Enhancing Crowdsourced Classification on Human Settlements Utilizing Logistic Regression Aggregation and Intrinsic Context Factors

Benjamin Herfort

GIScience Research Group

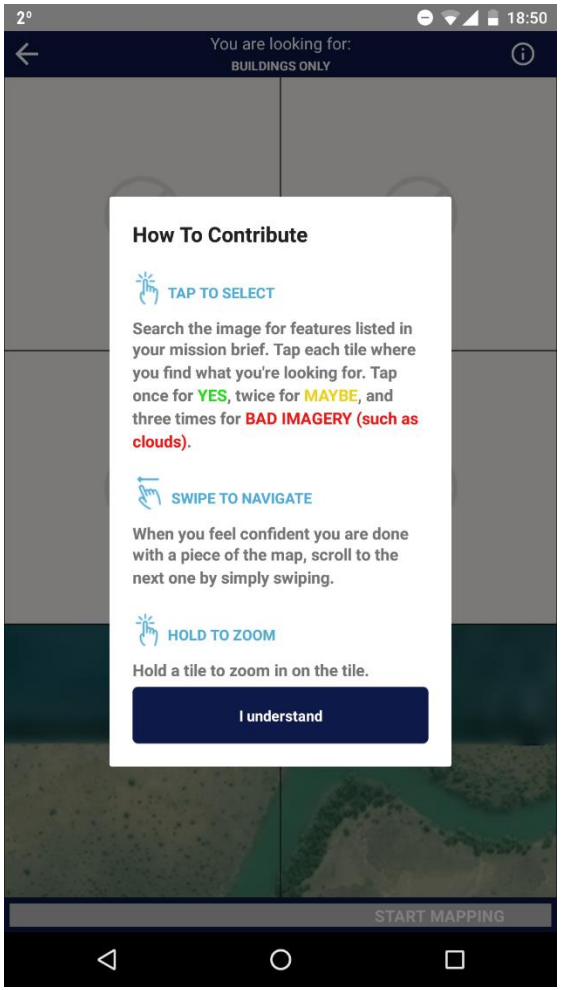
