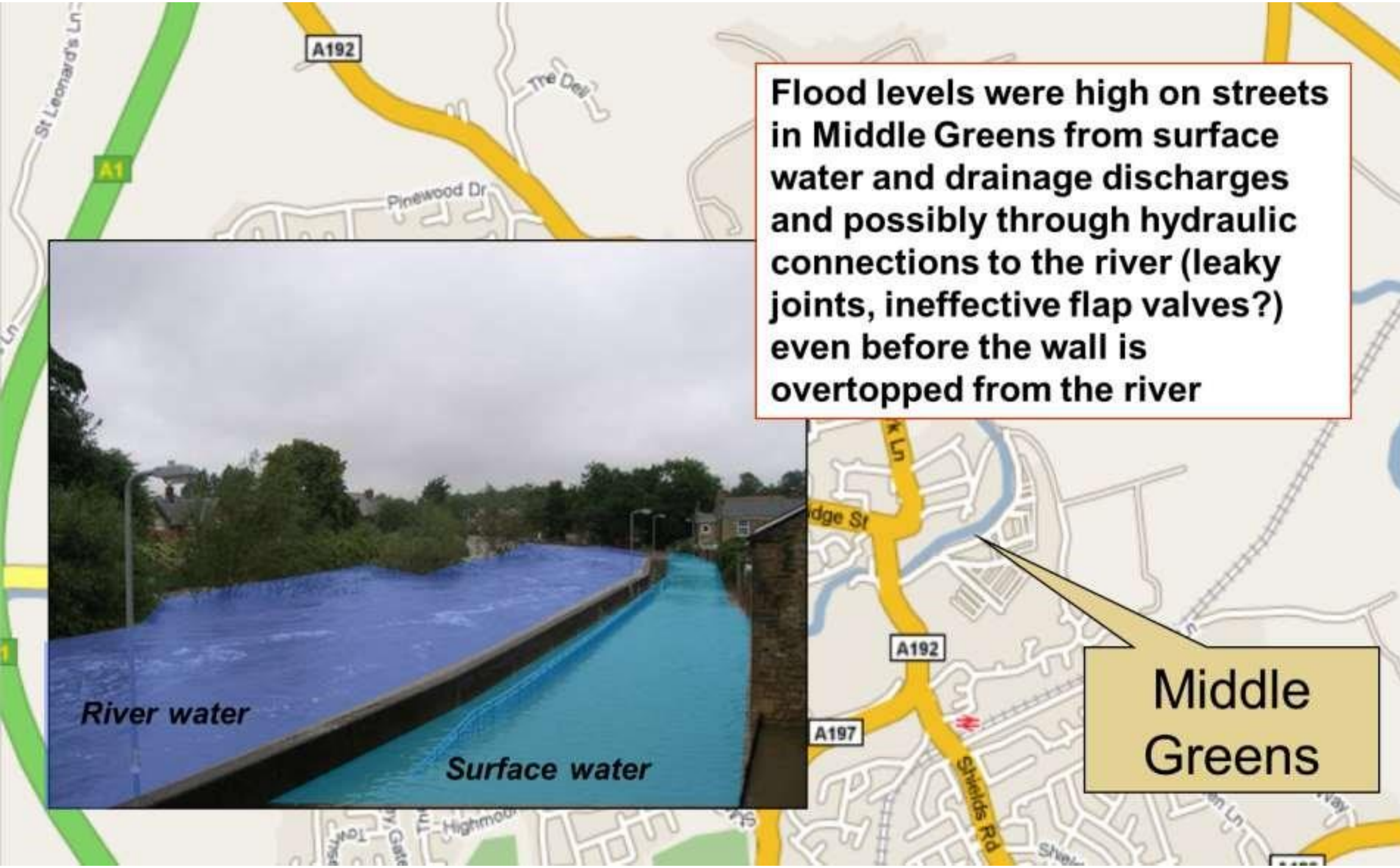


Interpolated observations (not modelled data)

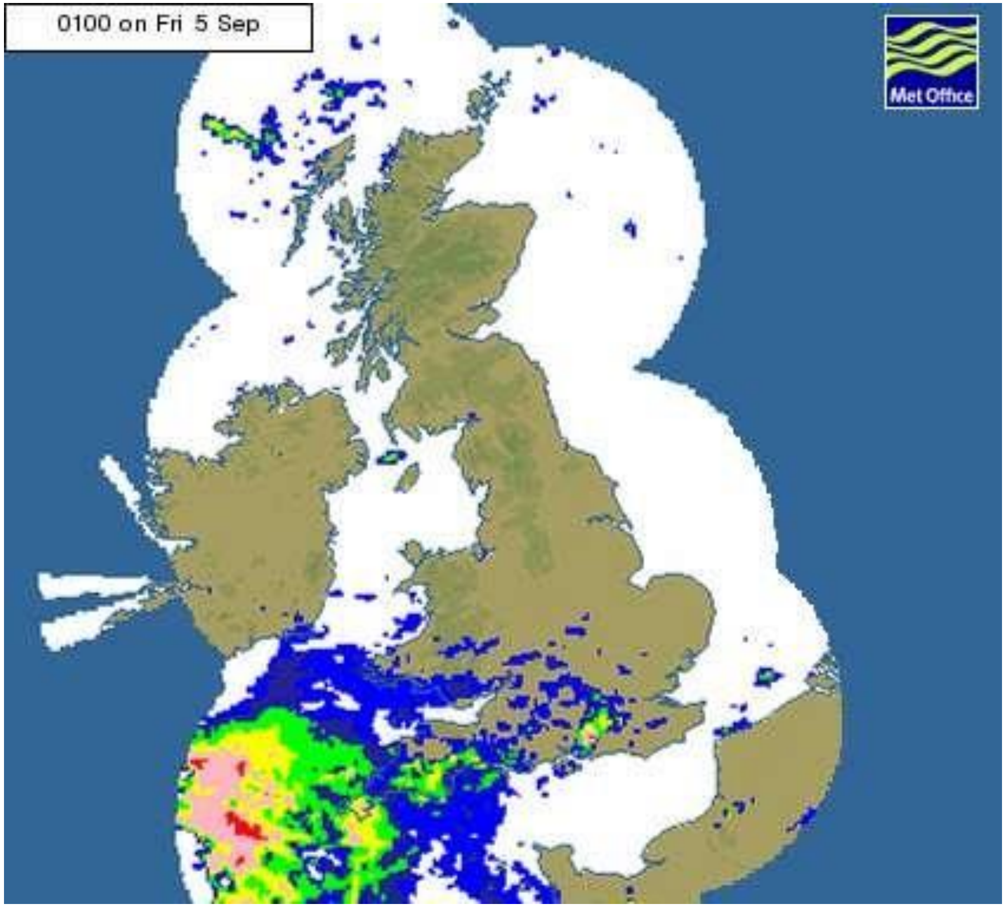


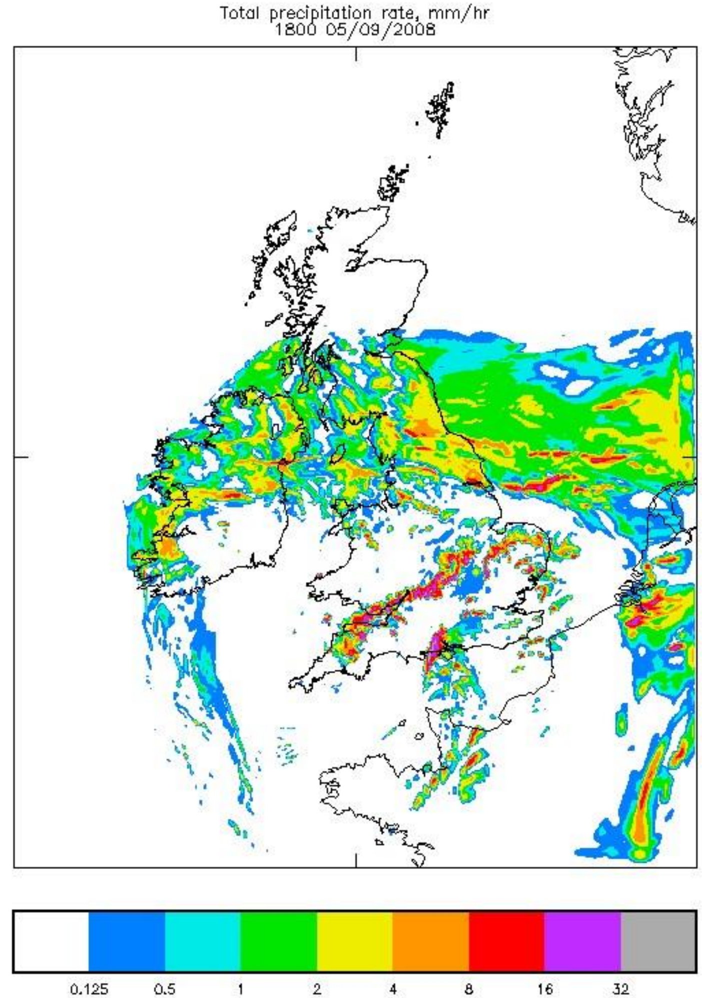
Additional value – interpretation of flood mechanisms

