

Small-area estimations from survey data for high-resolution maps of urban flood risk perception and evacuation behavior

Authors

Samuel Rufat, Peter D. Howe

Abstract

“Behavior-blind” risk assessments, mapping, and policy do not account for individual responses to risks, due to challenges in collecting accurate information at scales relevant to decision-making. There is useful spatial information in social survey data that is sometimes analyzed for spatial patterns despite potential biases. This paper explores whether risk perception and adaptive behavior can be inferred from census and exposure data with a specifically designed survey. An underlying question is what precautions surveys should take before mapping the results. We find that an hybrid multilevel regression and (synthetic) poststratification (MRP-MRSP) model can facilitate the transition from individual survey data to small-area estimations at different scales, including 200-meter grid cells. We demonstrate this model using municipal-level survey data collected in the Paris region, France. We find that model accuracy is not decreased at finer scales provided there is a strong spatial predictor such as hazard exposure. Our findings show that a wide range of risk perception and evacuation behavior can be estimated with such downscaling techniques. While not yet commonly used among geographers, our study suggests that this kind of modeling can improve mapping of survey results and, in particular, can provide spatially-explicit behavioral information for risk assessment and policy.

Keywords: survey data; small-area estimations; mapping; risk perception; evacuation behavior

30 1. Introduction

During the last decade, floods have caused the largest share of economic losses among all disasters and associated with other weather-related hazards they have affected 2bn people (IFRC, 2018). This has been attributed to population and economic growth in disaster-prone areas (Coronese et al. 2019), and the most recent IPCC report (2018) emphasized that climate change has increased flooding due to more intense rainfall. The costs associated with flood protection continue to rise along European rivers due to climate change (Blöschl et al. 2019). Despite the substantial public funds allocated to flood risk reduction, governmental investments in flood protection are often inadequate and planning policies are failing to reverse the trends of rising risk and the increasing number of exposed people. The evacuation of high-density districts to safer areas before a disaster is proving increasingly challenging (OECD, 2018). This is leading to a “behavioral turn” in disaster risk management (Kuhlicke et al., 2020), with recent strategies advocating that less protected households are individually responsible to look after themselves, as stretched public budgets are deemed unable to carry the costs for upgrading structural measures (Slavikova, 2018) and policy is increasingly relying on individual resilience (Begg et al., 2017) – i.e. the ability to individually prepare for, respond to and recover from disasters (Reghezza-Zitt & Rufat, 2019).

Whilst the perception and behavior of individuals, businesses and public services before, during and after a crisis has a significant impact on damages, recovery and resilience, current assessments fail to include such critical factors because they are not well understood (Ward et al., 2020). While there have been extensive efforts to model and map proxy indicators of hazard vulnerability from broad to fine scales) – e.g. SoVI (Cutter 1996) and other vulnerability or resilience index projects (Beccari

2016), vulnerability indices typically lack relevant self-reported information from

55 residents on topics like risk perception and preparedness behavior (de Sherbinin et al. 2019). Many previous efforts to measure disaster risk perception and preparedness take place either at broad scales, such as national surveys (Lee et al. 2015), or fine-scale case studies limited to individual communities or neighborhoods (Rufat et al. 2020).

Spatially disaggregating data from broad-scale surveys has limitations due to lack of
60 representativeness within geographic subunits (Lax & Phillips, 2009). Comparisons between community-scale studies can also be limited by varying times and modes of data collection or absence of data entirely in certain communities of interest.

Furthermore, even data from a study that is representative at the community level may mask important variations among people and places within communities.

65 Recent advances in statistical and spatial modeling for small-area estimation are one avenue to overcome some of these data and research design limitations. Multi-level regression and poststratification (MRP) and associated methods combine self-reported data from surveys with census or other data on population characteristics to systematically estimate responses across a study area at specified geographic scales.

70 MRP has been widely applied to estimate responses from national-level surveys to states, counties, or other administrative units (Caughey & Warshaw, 2019; Howe, 2018; Howe et al., 2015; Lax & Phillips, 2009; Mildemberger et al., 2016; Park et al., 2004). It has also been used to map public health indicators at similar scales (Downes et al., 2018; Howe et al., 2019a; Zhang et al., 2014, 2015). Although MRP has been typically
75 used to estimate responses in geographic subunits from national data, it is not limited to these scales and can conceptually be applied to estimate public responses from any

corpus of survey data within a defined study area, as long as appropriate population-level data are available at the geographic scale of interest.

The aim of this paper is to explore whether risk perception and adaptive behavior can be inferred from census and exposure data, for the purpose of integrating spatially explicit data on perceptions and behavior into risk, vulnerability, resilience or adaptation assessments. The ability to predict risk perception from census data would also be valuable in informing decision-makers, helping to develop targeted communication, tailoring risk communication strategies and campaigns. Predicting behavior and evacuation from census data would also prove critical for emergency management, helping to anticipate the magnitude of shelter and other special needs, pinpointing the areas and groups most likely to end up refusing to evacuate or coming back too soon, and bring into sharp focus the key role of spontaneous evacuation triggers, degraded living conditions exhaustion, or delayed panic reactions. Several research questions arise from such a goal; for example, which risk perceptions or behaviors are possible to estimate or infer from census data? An underlying question is what precautions users of survey data should take before jumping to mapping the results, not only to disentangle social and spatial variations but also to assess the relevance of mapping survey results while avoiding ecological bias and finding the right balance between the fineness of the scale and the robustness of the estimates. This requires consideration not only of the relevance of MRP to provide accurate estimates at different scales, including fine-scale uniform geographic units, but also the data requirements to validate such downscaled survey results.

Here we provide an example application of MRP using a large survey dataset on flood risk perceptions and behavior collected in Paris, France. We further model

responses at very high spatial resolution using population data at the municipality, census tract and the 200m grid scale. The remainder of the article is organized as follows. Section 2 present the current issues of “behavior-blind” risk assessments and mapping and how applying MRP to risk perception and behavior survey data could lead to a model improvement. Section 3 describes the context of our case study, while Section 4 details the methods to produce small-area estimation at different scales. The results are compared across scales and validated in Section 5. The discussion in Section 6 includes recommendations for using MRP and downscaling techniques to infer results from survey data while considering some precautions before mapping surveys results, and conclusions are provided in Section 7.

2. Literature review

2.1 “Behavior-blind” risk assessments and mapping

Whilst the perception and behavior of individuals, businesses, and public services before, during and after a crisis has a significant impact on damages, recovery and resilience, current assessments fail to include such critical factors because they are not well understood (Ward et al., 2020). The current focus on structural measures, monetary impacts and cost-benefit analyses (Hudson & Botzen, 2019) relies on some flawed underlying assumptions that leave aside social inequalities, actual behavior, underlying motivations, and capacities that can lead to significant differences in resilience across society (Rufat et al., 2020). The new EU Strategy on Adaptation to Climate Change is relying on the same “behavior-blind” assumptions (citizens will be aware, capable, motivated, abiding, etc.) and investing mostly on digital transformation and resilient infrastructure. A similar focus is present in recent U.S. government investments in the

Federal Emergency Management Agency's new Building Resilient Infrastructure and
125 Communities program (Mendelsohn et al., 2021). Such a narrow focus runs the risk of
hollowing out resilience by overlooking the perception, knowledge, capacity,
motivation, and behavior of citizens (Kuhlicke et al., 2020). This hinders the movement
towards more inclusive CCA & DRR called for by the UN Sendai Framework (2015-
2030) and the UN Sustainable Development Goals (SDGs 2030). They require a robust
130 knowledge base for action on behavior, resilience and adaptation to inform the
prioritization of actions, test policies and measure progress (Michel-Kerjan, 2015). The
vision of a 2050 disaster- and climate-resilient society cannot be achieved by relying on
"behavior-blind" assessments, decisions, and policy.

2.2 Risk perception and census data

135 One of the key challenges in risk, vulnerability and resilience is how to address the role
of risk perceptions and how perceptions influence behavior (Siegrist & Árvai, 2020).
Previous contradictory evidence on behavior hampers giving recommendations for
policy and risk management (Lechowska, 2018), such as the design of targeted risk
communication strategies (Höppner et al., 2012). Although numerous theoretical
140 frameworks have been developed (Kuhlicke, 2019), no definitive explanation has yet
been found (Siegrist & Árvai, 2020) and opposite conclusions can be reached from
different case studies (Wachinger et al., 2013). As they focus in turn on different
dimensions (sociological, economical, psychological, etc.), either on internal or
personal factors (gender, age, education, income, values, trust, etc.), external or
145 contextual factors (vulnerability, institutions, power, oppression, cultural backgrounds,
etc.), risk or environmental factors (perceived likelihood, experienced frequency, etc.),

or informational factors (media coverage, experts, risk management, etc.), they rely on hardly compatible qualitative and quantitative approaches, diverging sets of variables, and different methods.

150 Whilst most theories assume that high risk perception will lead to personal preparedness and then to risk mitigation behavior, the opposite has been repeatedly verified (Wachinger et al., 2013). The main sources of uncertainty include the many drivers beyond demographic, social, and cultural factors (Rufat et al., 2020), awareness, underestimation of risk (Poussin et al., 2014), place attachment (De Dominicis et al., 155 2015), previous hazard experience (Botzen et al., 2015) or the use of short-term horizons by households (Haer et al. 2017) and decision-makers in planning and risk management (Hartmann & Driessen, 2017). However, it is still challenging to disentangle which factors drive risk perception in a specific area or among specific groups (Rufat, 2014). One of the reasons stems from the interdisciplinary nature of the 160 field, shaped by different sets of assumptions, theories, methodologies, the lack of consideration for the spatial dimension, leading in turn often to diverging results (Rufat et al., 2015). A final challenge is the facts that risk perception, motivations, capacity and behavior are rarely linked, virtually never spatially analyzed (O'Neill et al., 2015), and studies usually overlook scale-dependency issues. This situation is hindering the 165 production of spatially explicit risk perception and evacuation baseline data that would be essential to emergency management. As a result, contemporary flood management (among other hazards) is inadequately informed by risk perception (Birkholz et al., 2014).

2.3 Evacuation behavior, risk perception and census data

170 The evacuation of high-density districts to safer areas before or during a disaster is
proving increasingly challenging, as demonstrated by the misplaced trust in public
preparedness (Baubion, 2015) and misleading lessons drawn from recent experiences,
such as before the 2016 flood in Paris (OECD, 2018), or the discarding of the
evacuation plan of London when Operation Sassoon was deemed unfeasible (Fekete &
175 Friedrich, 2018). While the importance of social and behavioral determinants of
vulnerability and resilience has been addressed by a wide range of disciplines (Fuchs &
Thaler, 2018), predicting the actual behavior of citizens is still a major challenge
(Poussin et al., 2015). It is equally difficult to appreciate how disputed factors in risk
perception influence households' or individuals' choices to take protective action or the
180 non-actions that may exacerbate the risk (Bamberg et al., 2017), as well as the trigger of
the evacuation decision, or even the refusal to evacuate (Kreibich et al., 2017). An
especially broad gap can be observed between risk perception and behavior leading to a
disconnection between decision-makers and affected people with respect to risk
communication, risk management and risk knowledge (Kellens et al., 2013). People can
185 have complex and sometimes risk-taking attitudes (Lazrus et al., 2016), and such a
multidimensional issue is quite challenging to implement in the current single-number
indices and aggregative assessment methods (Rufat, 2013). In addition, while most
assessments assume that vulnerability remains constant over time and most evacuation
plans suppose fixed behaviors, perception and behavior dynamics are influencing each
190 other, interacting with the flow of events, past experience and other's decisions,
sometimes in unpredictable ways (Wang et al., 2021).