Heidelberg University

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

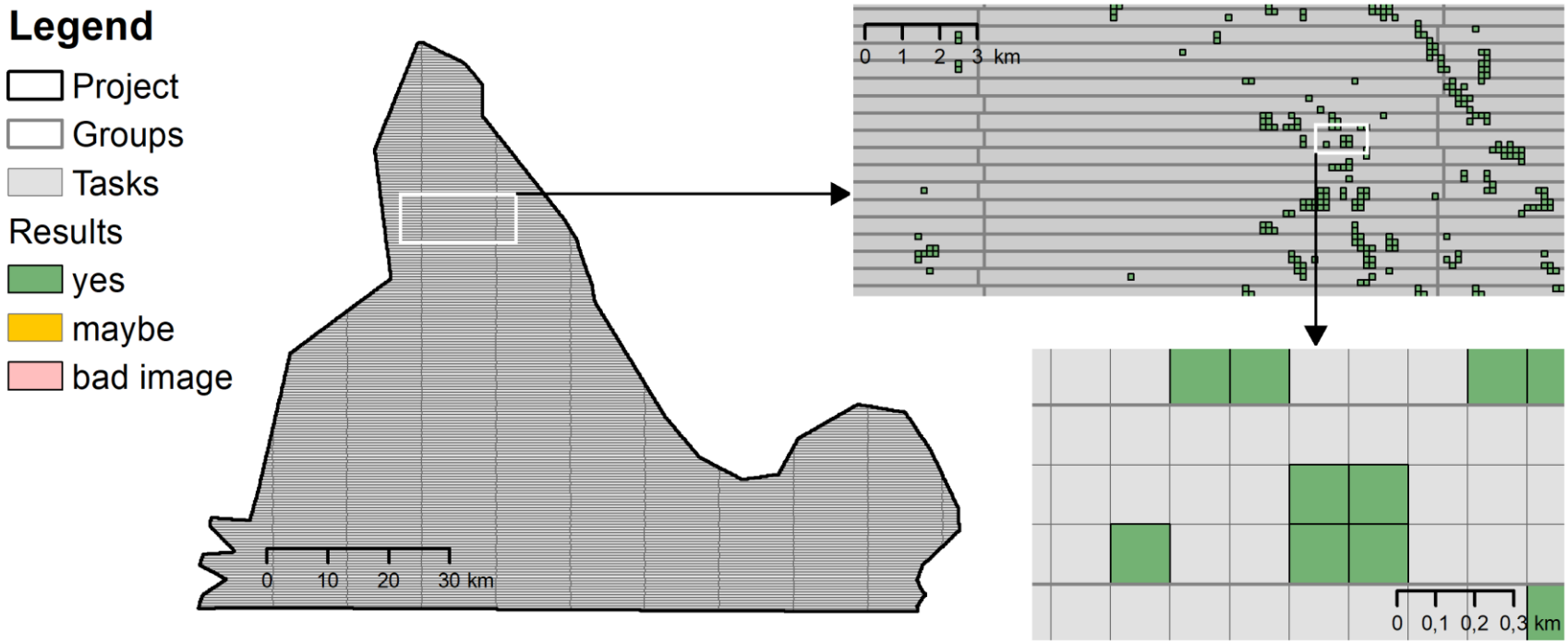
HeiGIT | HEIDELBERG INSTITUTE FOR GEOINFORMATION TECHNOLOGY