Additional value – interpretation of flood mechanisms

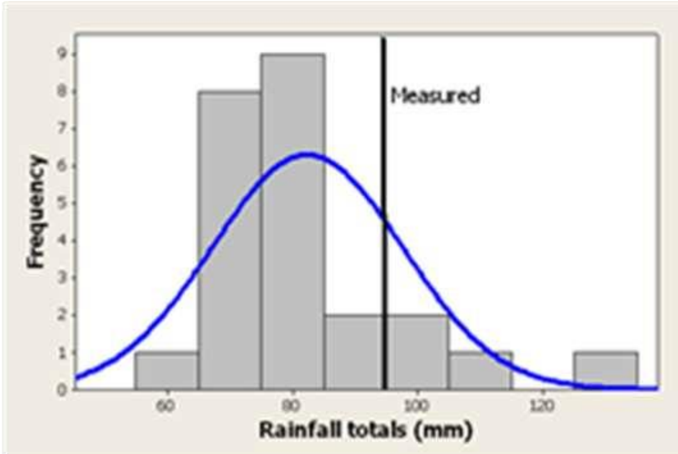


Rainfall: Radar and Mogreps 1.5km ensemble member

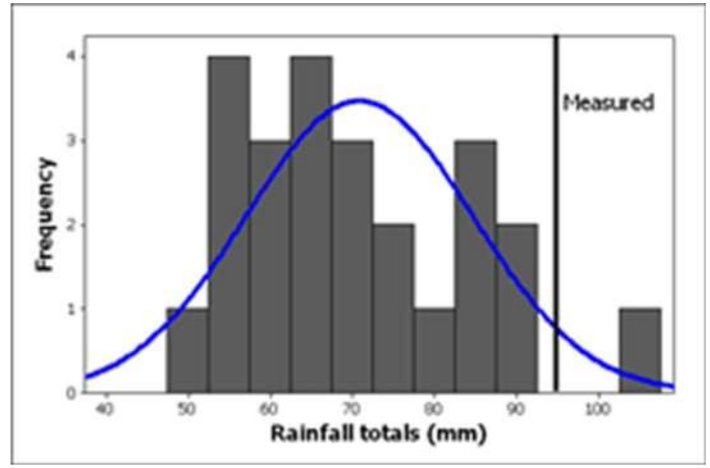




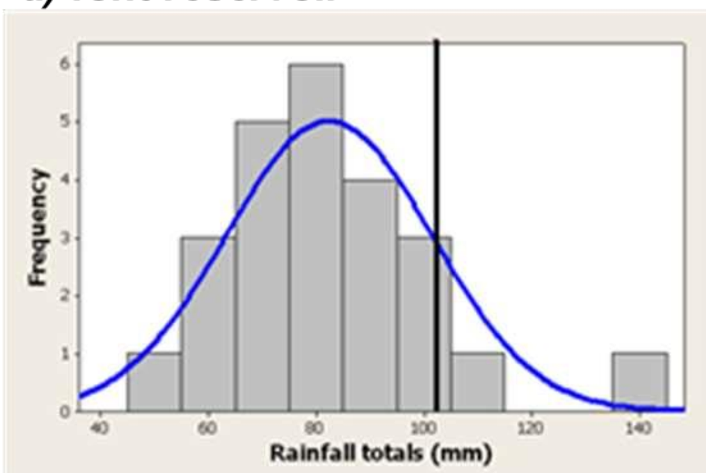
Mogreps ensemble member rainfall 30-hour totals



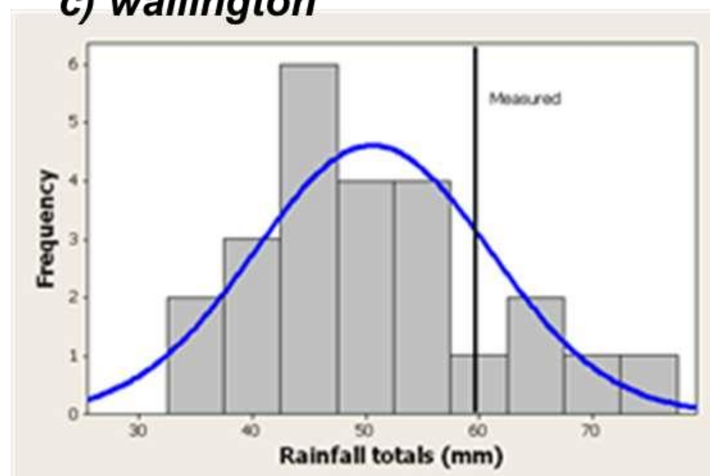
a) font reservoir



c) wallington



b) harwood

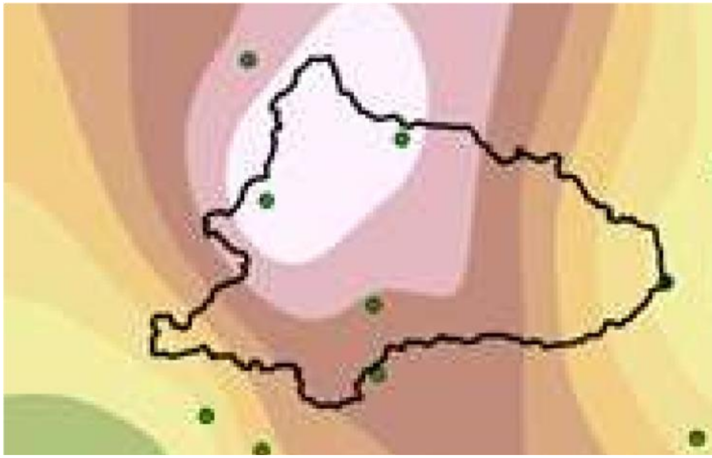


d) newbiggin

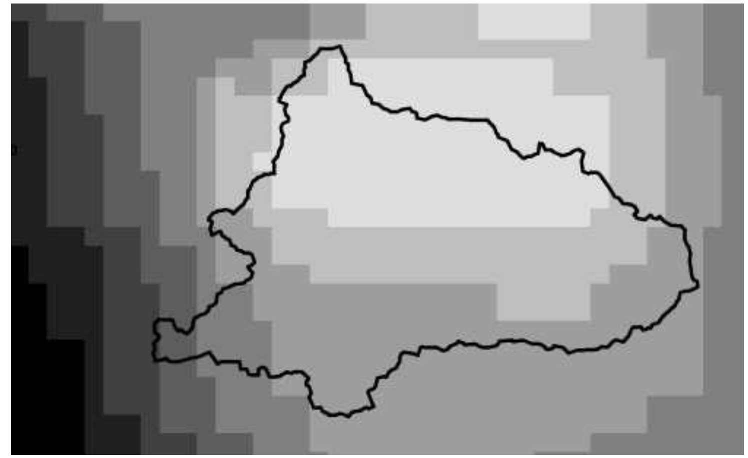
Spatial rainfall comparison

Rainfall pattern interpolated from daily raingauge totals

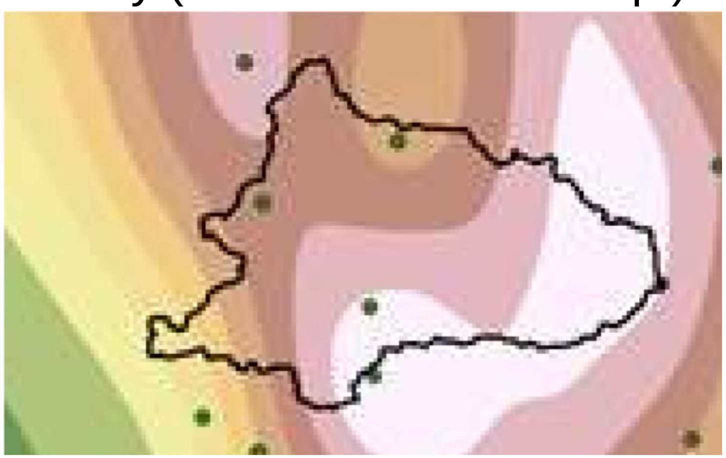
Pattern of mean 1km rainfall measured from probabilistic forecast model