The literature on evacuation is focused on compliance with protective action recommendations and mandatory evacuations, when decision-makers consider the risk to the population too great to permit them to remain at their place (Lindell et al., 2019).

195 Compliance behavior plays a key role in the success of emergency management and response operations (Kuhlicke, 2019). Previous studies found inconsistent and inconclusive results on the determinants of evacuation-related decisions; the relatively small number of empirical studies point towards highly hazard-specific and context-specific factors (Wang et al., 2021). Conversely, spontaneous or voluntary evacuation
200 refers to people leaving their location because of perceived risk without being directed to do so (Lindell et al., 2011). This dimension is even less often empirically studied (Kim & Oh, 2015), just as other critical factors are usually overlooked in models and simulations because of the lack of empirical knowledge (Aerts et al., 2018). These factors include people refusing to evacuate, panics or weariness to stay in degraded
205 living conditions after a few days, bursts of spontaneous evacuation that can overwhelm or even defeat the rescue, safety and recovery operations. Tackling the interplay between risk perception, motivations, capacity and individual behavioral dynamics including evacuation is challenging, and they are virtually never spatially analyzed (O'Neill et al., 2015). Digital technologies have recently prompted high expectations
210 about their ability to facilitate evacuation (Wang et al., 2019; Zou et al., 2018), however with the exception of hurricanes in the United States (Kim & Oh, 2015), little empirical knowledge has been collected on warning, communication and actual evacuation behavior, especially in large cities (Kreibich et al., 2017). A final challenge stems from risk-taking behavior during an emergency (Rufat et al., 2015). During floods most
215 studies highlight the fatalities associated with the use of motor vehicles (Hamilton et al.,

2020), the lack of trust in warnings and forecasts, and/or getting out to move a car, prevent or limit its immersion (Debionne et al., 2016). Finally, while the question of evacuation compliance and re-entry management are the most examined issues, evacuation refusals are less often explicitly addressed and might be linked to risk perception or vulnerability.

2.4 Statistical and spatial modeling for small-area estimation (MRP)

Hazard vulnerability and resilience are highly dependent on individual and household-level perceptions and behaviors, yet acquiring data on these topics typically requires household surveys that are resource-intensive. Survey samples are typically designed to achieve representativeness of a target population, and disaggregating these data across geographic subunits can result in highly uncertainty estimates, particularly across many units of varying population sizes (Lax & Phillips, 2009). Small-area estimation techniques like MRP can overcome these limitations to accurately estimate individual responses at subnational scales, and also inform understanding of the processes that influence geographic variation in responses (Howe et al., 2015). Survey sampling is often not perfectly representative due to overrepresentation of certain sociodemographic groups or geographic clustering in sample design (Groves et al., 2004). MRP can address these issues, and it can even provide accurate population-level estimates from highly non-representative polls (W. Wang et al., 2015). Researchers in political science, sociology, geography, and public health have adopted and validated MRP methods (Allan et al., 2020; Fowler, 2016; L. C. Hamilton et al., 2015; Howe et al., 2015, 2019a; Lax & Phillips, 2009; Mildenberger et al., 2016; Pacheco, 2011; Tausanovitch & Warshaw, 2013; Warshaw & Rodden, 2012; Zhang et al., 2014, 2015).

MRP takes place in two steps. First, a multilevel regression model is fit to an
240 individual-level dependent variable using a combination of individual-level
demographic predictors, grouped random effects, and group-level predictors across
geographic or other groupings of individuals within the dataset. In addition to using
fixed effects as in classical regression, the multilevel model uses random effects for
grouping predictors. Random effects are related to each other by their hierarchical
245 structure and thus are partially pooled towards their group mean. Greater pooling occurs
for smaller groups and when group-level variance is low. Group-level predictors are
also used to improve model fit (Buttice & Highton, 2013). Group-level predictors can
be extracted from external datasets based on respondents' geographic location.
Examples include county- or district-level census aggregated data (Howe et al., 2015;
250 Tausanovitch & Warshaw, 2013), or even finer resolution tract-level climate or land
cover data (Howe et al., 2019b). In the second step, the fitted model is poststratified to
the population of the study area across each demographic-geographic type within the
population, using census data cross-tabulated across each type. Estimates for each
demographic-geographic type can then be aggregated by their population to produce
255 estimates for every geographic unit. Importantly, poststratification can correct for
differences in representativeness between the survey sample and the population.

Recent examples applying MRP in the domain of climate change and hazards
include Mildenberger et al. (2016) and Howe et al. (2015), which produce estimates of
climate change opinion within Canada and the U.S., respectively, by fitting a multilevel
260 model to a dataset of multiple national surveys. In addition, the MRP models of climate
opinion in Howe et al. (2015) were shown to be unbiased when compared to a set of
independent county-level surveys (L. C. Hamilton et al., 2015). Recent work has also

validated MRP estimates of U.S. disaster preparedness behavior in metropolitan areas against American Housing Survey data and found the estimates to be unbiased and
265 similar in accuracy to representative surveys (Howe, 2018). Extending MRP to fine spatial scales within urban areas, another recent study applied and validated MRP to map risk perceptions of extreme heat hazards at the U.S. census tract level (Howe et al., 2019a). MRP requires population-level data with joint distributions (i.e. crosstabs) across individual-level predictors. However, recent advances using multilevel regression
270 with synthetic poststratification (MRSP) can eliminate the need for joint distributions and instead rely on marginal distributions to generate synthetic population crosstabs (Leemann & Wasserfallen, 2017).

3. Context

Two flood events in 2016 and 2018 and one major European flood exercise in 2016
275 occurred in rapid succession in the Paris metropolitan area in France. Each one was salient enough to receive international coverage for days. On the flood plain of the Seine River with an oceanic climate, Paris is exposed to river lowland floods, urban floods with drainage issues paralyzing the historical core and other weather-related hazards (Reghezza-Zitt & Rufat, 2015). Shortly before the floods, a study raised the
280 alarm about the fading flood memory and recommended the intensification of risk awareness (Baubion, 2015). After the 2016 and 2018 floods, a study by the Organization for Economic Co-operation and Development (OECD) estimated that a major flooding of the Seine River could affect up to 5 million residents and cause up to €30 billion worth of damage (OECD, 2018). After the major EU Sequana flood exercise
285 in March 2016, the June 2016 flood affected 486 municipalities over the course of

several weeks, causing 1bn € damages when the Seine River reached a 6.3m level in Paris: 18,000 were evacuated, 750 refused to evacuate and 4 people died (Richert et al., 2017). In January 2018, the Seine River reached a 5.9m level in Paris, causing 200 million € damages and affecting 242 municipalities over the course of several weeks: 2,500 people were evacuated and over 300 refused to evacuate (OECD, 2018).

4. Methods

4.1 Survey

We surveyed about flood risk perception and evacuation behavior from households 6 months after the last flood in Paris, France. The survey was administered face to face in the Paris metropolitan area from September to December 2018 to a representative sample ($n = 2,976$) of the population, with a random sampling and a spatial and social stratification to ensure the representativeness at different scales. Half of the sample are residents living in the flood zone, and one-third are living in the indirect impact zones, indicating that while their home might not be flooded, they might still face power, water or heating outages, sewer backflow and similar situations lasting several days. The rest of the sampled respondents live outside the direct and indirect exposure zones (Fig.1). The questionnaire was administered face to face with a geolocation of the place of residence, which is used to compute actual flood risk, with a 100m buffer to ensure privacy. The full questionnaire comprised 80 questions and required 15-20 minutes to complete.

4.2 Variables

We selected the answers to eleven questions representing a wide range of risk

perception and behavior as the dependent variables for this study. Two represent trust and flood predictions, five represent flood risk perception (worry, awareness, relative exposure, control over floods, self-assessed preparedness), and five capture flood evacuation behavior (abide official order, go out for car, spontaneous evacuation, no evacuation at all):

(1) *trust_forecast*: Is it possible to rely on flood forecasts?

(2) *predict_difficult*: In your neighborhood, are floods easy, or difficult to predict?

(3) *flood_worry*: Do you worry about flooding?

(4) *flood_awareness*: Is your home in a flood risk zone?

(5) *more_exposed*: Is your home more or less exposed to flooding than the rest of the Paris metropolitan region?

(6) *control_difficult*: In your neighborhood, are floods easy, or difficult, to prevent and control?

(7) *well_prepared*: In the event of a flood, how would you assess your household's level of preparedness?

(8) *evacuation_abide*: If you received an official evacuation instruction, but you couldn't see any flood, would you leave?

(9) *get_car*: In the event of a flood, would you go outside to get your car?

(10) *evacuation_spontaneous*: In case of flooding, the water is not visible, and you experience no effects. However, if 1 in 3 families on your street had already left (33%), would you spontaneously evacuate, without being instructed to do so?

(11) *evacuation_refuse*: respondents refusing to evacuate (this is not a direct question, but an indicator of responses where no evacuation question was answered with a “yes” under any scenario or circumstance).

4.3 MRP-MRSP models at three geographic levels

There is useful spatial information in social survey data that is sometimes analyzed explicitly for spatial patterns that may not accurately account for biases due to sampling or underlying demographic variation. Small-area estimation methods like MRP can reveal spatial patterns in survey data while also accounting for how representative the sample is of the broader population (Howe, 2018). However, the transition from individual data to aggregation in administrative units might subject interpretations to ecological bias. The “modifiable areal unit problem” (MAUP) commonly refers to two separate effects: variation in statistical results between different levels of aggregation and by the regrouping of data into different configurations at the same scale. In its traditional application MRP can be constrained by the MAUP since it relies on existing administrative units. These existing units, due to the way their boundaries are drawn, may propagate biases into subsequent small-area estimates. For example, U.S. congressional district boundaries are often drawn for partisan advantage (e.g. gerrymandering). In addition, administrative units at the same scale may vary dramatically in population (such as U.S. counties or states), which leads to varying uncertainties in small-area estimates for different units at the same scale, if the estimates are based on representative survey data (Howe et al. 2015) Here we show how MRP can be extended from larger-scale (municipality) to finer-scale administrative divisions (tracts) and uniform geographic units (grid cells with population counts) to map geographic variation in our outcome variable. The approach also frees us from relying on data often aggregated in administrative units. Another potential bias could be the “neighborhood effect averaging problem” (NEAP) referring to the traditional approach of measuring individual exposure by the static residence or by averaging it at the

neighborhood level (Kwan, 2018). The NEAP implies that the geographic contextual variables used as covariates in small-area-estimation models are inherently uncertain since they only incorporate data based on where respondents live—not where they work or may otherwise travel—and do so using arbitrary administrative units to define
360 respondents' neighborhoods. In extending our model to uniform geographic units (grid cells), we attempt to address the latter potential bias described by the NEAP. We do not have data on individual mobility to address potential uncertainties introduced by people being exposed to flooding beyond their residential neighborhood. However, we assume that effect of flood exposure on risk perceptions is likely to be strongest when
365 individuals' homes or neighborhoods are flooded. We therefore compute for all scales the share of each geographic unit inside the flood delineation zone as well as the exposure to indirect effects – i.e. residences that might not be directly flooded but might still face power, water or heating outages, sewer backflow and similar situations lasting several days during a major flooding.