GIScience

The MapSwipe App



MapSwipe Data Model: Project, Groups, Tasks, Results



Research Question

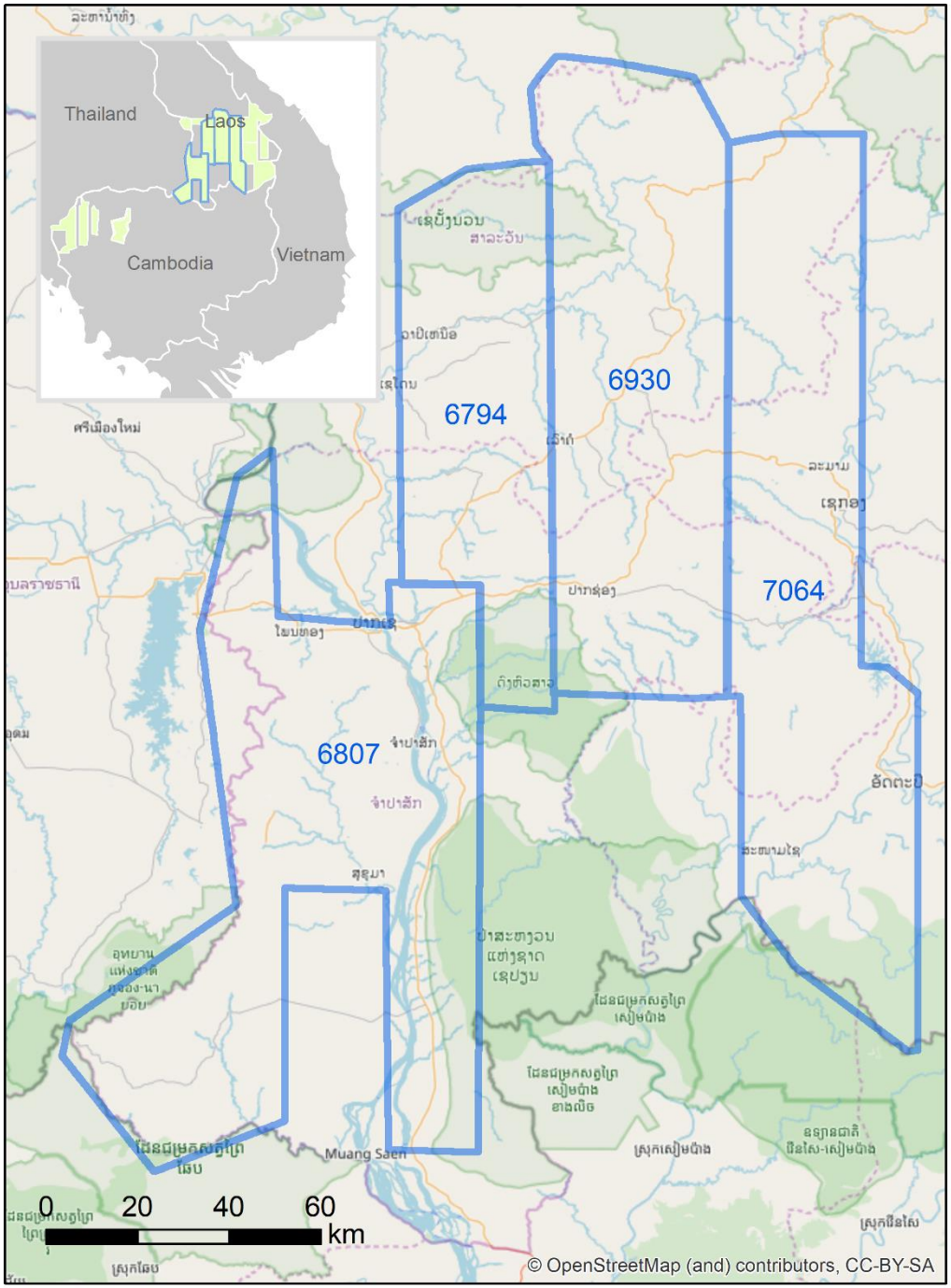
*To what degree can **automated classifiers** considering **intrinsic context factors** (user agreement, user characteristics and spatial characteristics) **enhance data quality** of aggregated **crowdsourced classifications?***