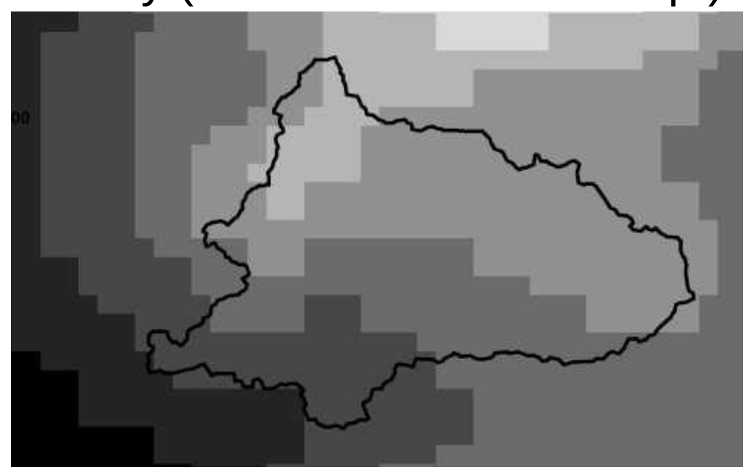
1st day (9am 5th to 9am 6th Sept)



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

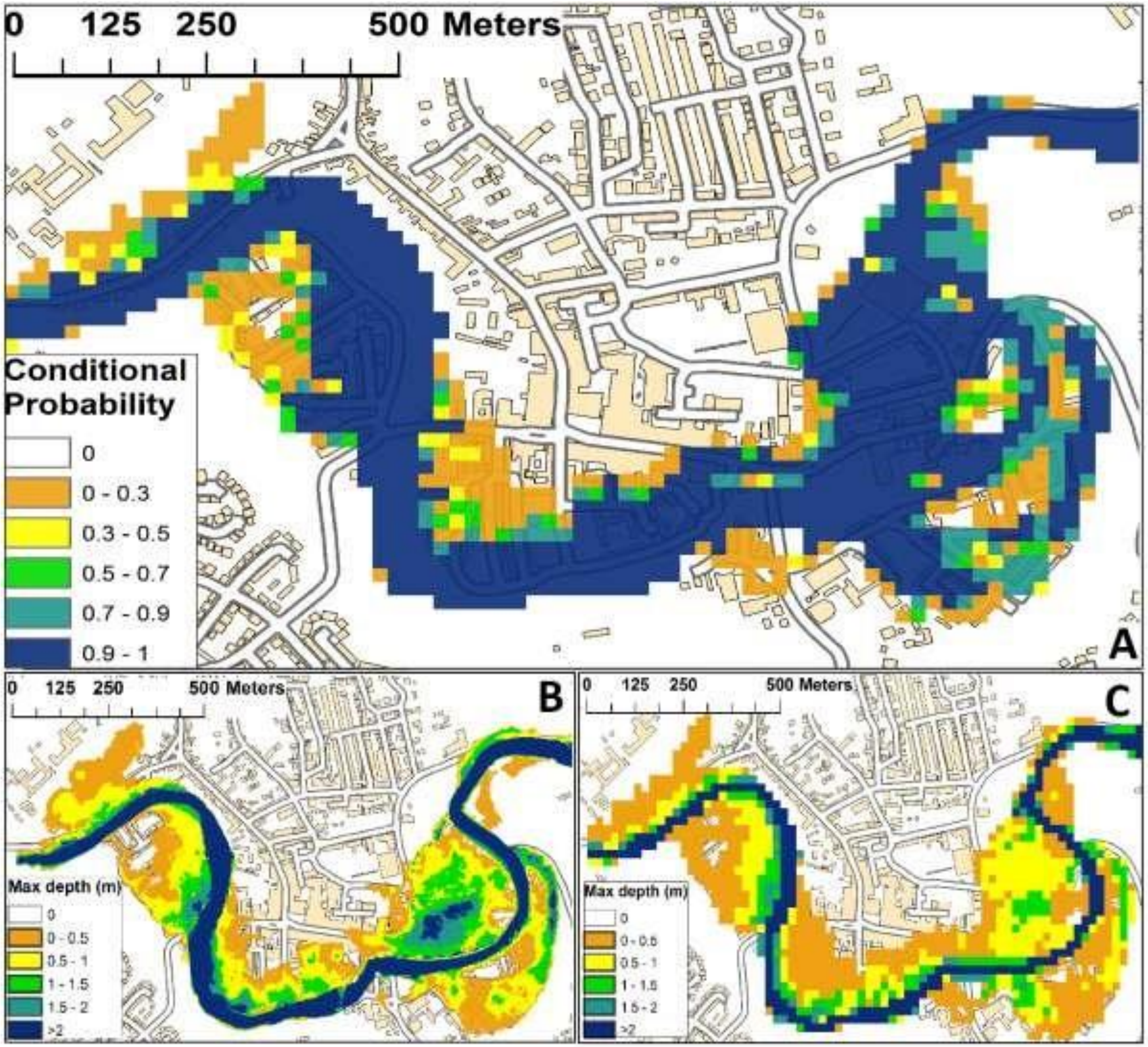


2nd day (9am 6th to 9am 7th Sept)



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

Flood modelling comparisons: extents and depths



0 125 250 500 Meters
Conditional Probability
0
0 - 0.3
0.3 - 0.5
0.5 - 0.7
0.7 - 0.9
0.9 - 1
A
0 125 250 500 Meters
B
Max depth (m)
0
0 - 0.5
0.5 - 1
1 - 1.5
1.5 - 2
>2
0 125 250 500 Meters
C
Max depth (m)
0
0 - 0.5
0.5 - 1
1 - 1.5
1.5 - 2
>2

Source

: Quinn et al. (in review)