370 We develop a hybrid MRP-MRSP (Leemann & Wasserfallen, 2017) model to estimate population proportions for our eleven variables at the census tract scale and 200m grid scale within the Paris metropolitan area (we also develop a municipality-scale model by aggregating tract-level model estimates). We first acquire tract-level census joint distributions for age (5 groups) and sex (2 groups) (INSEE, 2017a). We
375 also utilize estimates of total population at the 200m grid scale across the Paris metropolitan area (INSEE, 2015). At the grid scale, we estimate synthetic joint population distributions for age and sex for each grid cell by identifying the most-overlapping census tract polygon for each grid cell, and applying the tract's population proportions for each age by sex combination to the estimated population of the grid cell.

380 For each tract and associated grid cell we extract a variable termed *exposure* indicating whether the polygon has (0) low flood exposure; (1) indirect flood exposure; and (2) direct flood exposure (IIBRBS, 2013). We also extract a numeric variable termed *share_flood* indicating the proportion of the polygon in the zone of direct exposure to flooding. At the tract level we also use additional census data on income, poverty and
385 living standards as potential covariates (INSEE, 2017b).

We next fit a multilevel logistic regression model to the individual-level survey data for each dependent variable. The model includes random effects for the individual demographic variables (age, sex, and sex by age). It also includes nested geographic random effects for the census tract, municipality, and department. We then include
390 census tract-level fixed effects (or covariates) for *exposure* and *share_flood*. To determine these tract-level covariates, we iteratively tested sets of models using tract-level fixed effects for flood exposure (*exposure* and *share_flood*), per capita income, poverty rate, educational attainment (proportion with a less than high school education and proportion with at least a university-level education), national origin (proportion of
395 immigrants), home ownership (proportion of renters), household tenure (proportion who moved within the past two years), and family structure (proportion of households with children under 25 years and proportion of single-parent households). Starting with a base “null” model using only individual-level predictors and geographic random effects, we fit subsequent models using each possible geographic covariate individually and in
400 combination. We selected sets of covariates that resulted in a lower Bayesian Information Criterion (BIC) for the respective model, relative to the null model. Across all dependent variables, the only models that consistently returned lower BIC values were those that included our two flood exposure variables.

Our model estimates the probability Pr that a given respondent i had one of two
405 possible responses to the question being modeled, represented by $y_{[i]}$. The model for
each dependent variable is specified as follows at the individual level:

$$Pr(y_i = 1) = \text{logit}^{-1}$$

$$\alpha_j^{age} N(0, \sigma_{age}^2), \text{for } j = 1, \dots, 5$$

$$\alpha_k^{sex} N(0, \sigma_{sex}^2), \text{for } k = 1, \dots, 3$$
410
$$\alpha_l^{sex:age} N(0, \sigma_{sex:age}^2), \text{for } l = 1, \dots, 15$$

Each variable is indexed over individual I and over response categories $j, k, l, m,$
 $n,$ and o for age, gender, the interaction of age by gender, and geography (census tract)
respectively. Each variable is modeled as drawn from a normal distribution with mean
zero and estimated variance σ^2 . The census tract geography variable, m , is modeled as
415 follows:

$$\alpha_m^{tract} N(\alpha_{n[m]}^{muni} + \beta^{exposure} \cdot exposure_m + \beta^{share.flood}$$

$$\cdot share.flood_m, \sigma_{tract}^2), \text{for } m = 1, \dots, 1023$$

Exposure is a three-level ordinal variable indicating whether the census tract has
low, indirect, or direct flood exposure, and *share.flood* is the proportion of the census
420 tract in the zone of direct flooding exposure. *Muni* indicates the municipality within
which the census tract is nested, and is modeled as:

$$\alpha_n^{muni} N(\alpha_{o[n]}^{dep}, \sigma_{muni}^2), \text{for } n = 1, \dots, 111$$

Each municipality is further nested within eight broader *departement*-level (county) administrative areas in the Paris metropolitan region, which are modeled in

425 turn as:

$$\alpha_o^{dep} N(0, \sigma_{dep}^2), \text{ for } o = 1, \dots, 8$$

Models for each dependent variable were fit in R using the lme4 package's GLMER function (Bates et al., 2014).

In the next step, poststratification, we apply our fitted regression model to every
430 populated census tract and grid cell, resulting in estimated proportions for each age by sex population cell. We then weight these proportions by their population within each geographic unit, resulting in a total estimated proportion for each geographic unit (census tract or grid cell) for the modeled dependent variable. Municipal-level results are then calculated by aggregating proportions across census tracts by municipality.

435 Finally, we compare the model estimates at different scales by creating three new geographical variables for each dependent variable. Each time, we subtract the estimates from the top level to the bottom level estimates in order to highlight the discrepancies: the delta between municipal-level and tract-level estimates for each of the eleven dependent variables, the delta between municipal-level and grid cells, and the
440 delta between tract-level and grid cells. Positive values indicate that the upper-level estimates are higher than the lower-level estimates, negative values represent the opposite situation, while values close to zero indicate that the two estimate converge for these geographical units. This allows to assess the convergence or discrepancies of the results between the three geographical levels.

445 4.4 Validation

We validate our model estimates at the grid, tract, and municipality level using two cross-validation approaches. We first cross-validate at the municipality level by randomly selecting groups of sampled municipalities with at least 100 respondents across three urban-periurban categories with varying levels of flood exposure:

450 downtown Paris (3 groups), inner counties (2 groups), and outer counties (3 groups).

We chose municipalities within these categories for validation with the goal of representing a diverse set of neighborhoods with varying population densities and flood exposure, under the assumption that these contextual factors may be associated with flood risk perceptions. For each group, we create a new simulated dataset by removing
455 all respondents from that group, then fitting and poststratifying our MRP model to the simulated dataset. We then compare these MRP estimates for the selected group of municipalities to the actual proportion of responses from the selected area.

We next cross-validate at the grid cell and census tract level using a modified cross-validation technique that allows us to simulate the presence of an individual grid
460 cell or census tract with a large number of responses ($n \geq 50$), since our primary survey dataset lacks such response density in any one small geography. For each dependent variable we extract the distribution of model predictions across all grid cells (or census tracts). For each distribution, we identify values in the distribution where the number of respondents n in the sample who live in grid cells with the same predicted value,
465 rounded to the nearest whole number, is greater than 50 (the median number of residents per grid cell). For each value where $n \geq 50$, we iteratively create a simulated survey dataset that randomly sets aside 50 respondents who live in grid cells with that predicted value. The set-aside respondents then serve as a comparison group against

grid-level predictions from a model fit to the simulated dataset without these

470 respondents. After fitting this model, we calculate the mean absolute difference between the proportion in the set-aside sample and the mean predicted proportion across those respondent's respective grid cells. For comparison, we also calculate mean absolute difference between the proportion of respondents in the set-aside sample and the disaggregated proportion of remaining respondents in the simulated survey dataset as a
475 baseline indicator of model accuracy. We repeat this sampling and cross-validation process 99 times for each dependent variable at both the grid cell and census tract level.

5. Results

5.1 Statistical distribution of the estimates

Overall, the model estimates are well dispersed for all variables at all scales (Fig.2). The
480 only exception is *evacuation_refuse* summarizing the answers on evacuation, reflecting the fact that there might be no situation in which the respondents would consider to evacuate. The incidence of this indicator in the population is small as only a slim minority of respondents to the survey reflected such a behavior (4%). At all scales, the model estimates for this variable are both consistently very low and narrowly focused
485 around the median. Comparing scales, the distributions appear more dispersed and more symmetrical for the other variables at the tract level, followed by the grid level. The dispersion is lower and dissymmetry more pronounced at the municipality level (which includes fewer geographic units).

5.2 Estimates at the municipal scale ($n = 1,300$)

490 Figure 3 displays the spatial distribution of the estimates for all questions at the

municipal scale. This spatial distribution is both well differentiated and well distributed; it does not remain homogeneous throughout the area nor depend too much on the higher administrative level, except for *predict_difficult* and *evacuation_spontaneous* for which a county effect is noticeable. Overall, flood exposure seems to play a critical role in the models as the municipal level estimates often reflect the river network, especially towards the outskirts. The outlier is again the model for *evacuation_refuse* producing almost homogeneous estimates for most of the metropolitan area, except for some focal points downtown and in the north-east.

5.3 Estimates at the tract scale (n = 5,261)

Overall, the estimates at the census tract scale support the previous findings (Fig.4). The county effect is less noticeable for *predict_difficult* and *evacuation_spontaneous*, especially downtown. The dependence on the hydrographic network is quite noticeable for *flood_awareness*, which is to be expected, whereas *evacuation_refuse* on evacuation refusal remains an outlier with little variation and very low estimates except for some focal points downtown and in the north-east. For all other models, a diversity of factors seems to be at play and the estimates present a well-distributed spatial pattern.

5.4 Estimates at the grid scale (n = 73,166)

Figure 5 displays the spatial distribution of the MRP estimates for all questions at the 200m grid scale. Only inhabited cells are represented. The spatial distribution of estimates at the 200m grid scale is convergent with the previous results for all questions. Some county and municipal effects are still apparent, especially for *predict_difficult* and *evacuation_spontaneous*, *evacuation_refuse* on evacuation refusal remains an outlier,

whereas hazard exposure proves to be a more powerful predictor for *flood_awareness* on flood exposure awareness.

515 **5.5 Comparison between the estimates at the different scales**

We compare municipal and tract scale estimates by subtracting the upper level (municipal) results from the lower level (tract) ones for all questions. Higher estimates at the lower level (tract) than upper level (municipal) result in positive values represented in red on Figure 6, higher estimates at the upper level result in negative values and are represented in blue, while convergent estimates result in values close to zero represented in white. In general, the differences are quite small, except for *flood_awareness* on flood exposure awareness and *more_exposed* on perceived relative exposure. There is less convergence downtown though, where densities are higher and municipalities are subdivided into a greater number of tracts. Outwards, the outliers reflect specific places with very low densities, most notably the Fontainebleau forest, a very large tract with only around 200 residents in the Southeast. For the same tract, these small differences can be positive for one question and negative for another.

Figure 7 displays the subtraction of upper level (municipal) estimates from the lower level (grid) ones for all 10 questions using the same representation modalities, 200m cells intersecting the border of the upper level geographic units are removed to reduce false positives. Overall, the differences are quite small, except for *flood_awareness* on flood exposure awareness and *more_exposed* on perceived relative exposure, both of which are heavily impacted by the flood zone delineation. The flood zone delineation is also apparent in the small differences for *control_difficult* on perception of flood prevention and control. For the other models, a diversity of factors

seems to be at play and the estimates present varied spatial pattern. Once again, for the same geographic unit, these small differences can be positive for one question and negative for another.