Dataset

- Tasks: 941,589
- Results: 3,275,380
- Users: 1,534

Reference:

- Building Footprints from OpenStreetMap (area has been validated through Clinton Health Initiative)



Agreement

Equation 1: Scott's Pi

$$P_i = \frac{1}{n * (n-1)} * \sum_{J=1}^{k} n_{ij}^2 - n_{ij}$$

Equation 2 Building Classification Index

$$\mathrm{BI} = \frac{n_{building}}{n}$$

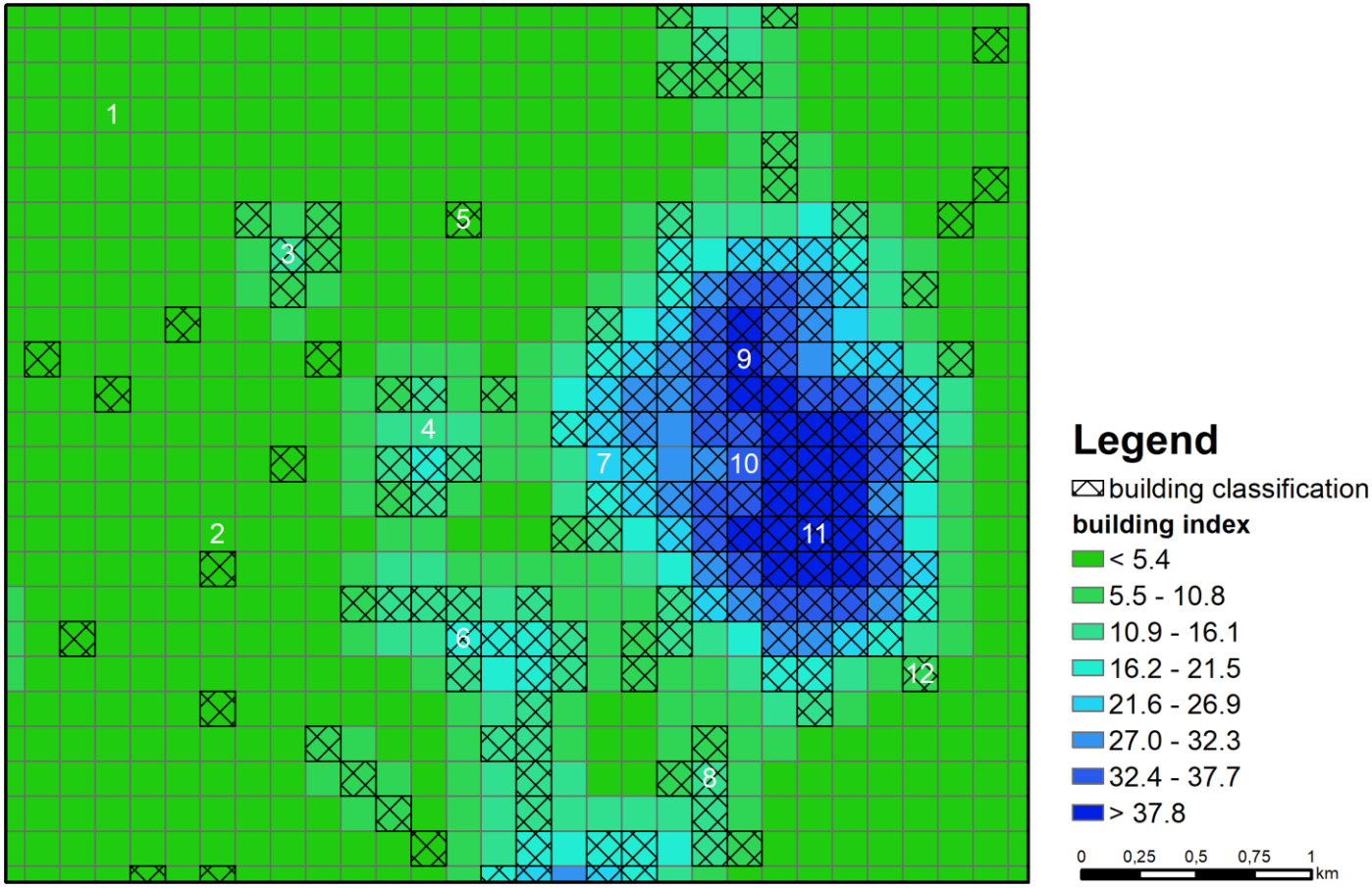
Equation 3 No Building Classification Index

$$NBI = \frac{n_{no\ building}}{n}$$

User Characteristics

User Activity	**User Performance**
number of contributions	accuracy
number of completed groups	(no building, building, bad image) sensitivity
number of different projects	(no building, building, bad image) precision
number of mapping sessions	(no building, building, bad image) f1 score.