Case study: Ethiopia

Community-based monitoring

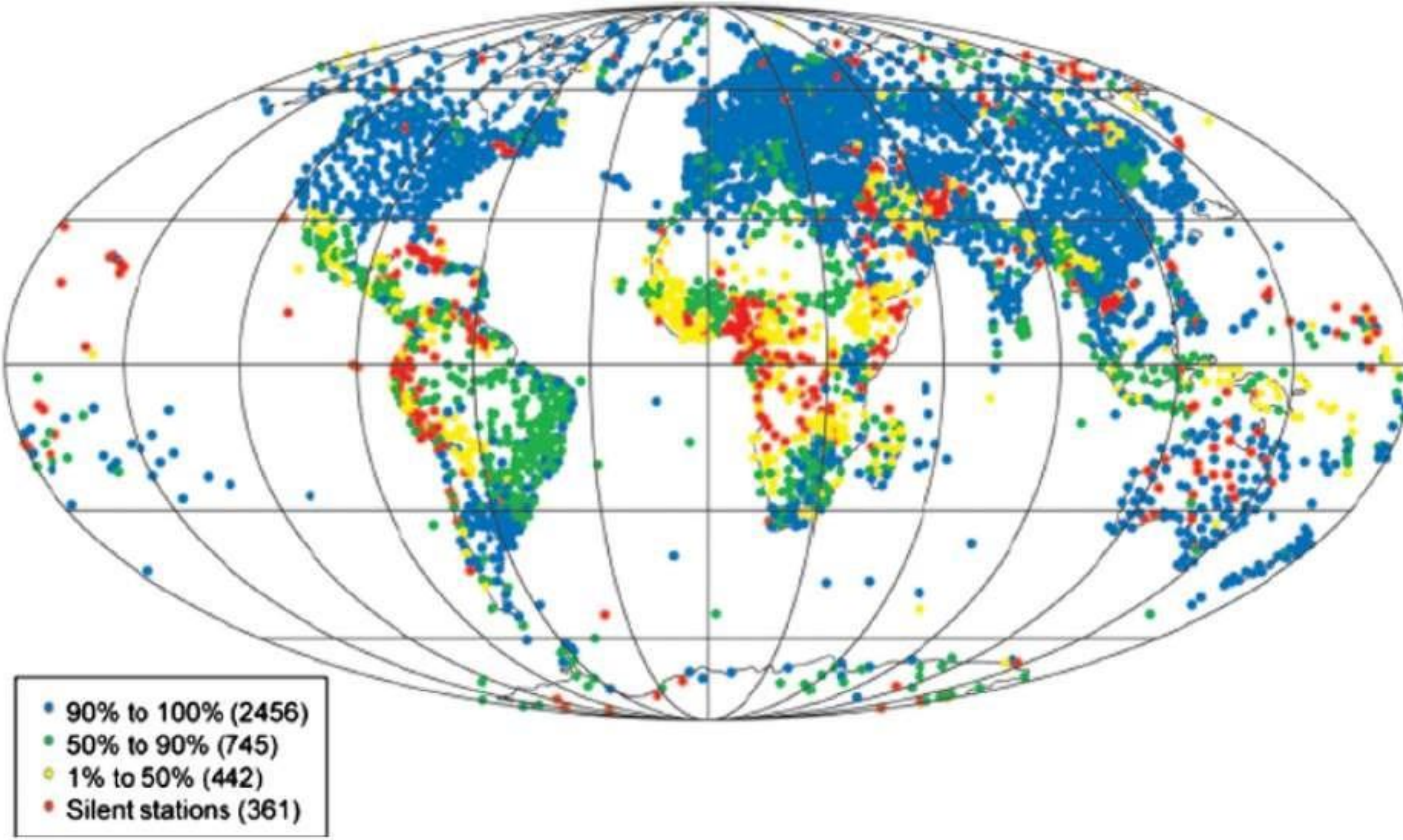
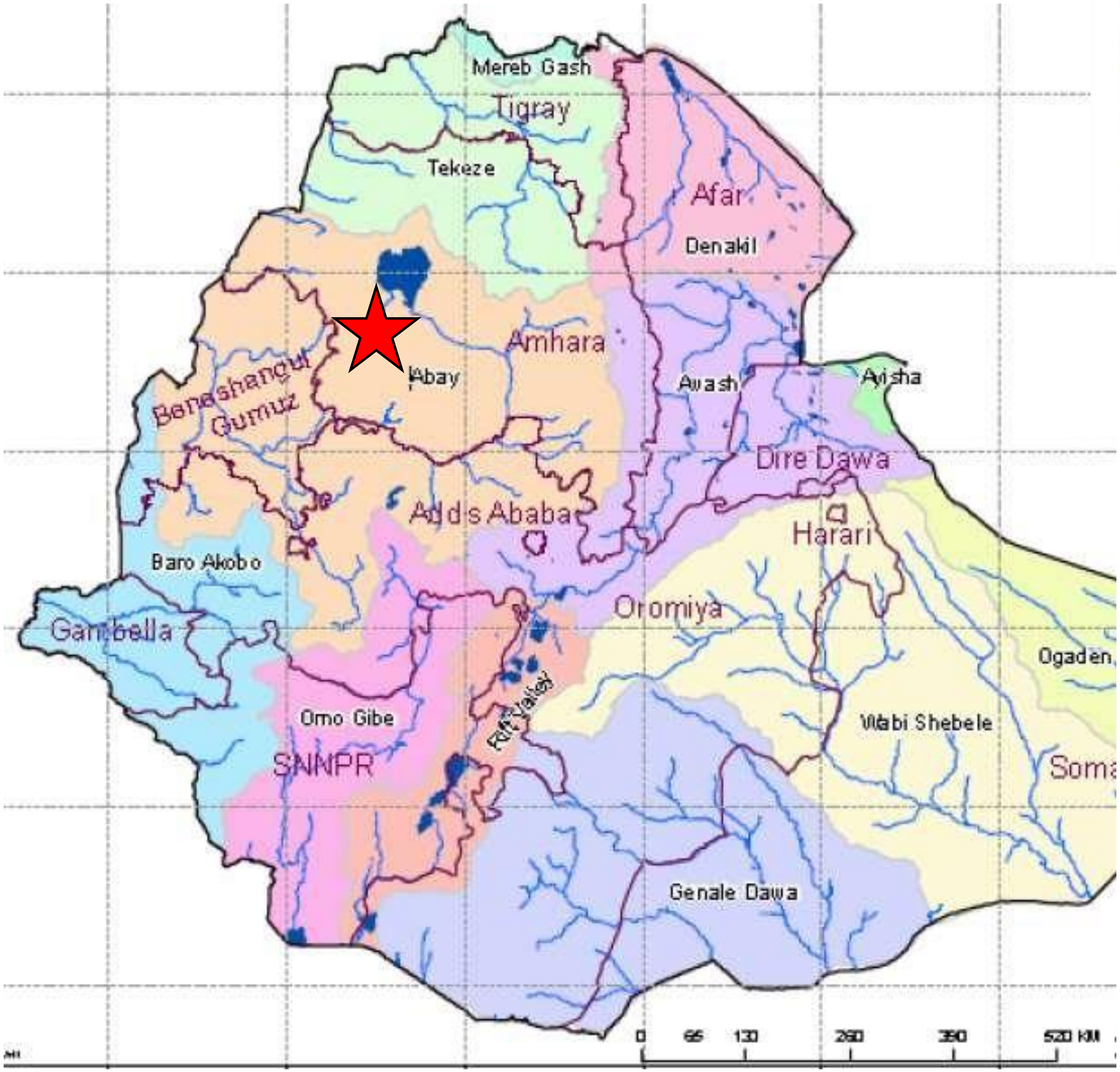


Fig. 1. The global network of World Weather Watch stations colour-coded to show reporting rates (WMO, 2003).

international example in a data-scarce region

: Walker et al. (2016)

Case study: Ethiopia



Reliability: Ethiopia study

Assessments made of community-based observations of rainfall,

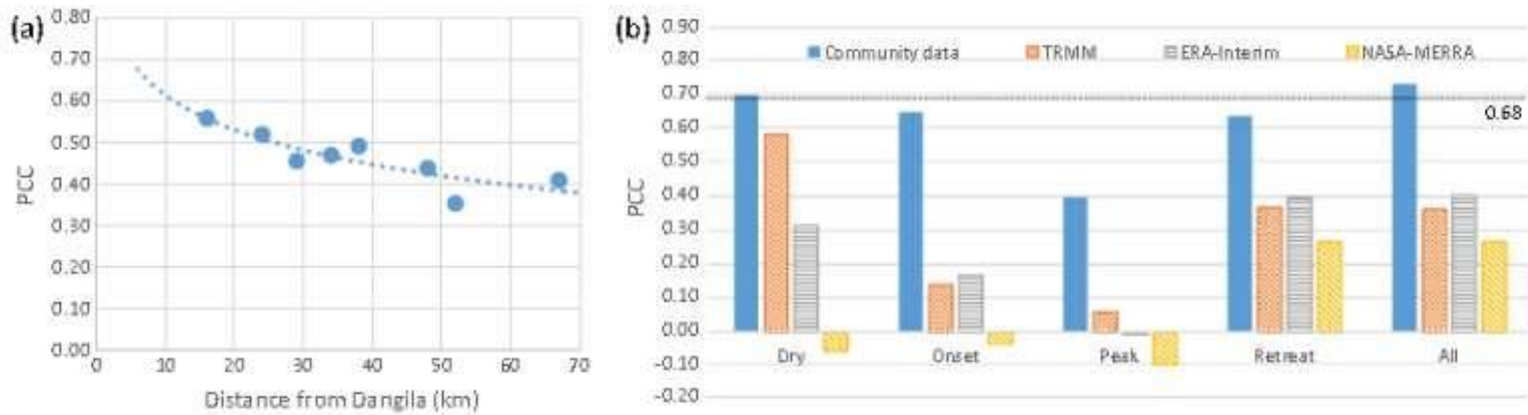


Fig. 6. Variation with distance of Pearson correlation coefficient (PCC) between daily rainfall from Dangila NMA rain gauge and other NMA rain gauges close to Dangila woreda (a). Pearson correlation coefficient (PCC) between daily rainfall from Dangila NMA rain gauge and alternative sources (b).