The comparison of tract and grid scale estimates with the same representation modalities in Figure 8 is convergent with the previous results.

5.6 Validation of the estimates at the different scales

We conducted cross-validation tests of our model estimates at the tract, grid, and municipality level, as described above. Averaging across each variable, the mean absolute error (MAE) was 6.7 ± 2.3 percentage points at the municipality level, 6.6 ± 1.4 points at the grid cell level, and 8.5 ± 1.9 points at the census tract level (Figure 9). The median MAE was 6.8 points at the municipality level, 7.0 points at the grid cell level, and 8.3 points at the census tract level. For the majority of variables, the municipality-level estimates had a lower MAE than the grid-level and tract-level estimates, and the grid-level estimates had a uniformly lower MAE than the tract-level estimates. The lowest MAE was found for the *evacuation_refuse* variable, which reflects the highly asymmetric distribution of responses to this variable (with only 4% of respondents indicating refusal to evacuate).

6. Discussion

Our main objective was to assess whether survey results on risk perception and adaptive behavior can be inferred from census and exposure data using a small-area estimation model, and to discuss their possible subsequent implementations in risk, vulnerability, resilience or adaptation models and assessments. We find that an hybrid MRP-MRSP model can transition from individual survey data to small-area estimation, relying on

individual demographic variables and flood exposure data to provide estimates at
560 different scales, including fine-scale uniform geographic units (e.g. 200m grid cells of
populated areas). While MRP has been applied to estimate responses from national-
level surveys to states, counties, or lower administrative units (Caughey & Warshaw,
2019; Downes et al., 2018; Howe, 2018; Howe et al., 2015; Mildemberger et al., 2016;
Zhang et al., 2015), our results show that an hybrid MRP-MRSP model is not limited to
565 these scales and can be applied to estimate public responses using survey data for fine
scales within municipalities and neighborhoods. Our findings show that a wide range of
risk perception and adaptive behavior can be estimated with this approach, as the model
estimates are well statistically dispersed for all variables at all scales while the spatial
distribution is both well differentiated and well distributed. This remains true for all
570 scales; estimates do not remain homogeneous throughout the area nor depend too much
on the higher administrative level. Therefore, one promising result of this study is that a
carefully designed survey within any defined study area allows to infer small-area
estimations of risk perception and adaptive behavior from census data, including for
fine-scale uniform geographic units.

575 Another aim of our study was to assess the relevance of mapping survey results
while avoiding ecological bias and finding the right balance between the fineness of the
scale and the robustness of the estimates. Our findings show convergence between
estimates at different scales, even if this is less consistent for flood awareness and
perceived relative flood exposure. As a result, we suggest that our approach – when
580 using uniform areal units – reduces potential MAUP biases in interpretation of MRP
results. In this particular case, the NEAP bias is less of a concern as flood exposure is
more residence-based than mobility-dependent and we have implemented both direct

flood exposure and exposure to indirect effects. The novelty of our approach is to compare small-area estimations at different scales, allowing to analyze the sensibility of MRP estimates to MAUP and NEAP biases, which has not been systematically studied yet. Furthermore, our iterative cross-validation process show that the resulting estimates are robust with a mean absolute error of about 7 percentage points at both the broadest geographic scale (municipality level) and the finest geographic scale (grid cell level). These results suggest that small-area estimation models can make robust fine-scale predictions when using appropriate geographic covariates that exhibit strong associations with the outcome variable, as is the case with flood exposure in our model. Researchers have validated MRP methods (Allan et al., 2020; Fowler, 2016; Hamilton et al., 2015; Howe et al., 2015, 2019a; Lax & Phillips, 2009; Mildenerger et al., 2016; Pacheco, 2011; Tausanovitch & Warshaw, 2013; Warshaw & Rodden, 2012; Zhang et al., 2014, 2015). MRP models of climate opinion in were shown to be unbiased when compared to a set of independent county-level surveys (Hamilton et al., 2015) while MRP estimates of U.S. disaster preparedness behavior were to be similar in accuracy to representative surveys (Howe, 2018). While we lack external datasets for comparable direct validation, our MRP model accuracy based on internal cross-validation was similar to that found in an internal cross-validation of estimates for U.S. counties and small-population states from a national MRP model (Howe et al., 2015).

Contrary to our expectations, the level of precision of small-area estimations is not decreased at finer scales, even as the number of geographic sub-units increases dramatically. The validation of estimates at grid level ($n = 73,166$) performed systematically better than estimates at the tract level ($n = 5,261$), and performed similarly to the higher level municipality estimates ($n = 1,300$). We suggest one reason

may be that population composition is more homogeneous at finer scales, making prediction easier. Performance at finer scales is less consistent, however, when the initial survey data for a specific question are skewed (as in our case for trust in
610 forecast), where respondents giving the same answer are spatially clustered (e.g. flood awareness), or the subpopulation giving a specific answer is low (e.g. evacuation refusal). While small-area estimation techniques like MRP can overcome the uncertainties deriving from disaggregating data across geographic subunits (Howe et al., 2015), particularly across many units of varying population sizes (Lax & Phillips,
615 2009), our findings suggest that robust estimates can be achieved at a very fine spatial scale, provided there is a strong spatial predictor. Further accuracy gains at fine geographic scales may also be achieved if more detailed census demographic data are available as joint distributions (i.e. crosstabs) at the geographic scales to be predicted in the poststratification stage, or alternatively through synthetic crosstabs based on
620 marginal distributions (Leemann & Wasserfallen, 2017). The choice of individual-level predictors for small-area estimation is context-dependent, and should be informed by previous research about predictors known to be associated with the outcome variable. For instance, race/ethnicity was used as an individual-level predictor in small-area estimates of climate change opinion in the U.S. (Howe et al., 2015), while first language
625 was used in a similar model in Canada (Mildenberger et al., 2016). Here, we used joint distributions of sex by age at the census tract and grid cell level, since more detailed demographic joint distributions were not available at these fine geographic scales (for example, data on race or ethnicity are not collected in France).

Small-area estimation techniques like MRP are not commonly used among
630 geographers. One promising result from our study is that this kind of modeling is a

fruitful direction to improve mapping of survey results. Survey sampling is often not perfectly representative due to overrepresentation of certain socio-demographic groups or geographic clustering in sample design (Downes et al., 2018; Groves et al., 2004; Wang et al., 2015). We argue that small-area estimation models, such as the hybrid
635 MRP-MRSP modeling employed here, can address these issues; it can even provide accurate population-level estimates from highly non-representative surveys (e.g. Wang et al. 2015).

An underlying question is what precautions surveys should take before jumping to mapping the results, disentangling social and spatial variations, inferring sample
640 survey results at different scales, while tackling scale-dependency issues. Our findings suggest that unless survey sampling is conducted to be representative at the scale needed, then it is important to model and adjust for population characteristics before making claims about geographic patterns in survey results.

Therefore, greater efforts to explicitly collect validation data from randomly
645 selected locations that are set aside for later validation analyses should be considered by future projects. Survey samples are typically designed to achieve representativeness of a target population; however our validation results show that is less straightforward when a respondents' subgroup is small or clustered. As a result, caution about interpreting model estimates like those we have produced here is required. An important step is to
650 understand the underlying survey data and sampling design. The estimates for small-population areas have more inherent uncertainty because they may not have respondents included in the survey sample, in which case their estimates are driven entirely by the modeled relationships present in the remaining data. This could lead to errors in the estimates for some small-population areas if they are different from the remaining

655 region in ways not known to the model. This limitation is why we indicate data-poor areas in our maps. By contrast, areas with larger populations are likely to have lower uncertainties because their residents are more likely to be present in a representatively-sampled dataset. We advise thus to increase the sample size for smaller subpopulations (e.g. vulnerable people or people refusing to evacuate) and/or less dense neighborhoods
660 (e.g. having a higher sampling ratio in suburban peripheries than in the city center), when accurate estimates of both high- and low-population areas are needed.

7. Conclusion

We find that a wide range of risk perception and adaptive behavior can be estimated with an hybrid MRP-MRSP model combining a specifically designed survey with
665 census and exposure data. The model estimates are statistically well dispersed for all variables and the spatial distribution is both well differentiated and well distributed at all scales, including fine-scale uniform geographic units. The small-area estimations do not remain homogeneous throughout the area nor depend too much on information at higher administrative levels. The primary limiting factor is with low-incidence
670 variables: when only a slim minority of respondents to the survey display a specific behavior. This limitation is evident in our measure of refusals to evacuate (4% total incidence), reflecting the fact that for very few respondents was there no situation in which they would consider evacuation.

While our measures of flood risk awareness and perceived flood exposure are, as
675 expected, highly dependent on location relative to the hydrographic network, questions on trust, confidence in forecast and control over floods are more challenging to spatially disaggregate since they are less predicted by flood exposure. These variables, especially at the higher levels with fewer spatial units, display greater dependence on information

at higher administrative levels. Performance at finer scales is less consistent when the
680 initial survey data for a specific question are skewed (as in our case for trust in
forecast), where respondents giving the same answer are spatially clustered (e.g. flood
awareness), or the subpopulation giving a specific answer is low (e.g. evacuation
refusal). For all other risk perception and adaptive behavior, a diversity of factors is at
play and the estimates have a spatial distribution with complex marquetry. We also find
685 convergence between estimates at different scales. Furthermore, our iterative cross-
validation process demonstrates that the model estimates are robust, with a low error
rate (7%) similar to that found for small-area estimates of U.S. counties from national
surveys (Marlon et al., 2020). Contrary to our expectations, the level of precision of
small-area estimations is not decreased at finer scales, even as the number of geographic
690 sub-units increases dramatically.

One promising result is that a carefully designed survey allows for small-area
estimation of risk perception and adaptive behavior from census data, including for fine-
scale uniform geographic units. We advise researchers, however, to consider
oversampling smaller subpopulations (e.g. vulnerable people or people refusing to
695 evacuate) and/or less dense neighborhoods (e.g. having a higher sampling ratio in
suburban peripheries than in the city center), which will likely improve accuracy of the
estimates across the area of interest. The resulting accurate population-level estimates
and multilevel estimates of risk perception and adaptive behavior can help overcome
current shortcomings in risk, vulnerability, resilience or adaptation indices that typically
700 lack relevant information on risk perception, adaptation or preparedness behavior. The
ability to infer perceptions and behavior from census data would be most valuable for
decision-makers and first-responders, helping to develop targeted communication,

tailoring risk communication strategies and campaigns, and helping to anticipate the magnitude of shelter and other special needs. For example, emergency managers could
705 pinpoint the areas and groups most likely to end up refusing to evacuate or coming back too soon after a disaster.

This study demonstrates that an hybrid MRP-MRSP model is not limited to broad geographic scales and can be applied to estimate public responses using survey data for fine scales within municipalities and neighborhoods. We show how small-area
710 estimation can be extended from existing administrative divisions to uniform geographic units (grid cells with population counts) to map geographic variation in survey outcome variables. Small-area estimation techniques using survey data are not commonly used among geographers. We argue that this kind of modeling is a fruitful direction to improve mapping of survey results. We raise the underlying question of
715 what precautions surveys should take before inferring sample results at different scales or mapping them. We advise specifically designing survey samples to be representative at the scales needed and to account for biases due to sampling or underlying demographic variation, i.e. to model and adjust for population characteristics before making claims about geographic patterns in unweighted survey results. Geographers
720 and other social scientists should strongly consider using small-area estimation techniques to infer and map results from survey data to understand underlying spatial patterns in ways that account for how representative the sample is of the broader population. This opens the door to other applications that require accurate population-level estimates and spatially disaggregated estimates from polls and survey data.