Spatial Characteristics



Logistic Regression

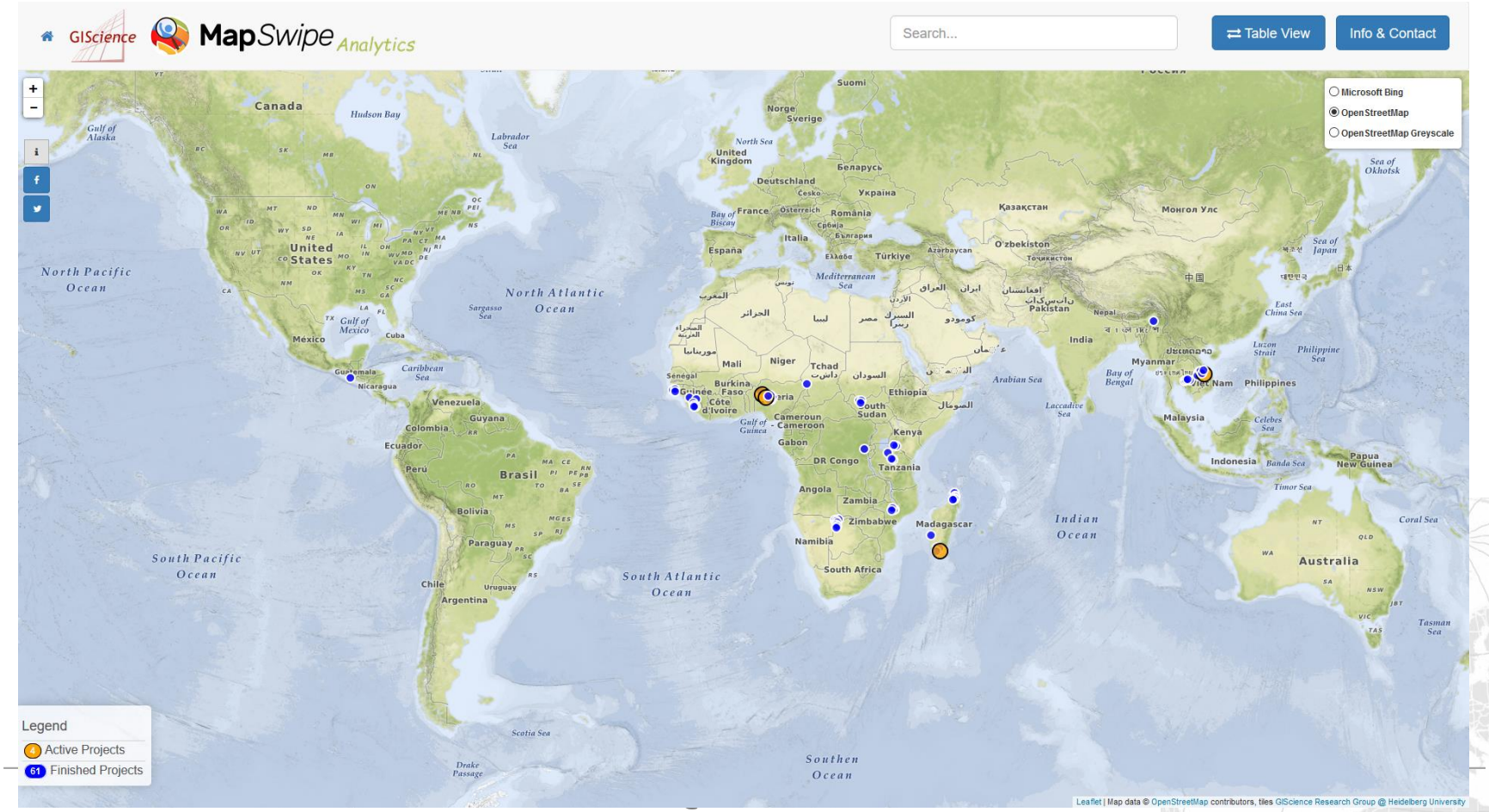
	Coeff.	StdEr	Sign.	Odds
Building Index	78.735	0.027	<0.005	2626.8337
Average Accuracy (No Building Results)	-84.139	0.076	<0.005	0.0002
Average Building Precision (Building Results)	62.469	0.051	<0.005	516.406
Average Bad Image Precision (Bad Image Results)	-14.723	0.059	<0.005	0.2294
No Building Class Density	-0.0148	0.001	<0.005	0.9853
Building Class Density	0.0919	0.001	<0.005	1.0962
Bad Image Class Density	-0.2360	0.005	<0.005	0.7897

Classification Performance

	Soft Majority Agreement	Logit Classifier
Overall Accuracy	0.9693	0.9906
Building Sensitivity	0.7838	0.9325
Building Precision	0.9115	0.9770
Building F1 Score	0.8428	0.9543

MapSwipeAnalytics

mapswipe.heigit.org



Thank You.

herfort@uni-heidelberg.de

HeiGIT | HEIDELBERG INSTITUTE FOR GEOINFORMATION TECHNOLOGY