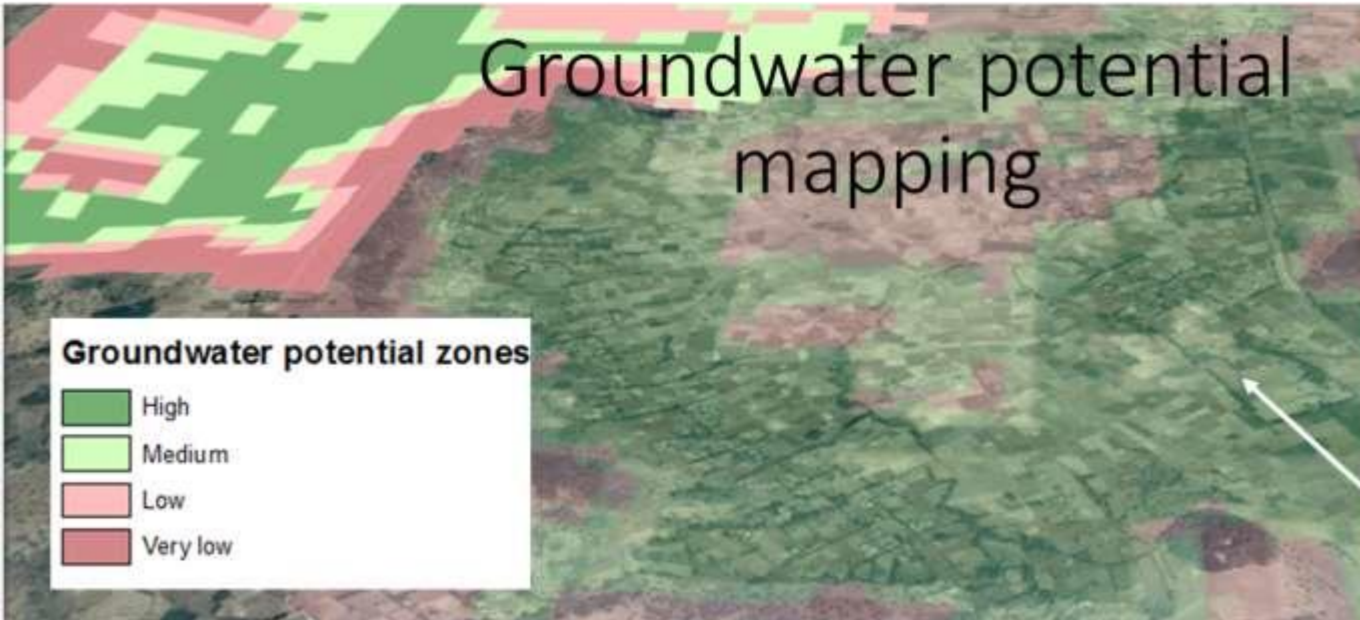
river flows, groundwater levels

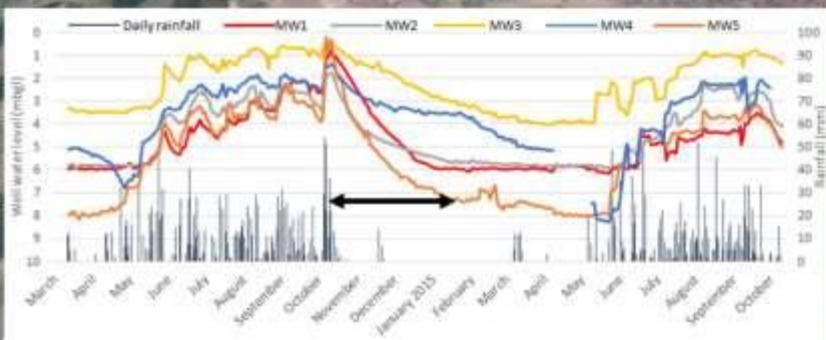
The quality of community-based rainfall observations in the Ethiopia study equals or exceeds that of formal monitoring network and gridded rainfall products, taking account of spatial variability

: Walker et al. (2016)

Source

Groundwater potential mapping





Duration of **groundwater recession** indicates potential

High potential areas match the darker more agricultural areas

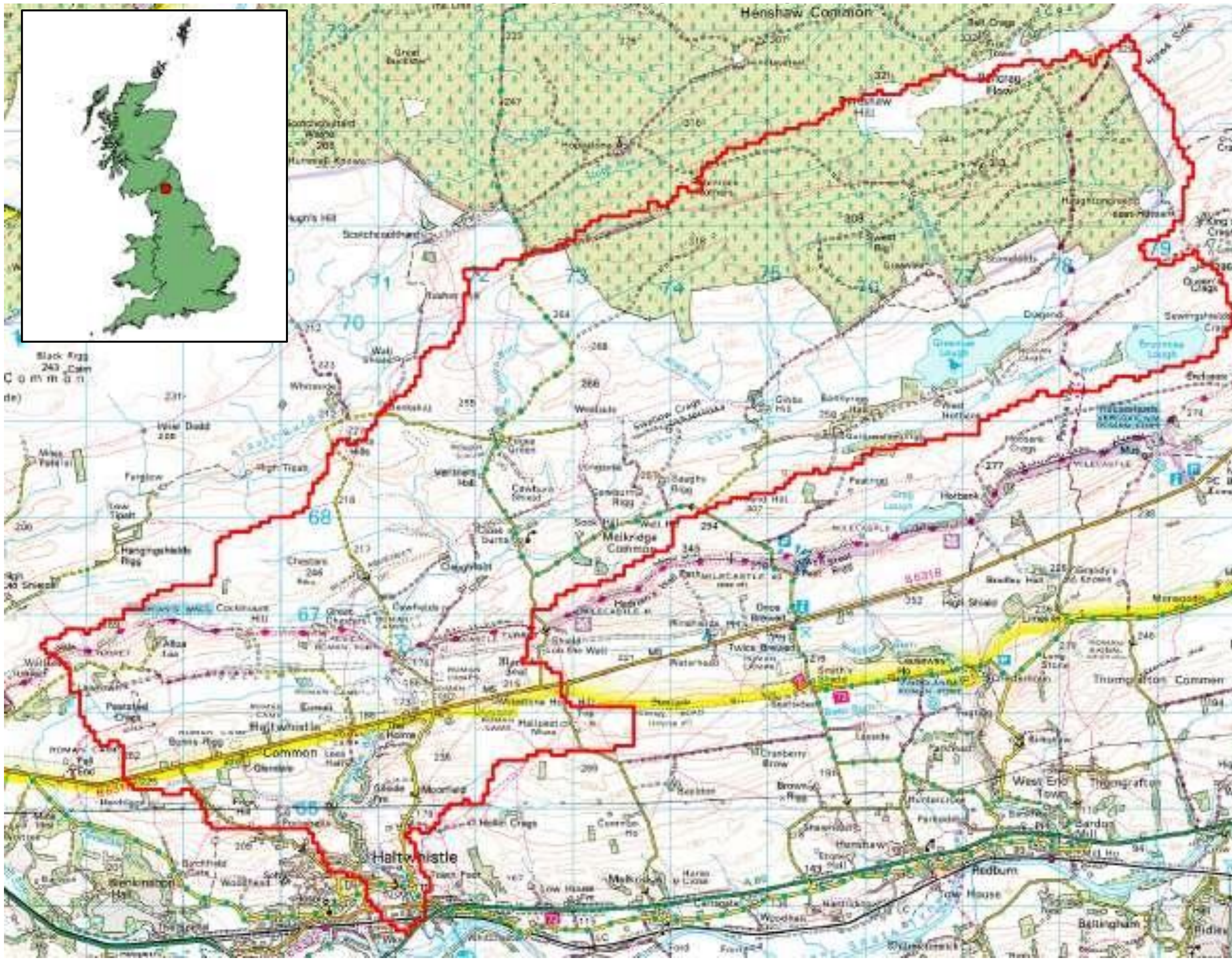
Gendered focus groups and participatory mapping

Based on high-resolution hydrological modelling

Use of community-based data for modelling

Case study: Haltwhistle

Community-based monitoring



PhD study, part of multi-purpose Catchment Restoration Fund (CRF) project

Case study: Haltwhistle Burn

Haltwhistle Burn
river watch

Simple, low-cost monitoring

Standardised data collection methods, minimal training, guidance documents

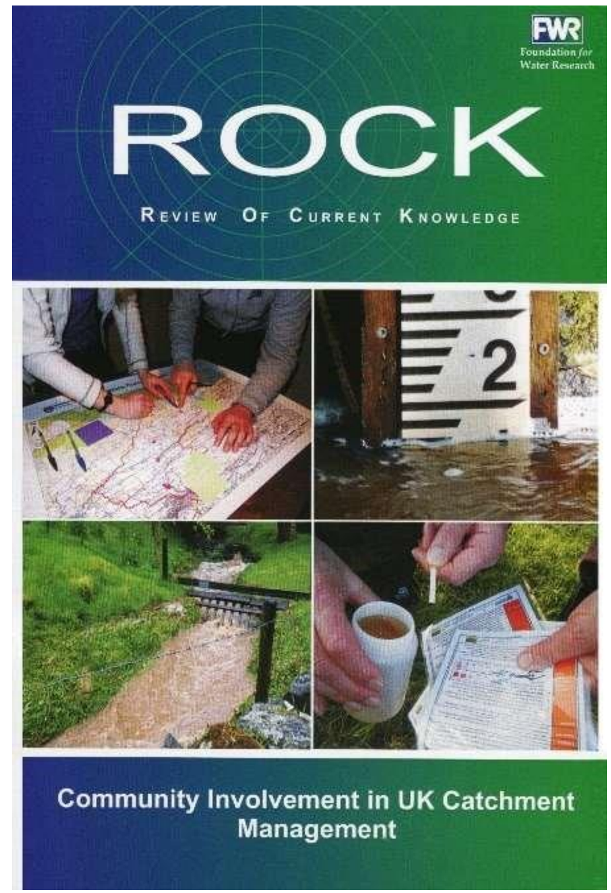
Case study: Haltwhistle Burn

Use of digital technologies

Case study: Haltwhistle Burn

Web host for data (feedback to communities)