725 Beyond hazard mitigation and preparedness, additional practical applications could include, for example, mapping health-related perceptions and behavior for public health

campaigns, mapping energy-related behaviors to optimize transmission infrastructure, or mapping local government policy preferences to inform decision-making.

730 References

- Aerts, J. C. J. H., Botzen, W. J., Clarke, K. C., Cutter, S. L., Hall, J. W., Merz, B., Michel-Kerjan, E., Mysiak, J., Surminski, S., & Kunreuther, H. (2018). Integrating human behaviour dynamics into flood disaster risk assessment. *Nature Climate Change*, 8(3), 193–199.
<https://doi.org/10.1038/s41558-018-0085-1>
- 735 Allan, J. N., Ripberger, J. T., Wehde, W., Krocak, M., Silva, C. L., & Jenkins-Smith, H. C. (2020). Geographic Distributions of Extreme Weather Risk Perceptions in the United States. *Risk Analysis*, 40(12), 2498–2508. <https://doi.org/10.1111/risa.13569>
- Bamberg, S., Masson, T., Brewitt, K., & Nemetschek, N. (2017). Threat, coping and flood prevention – A meta-analysis. *Journal of Environmental Psychology*, 54, 116–126.
<https://doi.org/10.1016/j.jenvp.2017.08.001>
- 740 Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Dai, B. (2014). *lme4: Linear mixed-effects models using Eigen and S4* (1.1-7) [Computer software].
<http://cran.r-project.org/web/packages/lme4/index.html>
- Baubion, C. (2015). Losing memory – the risk of a major flood in the Paris region: Improving
 745 prevention policies. *Water Policy*, 17(S1), 156–179. <https://doi.org/10.2166/wp.2015.008>
- Begg, C., Ueberham, M., Masson, T., & Kuhlicke, C. (2017). Interactions between citizen responsabilization, flood experience and household resilience: Insights from the 2013 flood in Germany. *International Journal of Water Resources Development*, 33(4), 591–608.
<https://doi.org/10.1080/07900627.2016.1200961>
- 750 Birkholz, S., Muro, M., Jeffrey, P., & Smith, H. M. (2014). Rethinking the relationship between flood risk perception and flood management. *Science of The Total Environment*, 478, 12–20.
<https://doi.org/10.1016/j.scitotenv.2014.01.061>
- Botzen, W., Kunreuther, H., & Michel-Kerjan, E. (2015). Divergence between individual perceptions and objective indicators of tail risks: Evidence from floodplain residents in New York City.
 755 *Judgment and Decision Making*, 10(4), 365–385. <http://journal.sjdm.org/15/15415/jdm15415.html>

- Buttice, M. K., & Highton, B. (2013). How Does Multilevel Regression and Poststratification Perform with Conventional National Surveys? *Political Analysis*, 21(4), 449–467.
<https://doi.org/10.1093/pan/mpt017>
- 760 Caughey, D., & Warshaw, C. (2019). Public Opinion in Subnational Politics. *The Journal of Politics*, 81(1), 352–363. <https://doi.org/10.1086/700723>
- De Dominicis, S., Fornara, F., Ganucci Cancellieri, U., Twigger-Ross, C., & Bonaiuto, M. (2015). We are at risk, and so what? Place attachment, environmental risk perceptions and preventive coping behaviours. *Journal of Environmental Psychology*, 43, 66–78.
<https://doi.org/10.1016/j.jenvp.2015.05.010>
- 765 De Sherbinin, A., Bukvic, A., Rohat, G., Gall, M., McCusker, B., Preston, B., & Zhang, S. (2019). Climate vulnerability mapping: A systematic review and future prospects. *WIREs Climate Change*, 10(5), e600. <https://doi.org/10.1002/wcc.600>
- Debionne, S., Ruin, I., Shabou, S., Lutoff, C., & Creutin, J.-D. (2016). Assessment of commuters' daily exposure to flash flooding over the roads of the Gard region, France. *Journal of Hydrology*, 770 541, 636–648. <https://doi.org/10.1016/j.jhydrol.2016.01.064>
- Downes, M., Gurrin, L. C., English, D. R., Pirkis, J., Currier, D., Spittal, M. J., & Carlin, J. B. (2018). Multilevel Regression and Poststratification: A Modeling Approach to Estimating Population Quantities From Highly Selected Survey Samples. *American Journal of Epidemiology*, 187(8), 1780–1790. <https://doi.org/10.1093/aje/kwy070>
- 775 Fekete, A., & Fiedrich, F. (Eds.). (2018). *Urban disaster resilience and security: Addressing risks in societies*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-68606-6>
- Fowler, L. (2016). The states of public opinion on the environment. *Environmental Politics*, 25(2), 315–337. <https://doi.org/10.1080/09644016.2015.1102351>
- Fuchs, S., & Thaler, T. (Eds.). (2018). *Vulnerability and resilience to natural hazards*. Cambridge 780 University Press.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology* (pp. xix, 424 p). J. Wiley.

- Hamilton, K., Demant, D., Peden, A. E., & Hagger, M. S. (2020). A systematic review of human behaviour in and around floodwater. *International Journal of Disaster Risk Reduction*, 47, 101561. <https://doi.org/10.1016/j.ijdr.2020.101561>
- Hamilton, L. C., Hartter, J., & Safford, T. G. (2015). Validity of county-level estimates of climate change beliefs. *Nature Climate Change*, 5(8), 704–704. <https://doi.org/10.1038/nclimate2720>
- Hartmann, T., & Driessen, P. (2017). The flood risk management plan: Towards spatial water governance: Towards spatial water governance. *Journal of Flood Risk Management*, 10(2), 145–154. <https://doi.org/10.1111/jfr3.12077>
- Höppner, C., Whittle, R., Bründl, M., & Buchecker, M. (2012). Linking social capacities and risk communication in Europe: A gap between theory and practice? *Natural Hazards*, 64(2), 1753–1778. <https://doi.org/10.1007/s11069-012-0356-5>
- Howe, P. D. (2018). Modeling geographic variation in household disaster preparedness across u. S. States and metropolitan areas. *The Professional Geographer*, 70(3), 491–503. <https://doi.org/10.1080/00330124.2017.1416301>
- Howe, P. D., Marlon, J. R., Wang, X., & Leiserowitz, A. (2019a). Public perceptions of the health risks of extreme heat across US states, counties, and neighborhoods. *Proceedings of the National Academy of Sciences*, 116(14), 6743–6748. <https://doi.org/10.1073/pnas.1813145116>
- Howe, P. D., Marlon, J. R., Wang, X., & Leiserowitz, A. (2019b). Public perceptions of the health risks of extreme heat across US states, counties, and neighborhoods. *Proceedings of the National Academy of Sciences*, 116(14), 6743–6748. <https://doi.org/10.1073/pnas.1813145116>
- Howe, P. D., Mildenerger, M., Marlon, J. R., & Leiserowitz, A. (2015). Geographic variation in opinions on climate change at state and local scales in the USA. *Nature Climate Change*, 5(6), 596–603. <https://doi.org/10.1038/nclimate2583>
- Hudson, P., & Botzen, W. J. W. (2019). Cost–benefit analysis of flood-zoning policies: A review of current practice. *WIREs Water*, 6(6). <https://doi.org/10.1002/wat2.1387>
- IIBRBS (2013). *Crues de reference et zones inondables* [Reference floods and flood zones]. Institution interdépartementale des barrages réservoirs du bassin de la Seine – Hydratec, France. Available from <https://www.data.gouv.fr/fr/datasets/crues-de-references-idf/>

- INSEE (2015). *Données Carroyées 200m Filosofi 2015. Fichier Localisé Social et Fiscal* [Filosofi 2015 200m grid data. Localized Social and Fiscal Database]. Institut National de la Statistique et des Etudes Economiques, France. Available from <https://www.insee.fr/fr/statistiques/4176305>
- INSEE (2017a). *Recensement de la Population 2017. Base Infracommunale IRIS* [2017 Population Census. Infra-Municipal IRIS Database]. Institut National de la Statistique et des Etudes Economiques, France. Available from <https://www.insee.fr/fr/statistiques/4799309>
- INSEE (2017b). *Revenus, pauvreté et niveau de vie en 2017. Base Infracommunale Filosofi* [Income, poverty and living standards in 2017. Infra-Municipal Filosofi Database]. Institut National de la Statistique et des Etudes Economiques, France. Available from <https://www.insee.fr/fr/statistiques/4479212>
- Kellens, W., Terpstra, T., & De Maeyer, P. (2013). Perception and communication of flood risks: A systematic review of empirical research: perception and communication of flood risks. *Risk Analysis*, 33(1), 24–49. <https://doi.org/10.1111/j.1539-6924.2012.01844.x>
- Kim, J., & Oh, S. S. (2015). Confidence, knowledge, and compliance with emergency evacuation. *Journal of Risk Research*, 18(1), 111–126. <https://doi.org/10.1080/13669877.2014.880728>
- Kreibich, H., Müller, M., Schröter, K., & Thielen, A. H. (2017). New insights into flood warning reception and emergency response by affected parties. *Natural Hazards and Earth System Sciences*, 17(12), 2075–2092. <https://doi.org/10.5194/nhess-17-2075-2017>
- Kreibich, H., Pech, I., Schröter, K., Müller, M., & Thielen, A. H. (2016). *New insights into flood warning and emergency response from the perspective of affected parties* [Preprint]. Hydrological Hazards. <https://doi.org/10.5194/nhess-2016-133>
- Kuhlicke, C. (2019). Risk and resilience in the management and governance of natural hazards. In C. Kuhlicke, *Oxford Research Encyclopedia of Natural Hazard Science*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199389407.013.299>
- Kuhlicke, C., Seebauer, S., Hudson, P., Begg, C., Bubeck, P., Dittmer, C., Grothmann, T., Heidenreich, A., Kreibich, H., Lorenz, D. F., Masson, T., Reiter, J., Thaler, T., Thielen, A. H., & Bamberg, S. (2020). The behavioral turn in flood risk management, its assumptions and potential implications. *WIREs Water*, 7(3). <https://doi.org/10.1002/wat2.1418>