http://research.ncl.ac.uk/haltwhistleburn/

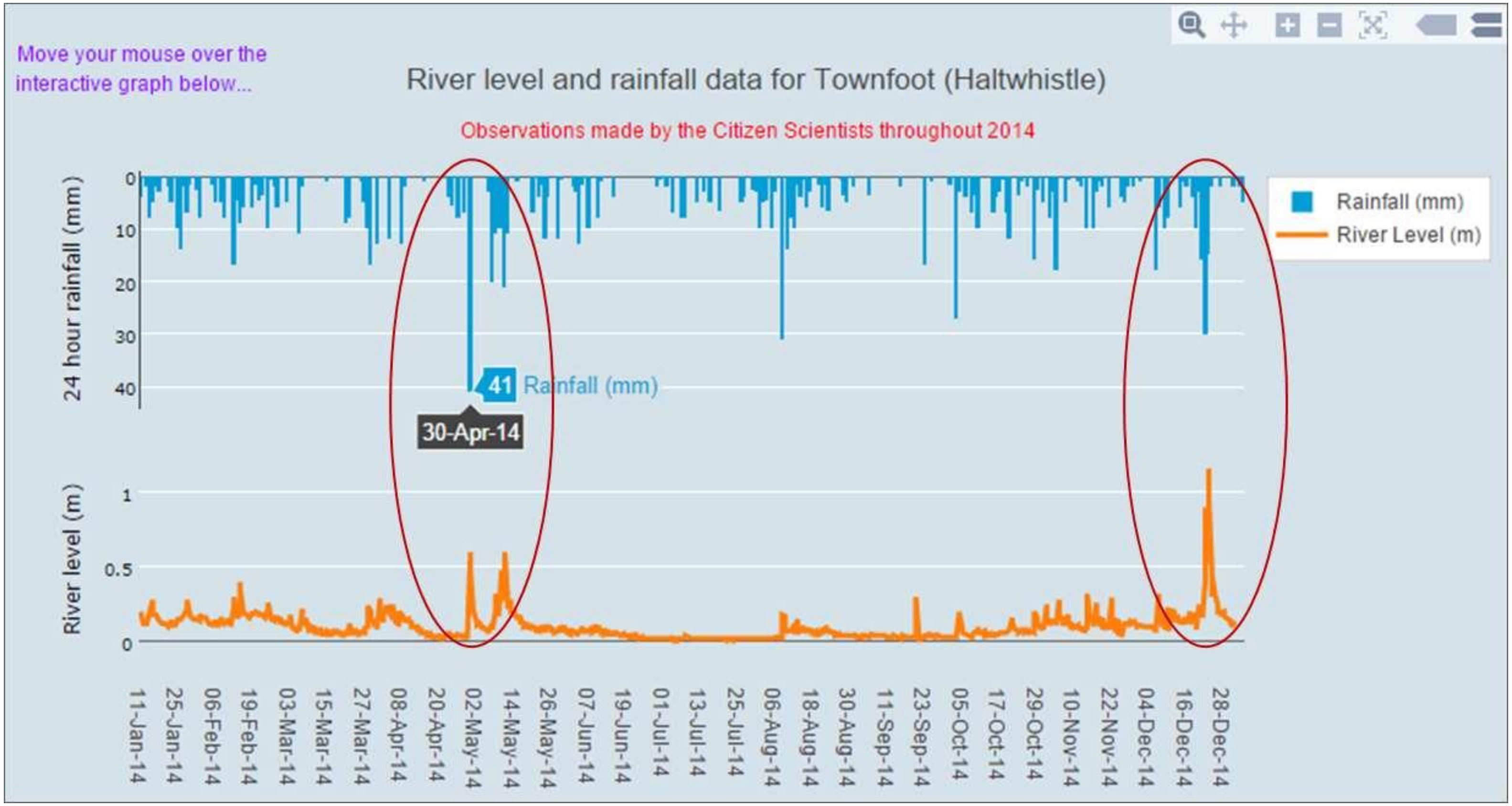


Background guidance document produced on approaches and methods (>8,000 purchases and downloads)

http://www.fwr.org/Catchment/frr0021.pdf

Case study: Haltwhistle Burn

Knowledge exchange ⇨ Collect data ⇨ Submit ⇨ Share ⇨ Feedback



Source

: Starkey (2017b)

Source: Starkey (2017a)

Data use: qualitative understanding

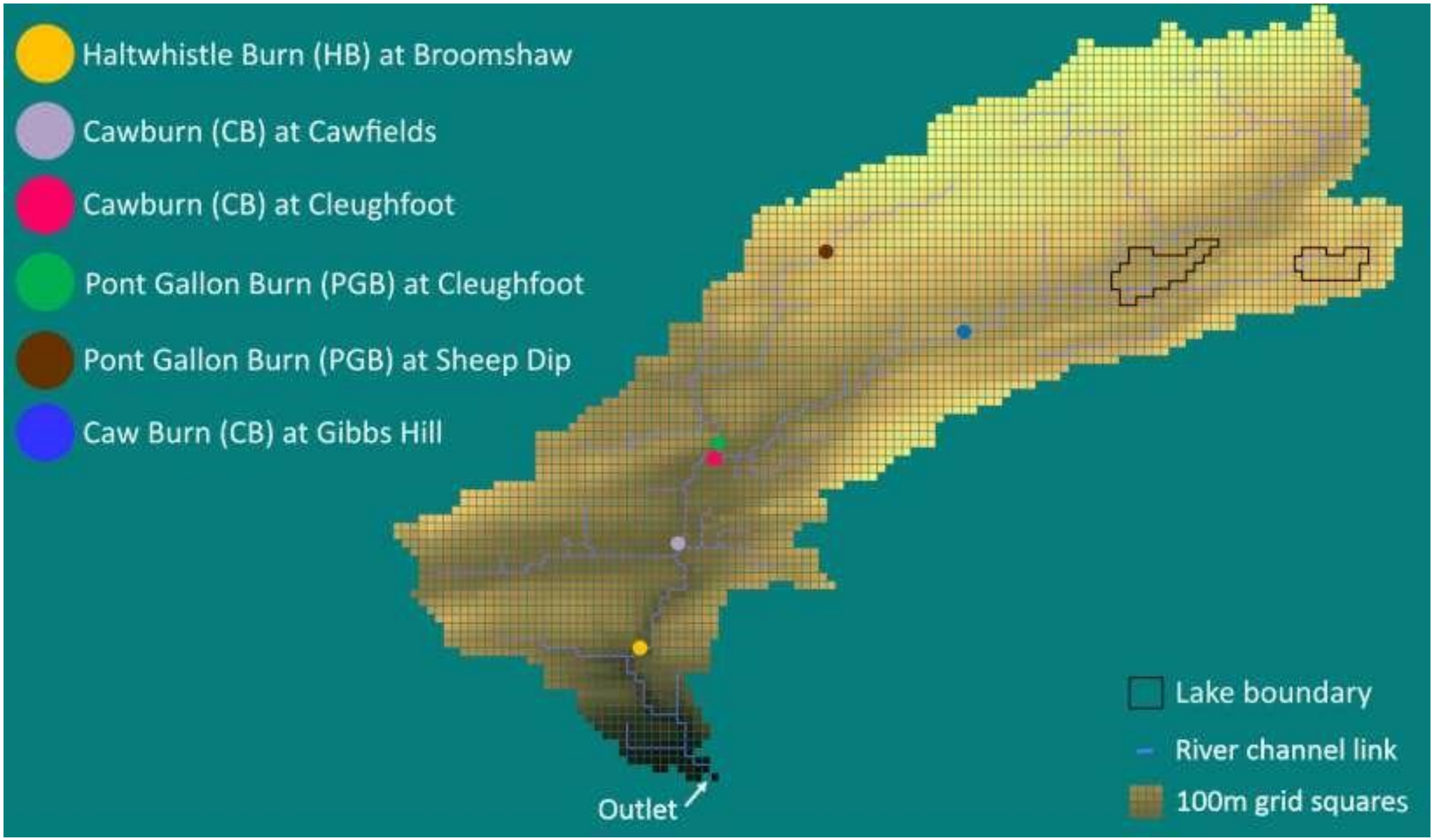
Source

Community-based timeline for 30th April 2014 event

Source: Starkey (2017a)

: Starkey (2017b)

Data use in hydrological modelling



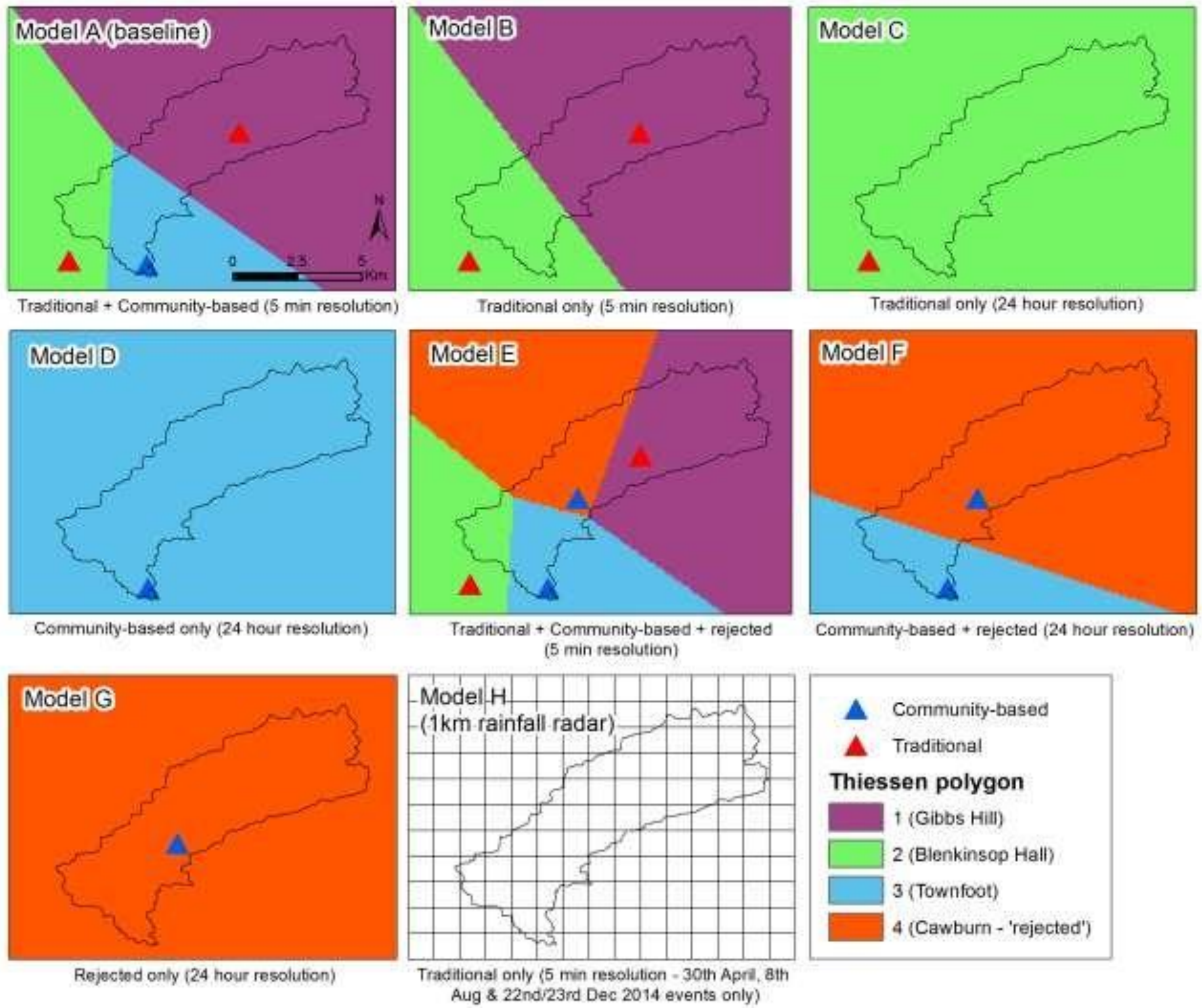
Source

Shetran catchment model setup with river gauge locations

Source: Starkey (2017a)

Data use in hydrological modelling

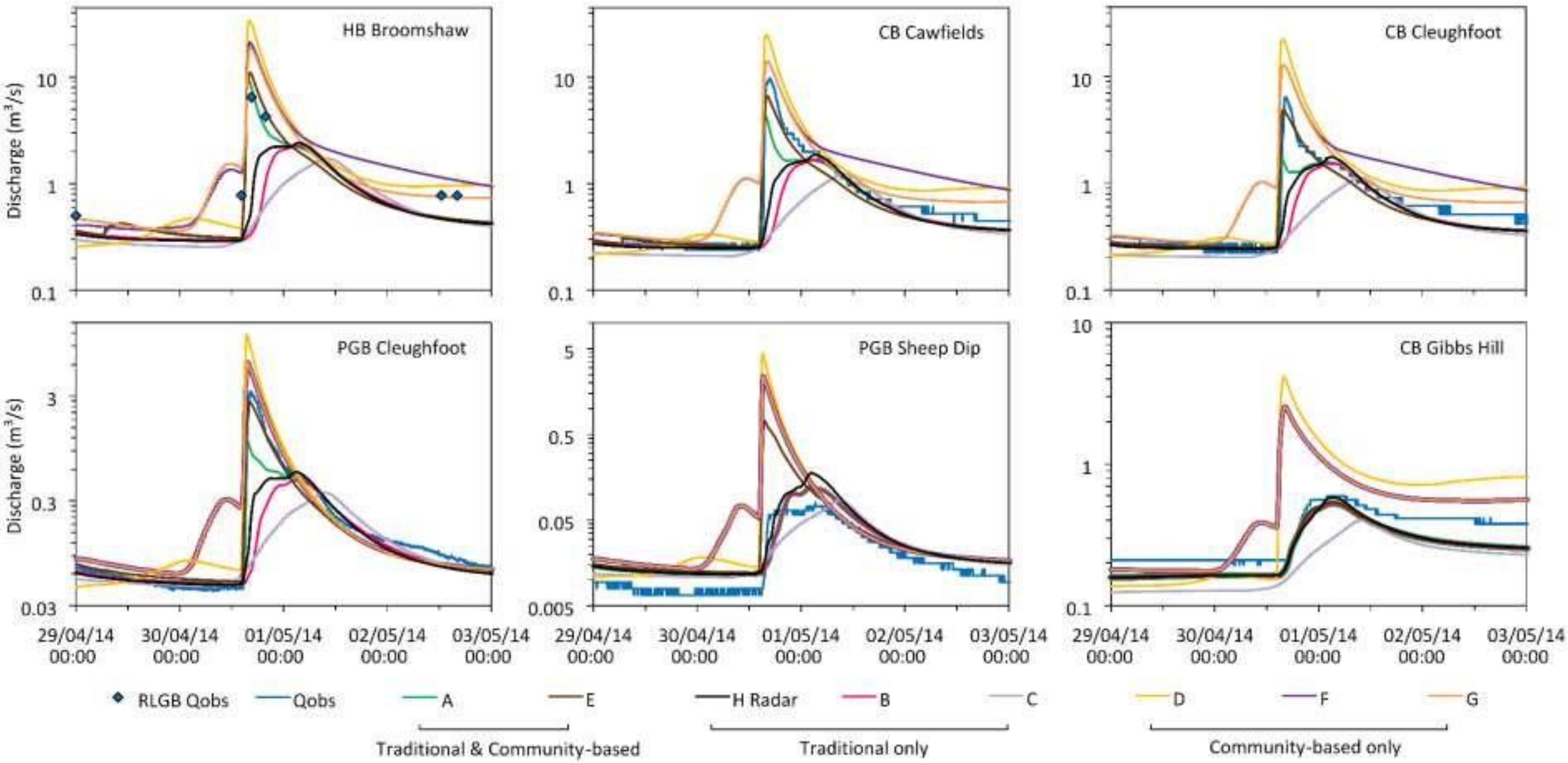
Source: Starkey (2017a)



Source: Starkey (2017a)

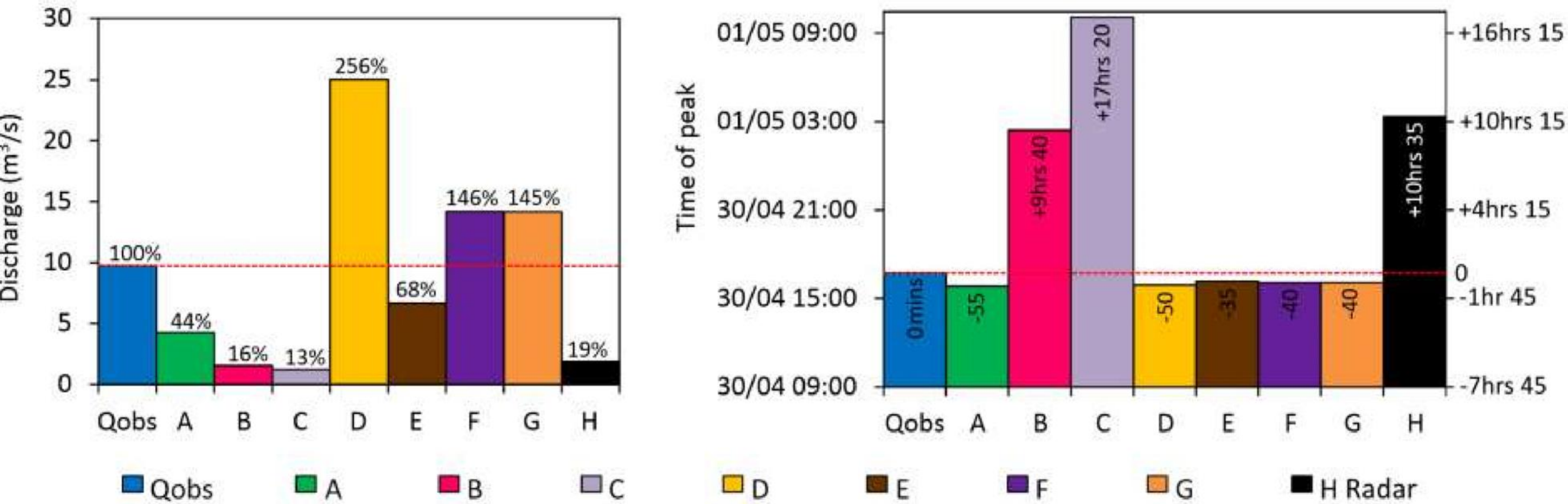
Modelled scenarios (one-at-a-time excluded data)