- Lax, J. R., & Phillips, J. H. (2009). How Should We Estimate Public Opinion in the States? *American Journal of Political Science*, 53(1), 107–121.
- 840 Lazrus, H., Morss, R. E., Demuth, J. L., Lazo, J. K., & Bostrom, A. (2016). “Know what to do if you encounter a flash flood”: Mental models analysis for improving flash flood risk communication and public decision making: “know what to do if you encounter a flash flood.” *Risk Analysis*, 36(2), 411–427. <https://doi.org/10.1111/risa.12480>
- 845 Lechowska, E. (2018). What determines flood risk perception? A review of factors of flood risk perception and relations between its basic elements. *Natural Hazards*, 94(3), 1341–1366. <https://doi.org/10.1007/s11069-018-3480-z>
- Lee, T. M., Markowitz, E. M., Howe, P. D., Ko, C.-Y., & Leiserowitz, A. A. (2015). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change*, 5(11), 1014–1020. <https://doi.org/10.1038/nclimate2728>
- 850 Leemann, L., & Wasserfallen, F. (2017). Extending the use and prediction precision of subnational public opinion estimation: Extending use and precision of mrp. *American Journal of Political Science*, 61(4), 1003–1022. <https://doi.org/10.1111/ajps.12319>
- Lindell, M. K., Kang, J. E., & Prater, C. S. (2011). The logistics of household hurricane evacuation. *Natural Hazards*, 58(3), 1093–1109. <https://doi.org/10.1007/s11069-011-9715-x>
- 855 Lindell, M. K., Murray-Tuite, P. M., Lindell, M. K., Wolshon, P. B., & Baker, E. J. (2019). *Large-scale evacuation: The analysis, modeling, and management of emergency relocation from hazardous areas*. Routledge/Taylor & Francis Group.
- Marlon, J., Howe, P. D., Mildenerberger, M., Leiserowitz, A., & Wang, X. (2020, September 2). *Yale Climate Opinion Maps—U.S. 2020*. Yale Program on Climate Change Communication. <http://climatecommunication.yale.edu/visualizations-data/ycom-us/>
- 860 Mendelsohn, J., Johnson, G., Klima, K., Steratore, R., Cohen, S., Kirkwood, G., Dixon, L., Hastings, J. L., & Steinberg, P. S. (2021). *Developing Metrics and Scoring Procedures to Support Mitigation Grant Program Decisionmaking*. RAND Corporation.
- 865 https://www.rand.org/pubs/research_reports/RRA377-1.html

Hastings, J. L., & Steinberg, P. S. (2021). *Developing Metrics and Scoring Procedures to Support Mitigation Grant Program Decisionmaking*. RAND Corporation.

https://www.rand.org/pubs/research_reports/RRA377-1.html

Michel-Kerjan, E. (2015). We must build resilience into our communities. *Nature*, 524(7566), 389–

389. <https://doi.org/10.1038/524389a>

Mildenberger, M., Howe, P., Lachapelle, E., Stokes, L., Marlon, J., & Gravelle, T. (2016). The Distribution of Climate Change Public Opinion in Canada. *PLOS ONE*, 11(8), e0159774.

<https://doi.org/10.1371/journal.pone.0159774>

OECD. (2018). *Preventing the Flooding of the Seine in the Paris–Ile de France Region*. *Progress*

Made and Future Challenges (OECD Publishing). Organisation for Economic Co-operation and Development (OECD). <https://www.oecd.org/governance/risk/preventing-the-flooding-of-the-seine-2018.pdf>

O'Neill, E., Brennan, M., Brereton, F., & Shahumyan, H. (2015). Exploring a spatial statistical approach to quantify flood risk perception using cognitive maps. *Natural Hazards*, 76(3), 1573–

1601. <https://doi.org/10.1007/s11069-014-1559-8>

Pacheco, J. (2011). Using National Surveys to Measure Dynamic US State Public Opinion A Guideline for Scholars and an Application. *State Politics & Policy Quarterly*, 11(4), 415–439.

Park, D. K., Gelman, A., & Bafumi, J. (2004). Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls. *Political Analysis*, 12(4), 375–385.

[https://doi.org/10.1093/pan/12\(4\)375](https://doi.org/10.1093/pan/12(4)375)

Poussin, J. K., Botzen, W. J. W., & Aerts, J. C. J. H. (2014). Factors of influence on flood damage mitigation behaviour by households. *Environmental Science & Policy*, 40, 69–77.

<https://doi.org/10.1016/j.envsci.2014.01.013>

Poussin, J. K., Wouter Botzen, W. J., & Aerts, J. C. J. H. (2015). Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters. *Global Environmental Change*, 31, 74–84. <https://doi.org/10.1016/j.gloenvcha.2014.12.007>

Reghezza-Zitt, M., & Rufat, S. (2015). *Resilience Imperative: Uncertainty, risks and disasters*. Elsevier.

- Reghezza-Zitt, M., & Rufat, S. (2019). Disentangling the range of responses to threats, hazards and
895 disasters. Vulnerability, resilience and adaptation in question. *Cybergeo*.
<https://doi.org/10.4000/cybergeo.32917>
- Richert, C., Erdlenbruch, K., & Figuières, C. (2017). The determinants of households' flood
mitigation decisions in France—On the possibility of feedback effects from past investments.
Ecological Economics, 131, 342–352. <https://doi.org/10.1016/j.ecolecon.2016.09.014>
- 900 Rufat, S. (2013). Spectroscopy of urban vulnerability. *Annals of the Association of American
Geographers*, 103(3), 505–525. <https://doi.org/10.1080/00045608.2012.702485>
- Rufat, S. (2014). Towards a social and spatial risk perception framework. *Cybergeo*.
<https://doi.org/10.4000/cybergeo.27010>
- Rufat, S., Fekete, A., Armaş, I., Hartmann, T., Kuhlicke, C., Prior, T., Thaler, T., & Wisner, B. (2020).
905 Swimming alone? Why linking flood risk perception and behavior requires more than “it’s the
individual, stupid.” *WIREs Water*, 7(5). <https://doi.org/10.1002/wat2.1462>
- Rufat, S., Tate, E., Burton, C. G., & Maroof, A. S. (2015). Social vulnerability to floods: Review of
case studies and implications for measurement. *International Journal of Disaster Risk Reduction*,
14, 470–486. <https://doi.org/10.1016/j.ijdr.2015.09.013>
- 910 Siegrist, M., & Árvai, J. (2020). Risk perception: Reflections on 40 years of research. *Risk Analysis*,
40(S1), 2191–2206. <https://doi.org/10.1111/risa.13599>
- Tausanovitch, C., & Warshaw, C. (2013). Measuring Constituent Policy Preferences in Congress,
State Legislatures, and Cities. *The Journal of Politics*, 75(02), 330–342.
<https://doi.org/10.1017/S0022381613000042>
- 915 Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox-implications
for governance and communication of natural hazards: The risk perception paradox. *Risk Analysis*,
33(6), 1049–1065. <https://doi.org/10.1111/j.1539-6924.2012.01942.x>
- Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-
representative polls. *International Journal of Forecasting*, 31(3), 980–991.
920 <https://doi.org/10.1016/j.ijforecast.2014.06.001>

Wang, Y., Kyriakidis, M., & Dang, V. N. (2021). Incorporating human factors in emergency evacuation – an overview of behavioral factors and models. *International Journal of Disaster Risk Reduction*, 102254. <https://doi.org/10.1016/j.ijdrr.2021.102254>

925 Wang, Z., Lam, N. S. N., Obradovich, N., & Ye, X. (2019). Are vulnerable communities digitally left behind in social responses to natural disasters? An evidence from Hurricane Sandy with Twitter data. *Applied Geography*, 108, 1–8. <https://doi.org/10.1016/j.apgeog.2019.05.001>

Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., Emberson, R., Jenkins, S. F., Kirschbaum, D., Kunz, M., Mohr, S., Muis, S., Riddell, G. A., Schäfer, A., Stanley, T., Veldkamp, T. I. E., & Winsemius, H. C. (2020). Review article: Natural hazard risk assessments at the global scale. *Natural Hazards and Earth System Sciences*, 20(4), 1069–1096. <https://doi.org/10.5194/nhess-20-1069-2020>

Warshaw, C., & Rodden, J. (2012). How Should We Measure District-Level Public Opinion on Individual Issues? *The Journal of Politics*, 74(01), 203–219. <https://doi.org/10.1017/S0022381611001204>

935 Zhang, X., Holt, J. B., Lu, H., Wheaton, A. G., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2014). Multilevel regression and poststratification for small-area estimation of population health outcomes: A case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. *American Journal of Epidemiology*, 179(8), 1025–1033. <https://doi.org/10.1093/aje/kwu018>

940 Zhang, X., Holt, J. B., Yun, S., Lu, H., Greenlund, K. J., & Croft, J. B. (2015). Validation of Multilevel Regression and Poststratification Methodology for Small Area Estimation of Health Indicators From the Behavioral Risk Factor Surveillance System. *American Journal of Epidemiology*, 182(2), 127–137. <https://doi.org/10.1093/aje/kwv002>

Zou, L., Lam, N. S. N., Cai, H., & Qiang, Y. (2018). Mining twitter data for improved understanding of disaster resilience. *Annals of the American Association of Geographers*, 108(5), 1422–1441. <https://doi.org/10.1080/24694452.2017.1421897>

Figures

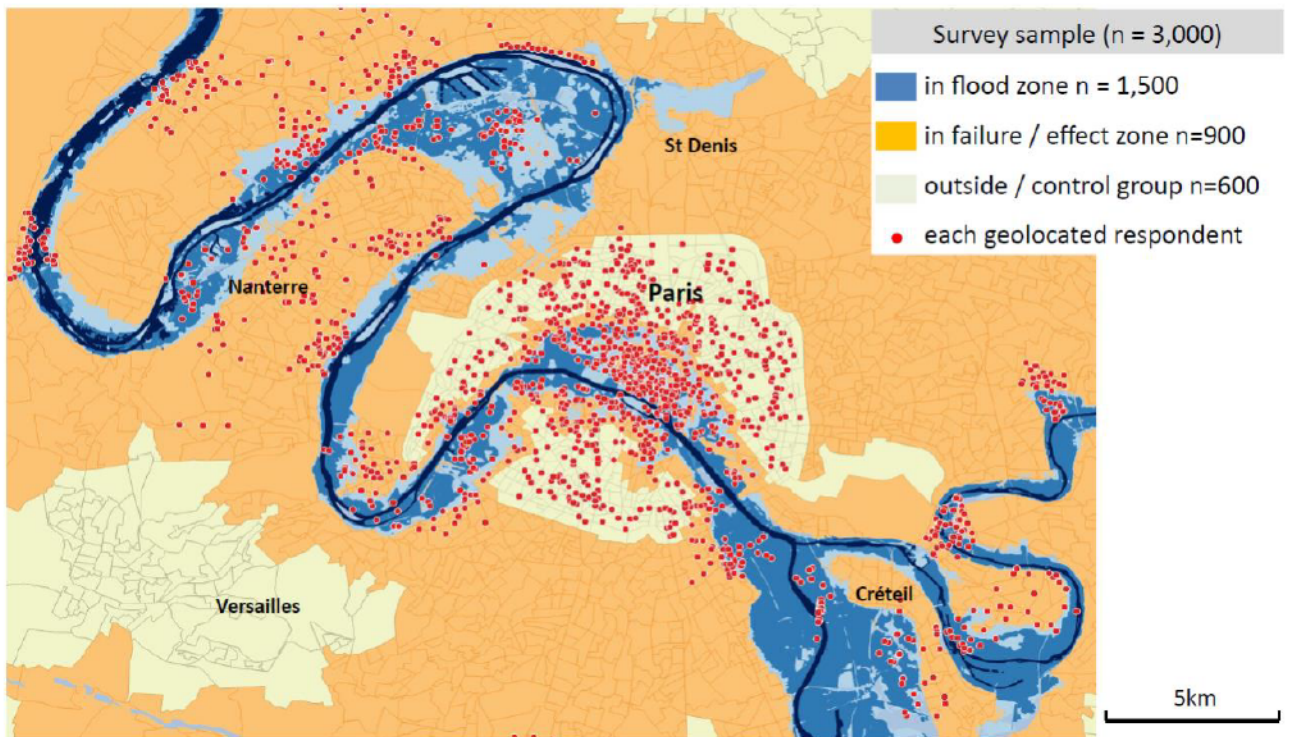


Figure 1. Map presenting the sampling method, flood zone and geotagged respondents.