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

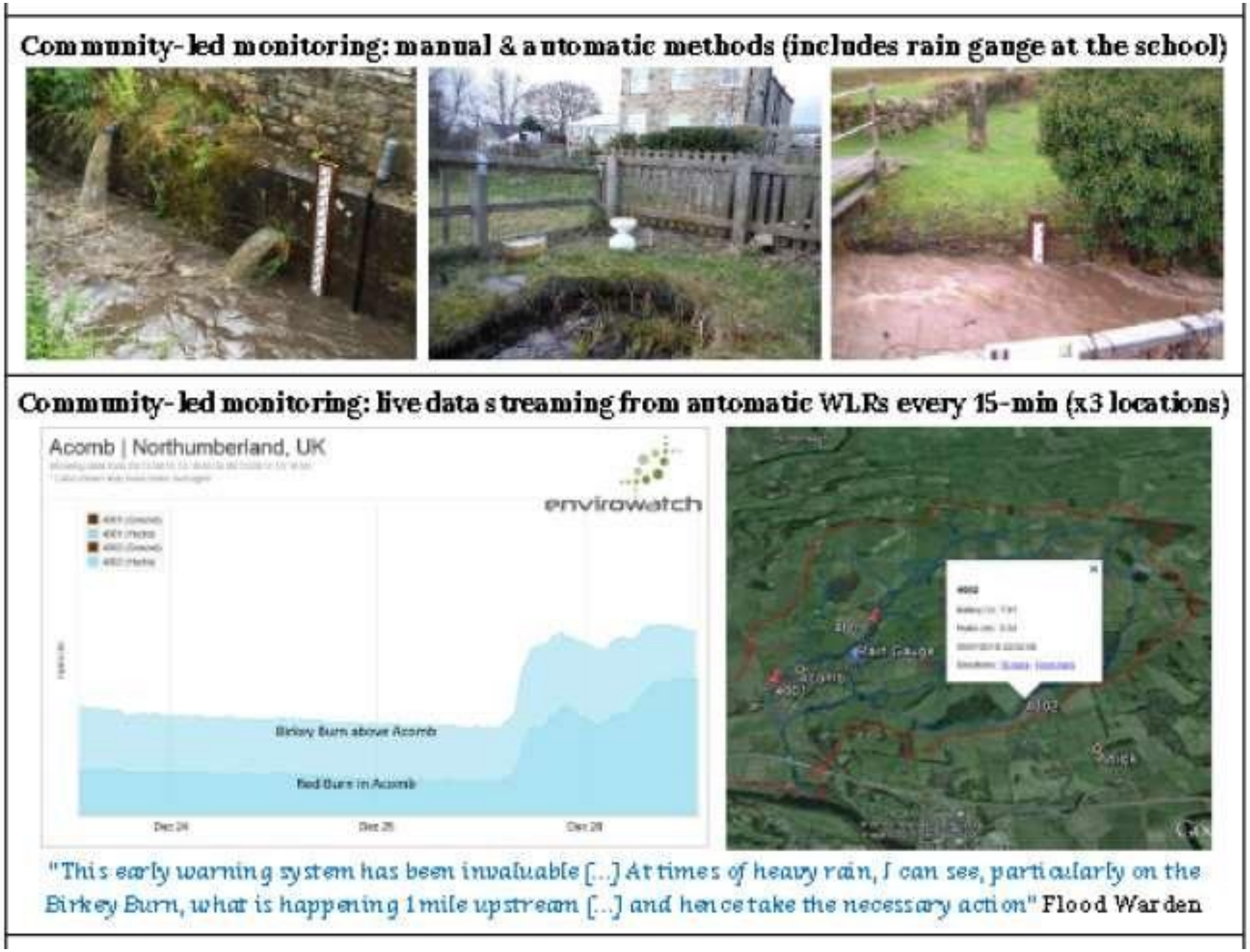
Some survey findings - community willingness to participate:

- 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



Source: Starkey (2017b)

Potential for use of community-based observations


Defra
Environment Agency
Natural England
Lead Local Flood Authority
Water & Sewerage companies
Highways
District Councils
Internal Drainage Boards
River Trusts
Wildlife Trusts
National Parks
Researchers
River Watch Groups
Flood Warden / Action Groups
Other environmental groups
Local land owners
People affected by catchment issues
Wider community
'Drivers'
'Professionals'
Local community
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:

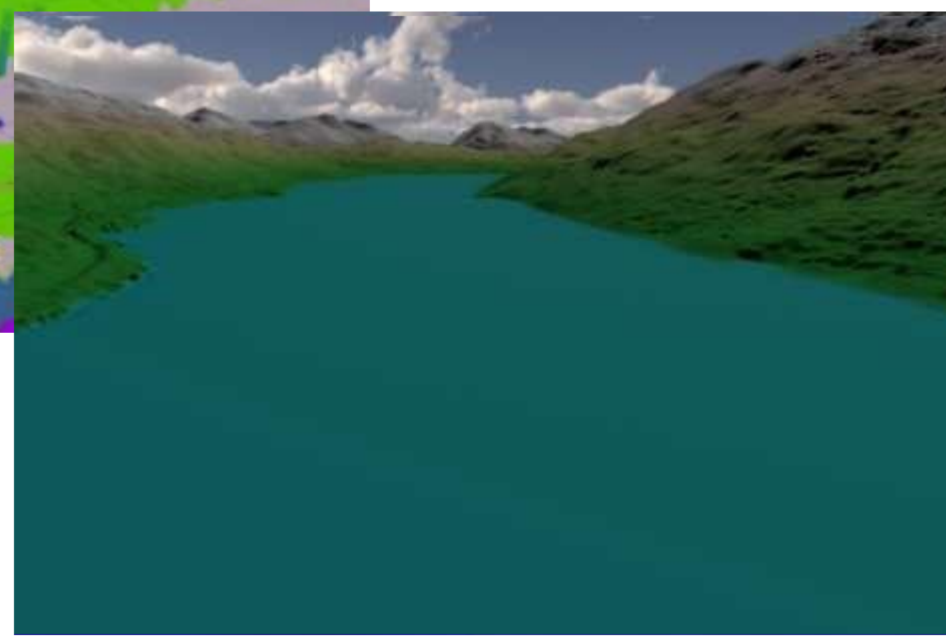
- Developing new low-cost technologies (eg camera-based direct river flow measurements)
- Data collection with a purpose

- 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 communitybased data

Visualisation: Geovisionary software

Scalable overlays of multiple datasets including near-real-time data streaming



Realistic visualisations of data and model scenarios

Thank you for listening
Any questions?

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

References

Archer, D.R., Parkin, G. and Fowler, H.J. (2016). Assessing long term flash flooding frequency using historical information. Hydrology Research, DOI: 10.2166/nh.2016.031.

Parkin, G. (2010). The September 2008 Morpeth flood: Information gathering for dynamic flood reconstruction. School of Engineering, Newcastle University, Newcastle upon Tyne, UK.

Quinn, N., Coxon, G., Dunne, T., Freer, J., Neal, J., Bates, P., Sampson, C., Smith, A., Parkin, G. (in review). Development of a flexible coupled hydrological – hydrodynamic model for catchment flood inundation modelling.

Starkey, E. and Parkin, G. (2015). Community Involvement in UK Catchment Management, Review of Current Knowledge (ROCK) report, Foundation for Water Research, http://www.fwr.org/Catchment/frr0021.pdf

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., Gibson, C. (2017a). Demonstrating the value of community-based (‘citizen science’) observations for catchment modelling and characterisation. Journal of Hydrology, doi: http://doi.org/10.1016/j.jhydrol.2017.03.019

Starkey, E. (2017b). Community-based ('citizen science') monitoring for catchment characterisation, modelling and management. PhD Thesis, School of Engineering, Newcastle University.

Walker D, Forsythe N, Parkin G, Gowing J (2016). Filling the observational void; scientific value and quantitative validation of hydro-meteorological data from a community-based monitoring programme in highland Ethiopia. Journal of Hydrology 538, 713–725.