950

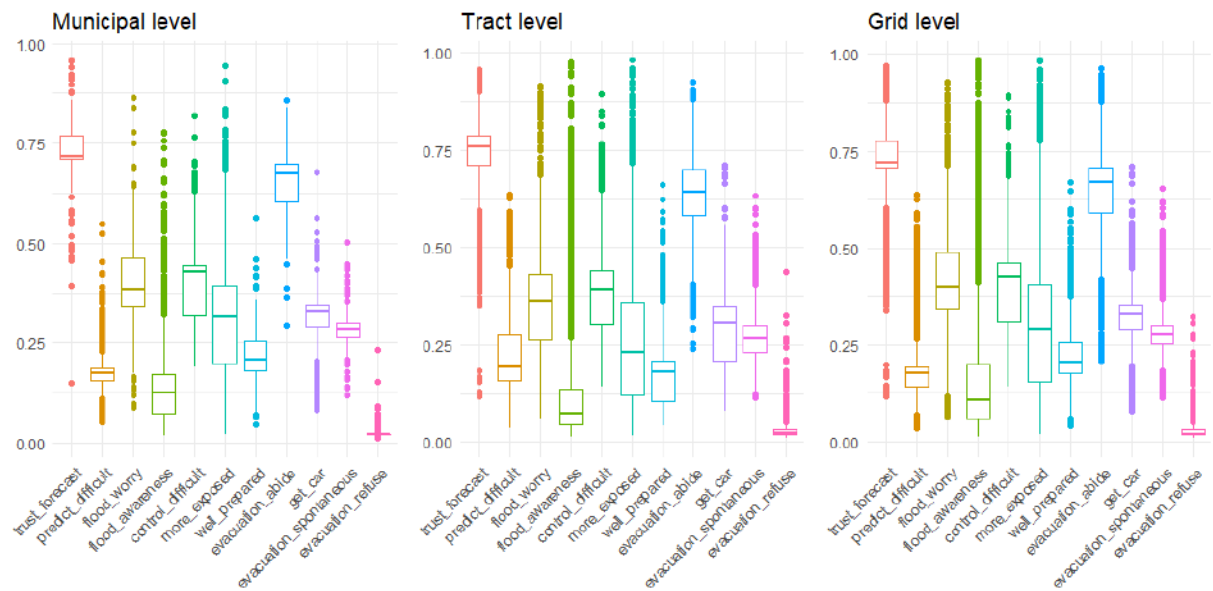


Figure 2. Boxplot of the estimates for all questions at the three scales.

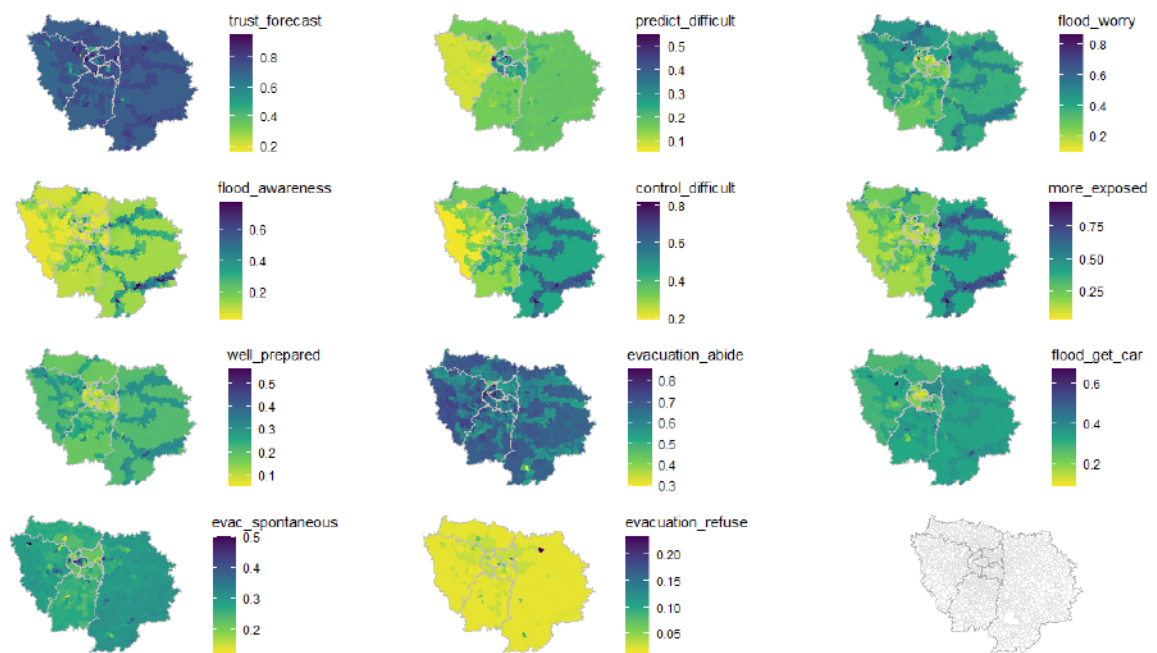


Figure 3. Map of the results for the 11 questions at municipal level.

955

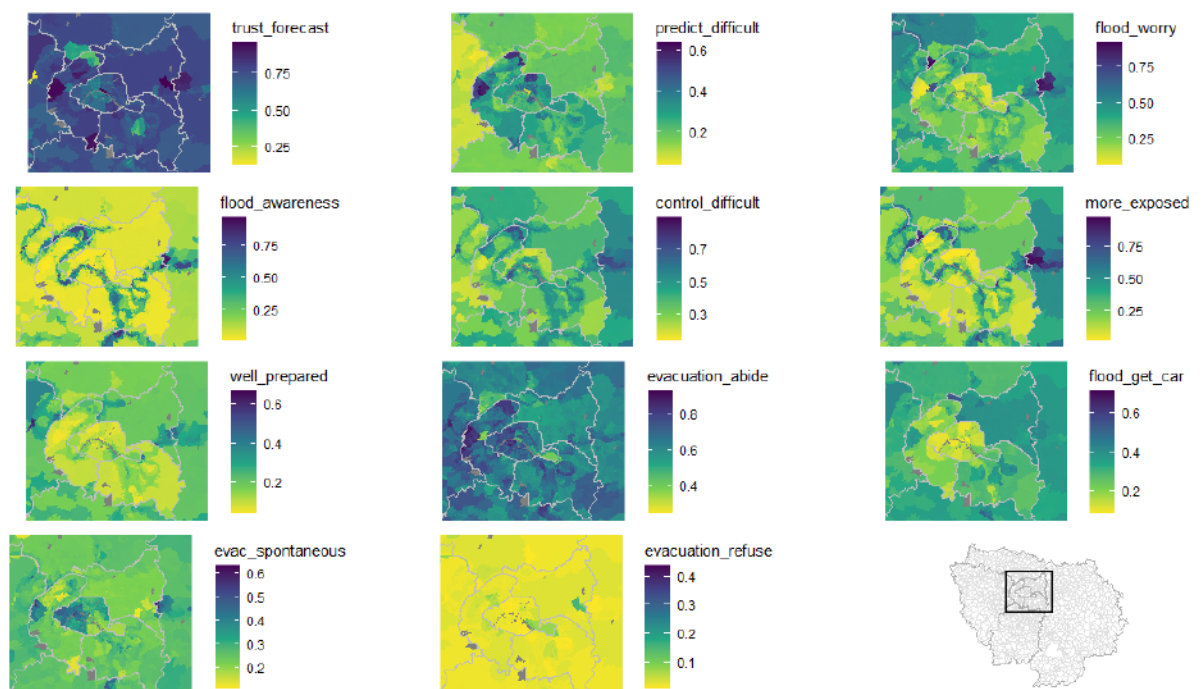


Figure 4. Map of the results for all questions at census tract level (downtown zoom).

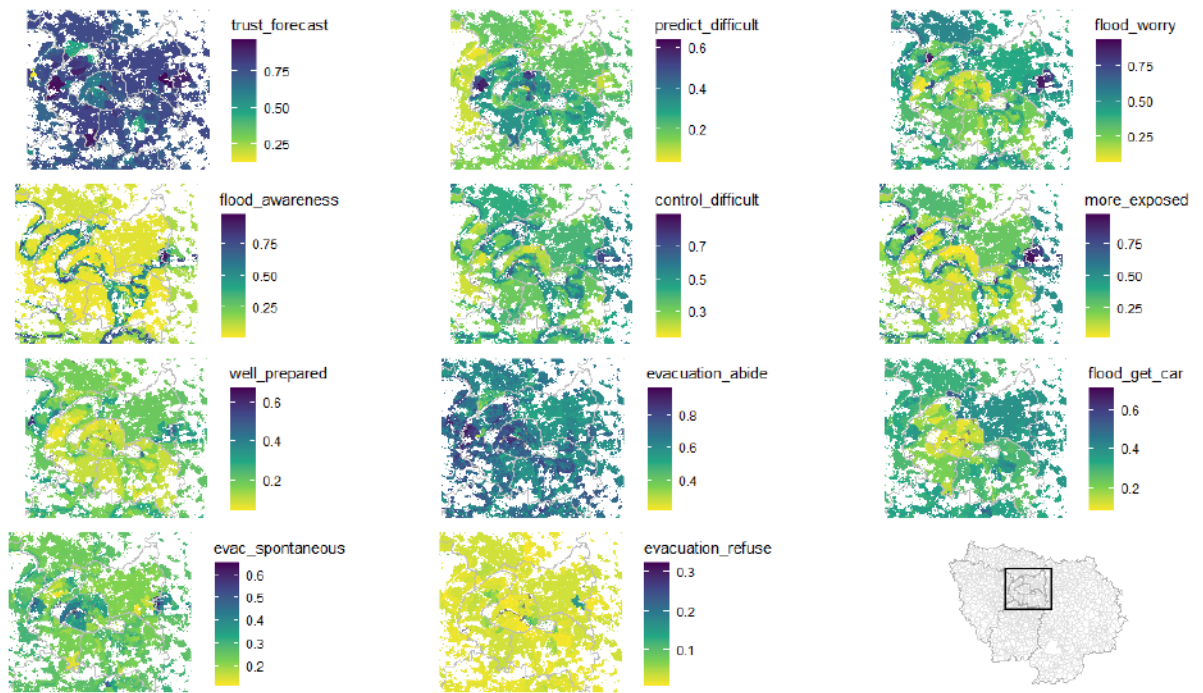


Figure 5. Map of the results for all questions at the 200m grid scale (downtown zoom).

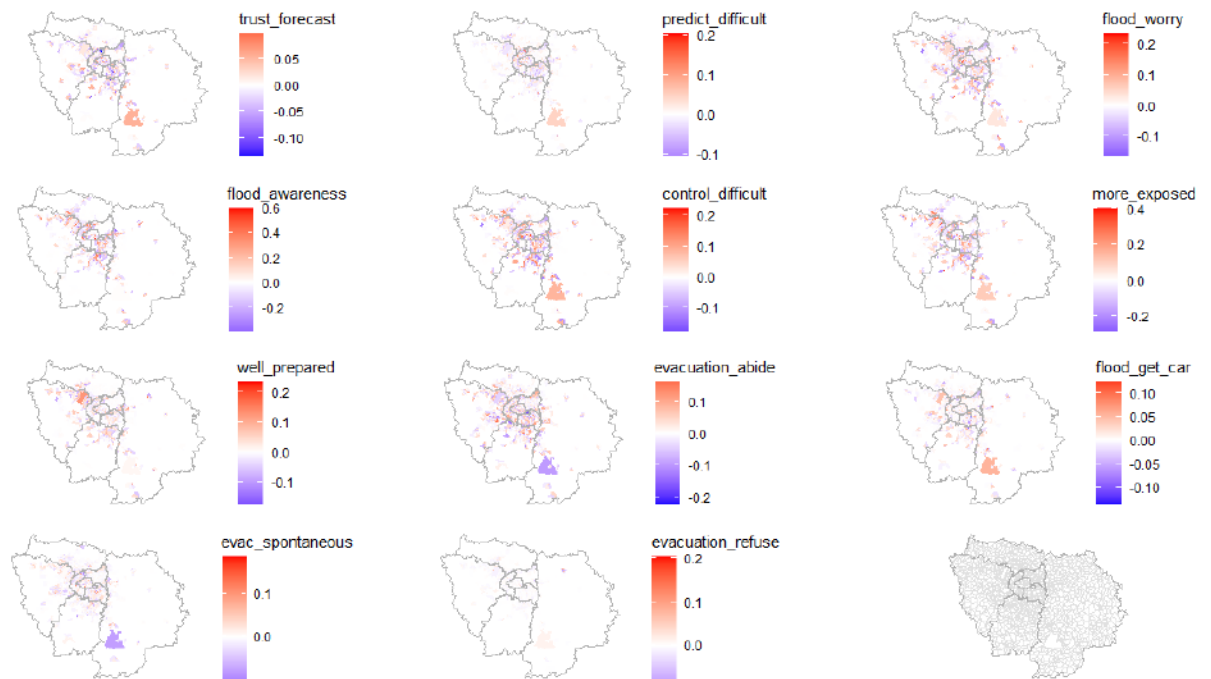


Figure 6. Map comparing estimates for the 11 questions at the municipal and tract scale.

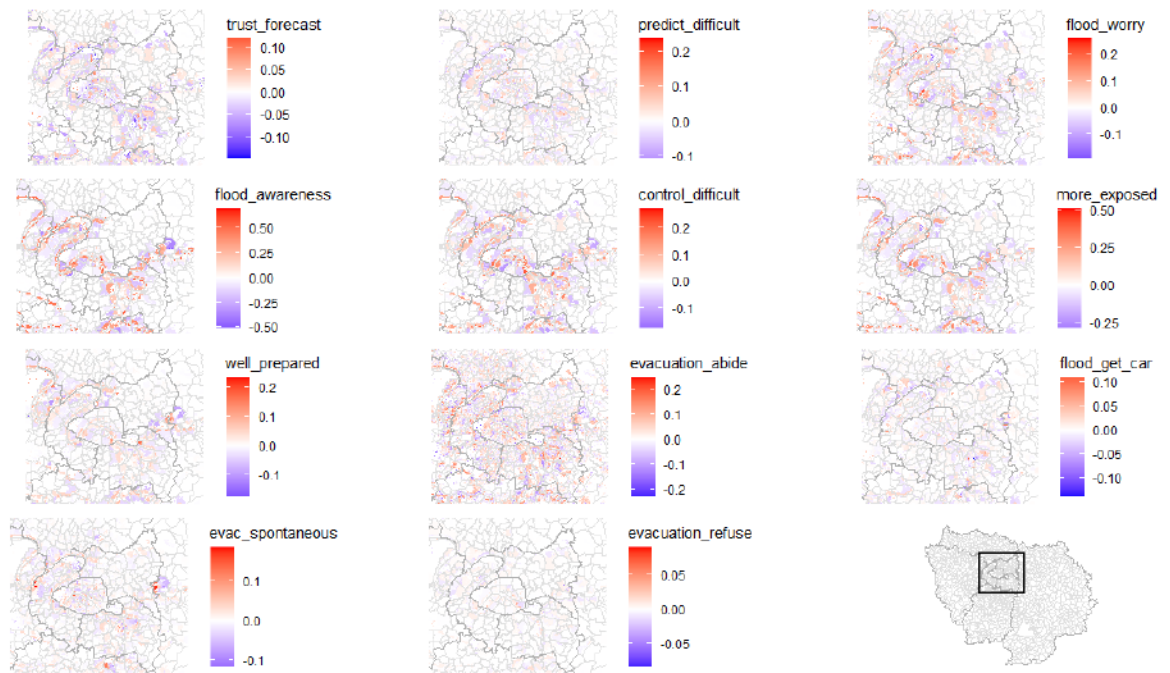


Figure 7. Map comparing estimates for the 11 questions at the municipal and 200m grid scale (downtown zoom).

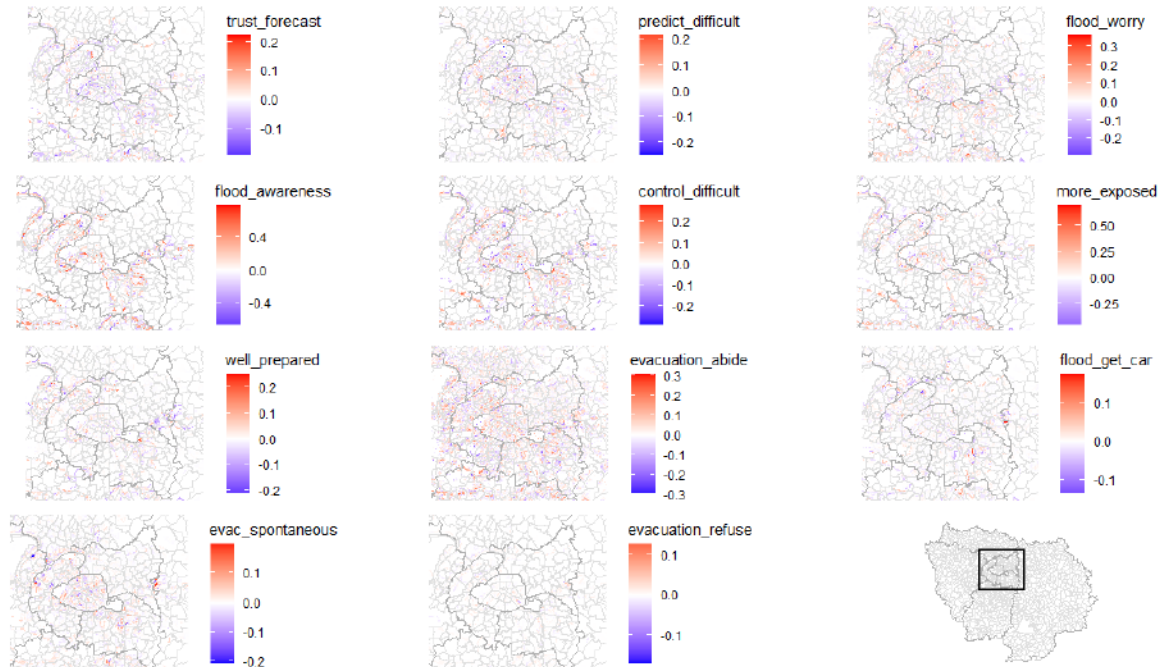


Figure 8. Map comparing estimates for the 11 questions at the tract and 200m grid scale (downtown zoom).

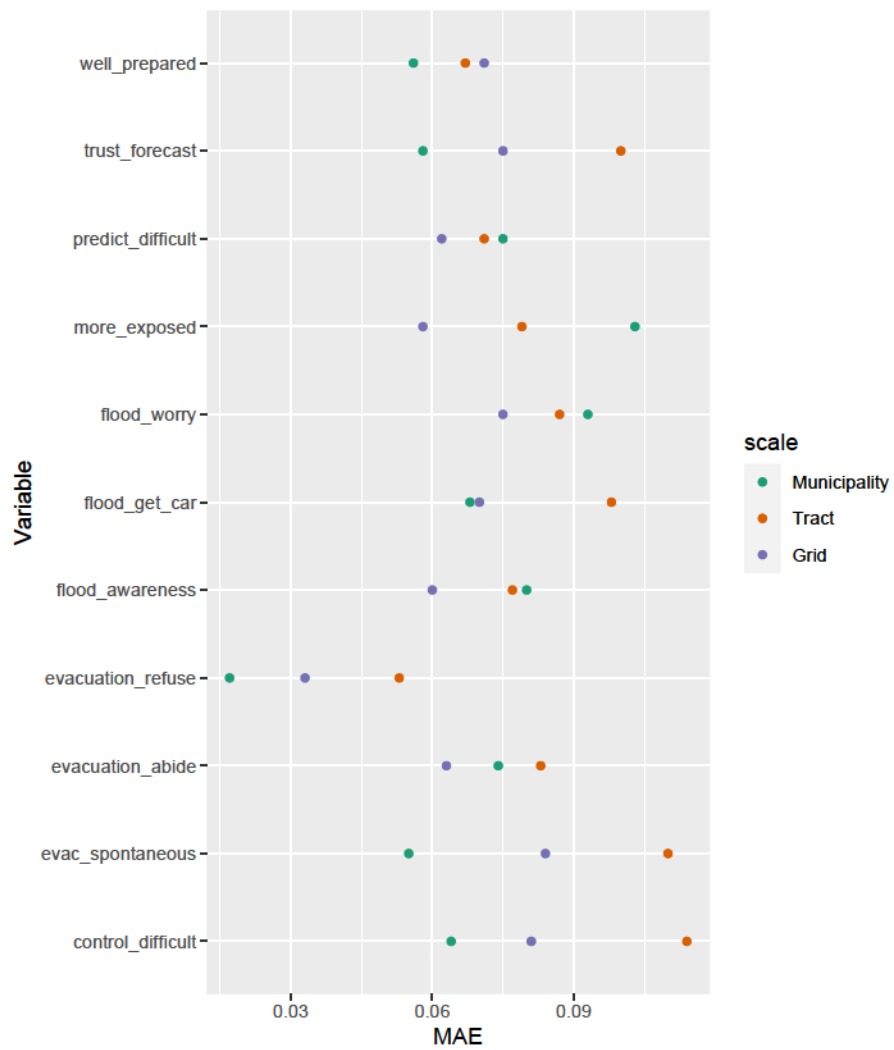


Figure 9. Graph of mean absolute error from validation tests for each variable at municipality, tract, and grid-cell scales.