

Measuring well-being at local level using remote sensing and official statistics data

Charlotte Articus, Christopher Caratiola, Hanna Dieckmann,
Max Gerhards, Ralf Münnich, Thomas Udelhoven ¹

Abstract

Measuring societal well-being as a multi-dimensional perspective on the conditions of people's life satisfaction has evolved to be an important task of European Official Statistics. Routinely, the focus of analysis is on country-comparisons. As central dimensions of well-being also vary on local level, we complement these insights by measuring well-being on the very low level of city districts and 100 metre grid cells. To achieve this, we combine high-resolution remote sensing data products with data from official statistics. As the data from different sources often have different scales, we discuss several scaling methods both from the field of geospatial research and from small area estimation. We calculate a composite well-being indicator on district and grid cell level for the city of Cologne and assess the influence of scaling methods and other construction decisions in a sensitivity analysis.

Keywords: composite indicator, kriging, small area estimation, sensitivity analysis.

¹ Charlotte Articus (articus@uni-trier.de); Christopher Caratiola (caratiola@uni-trier.de); Hanna Dieckmann (dieckmann@uni-trier.de); Max Gerhards (gerhardsm@uni-trier.de); Ralf Münnich (muennich@uni-trier.de); Thomas Udelhoven (udelhove@uni-trier.de), Trier University, Germany. Economic and Social Statistics Department and Environmental Remote Sensing and Geoinformatics.

The views and opinions expressed are those of the authors and do not necessarily reflect the official policy or position of the Italian National Institute of Statistics – Istat.

1. Introduction and motivation

The concept of well-being is gaining in importance within the global indicator framework for the Sustainable Development Goals (SDGs). The United Nations member states adopted 17 integrated SDGs in 2015. Well-being is directly included in the third goal, which is to ensure healthy lives and promote well-being for all at all ages (United Nations, 2019). For a long time, well-being was measured as gross domestic product (GDP) per capita. Easterlin (1974) has induced a paradigm shift from GDP as a proxy for well-being to the concept of relative income and the incorporation of aspects beyond income by showing that higher income leads to a higher perception of well-being only up to a certain point. Other initiatives, such as the Stiglitz-Sen-Fitoussi Commission Report (2009) and the European Commission GDP and beyond communication (2009), have supported the development from GDP per capita as a measure for well-being towards multidimensional well-being measures. Eurostat (2019) defines 8+1 dimensions as an overarching framework for the measurement of well-being:

- Material living conditions;
- Productive or main activity;
- Health;
- Education;
- Leisure and social interactions;
- Economic and physical safety;
- Governance and basic rights;
- Natural and living environment;
- Overall experience of life.

Stiglitz *et al.* (2009) identify material living standards, personal insecurity, social connections and relationships, environmental conditions and political voice and governance as main dimensions in the Stiglitz-Sen-Fitoussi report. The OECD (2020) defines human well-being in terms of eleven dimensions under the themes of material conditions and quality of life. All definitions comprise both individual and place-related factors. Thus, well-being highly depends on the living environment and differs not only from country to country, but rather is affected by local living conditions.

The traditional initiatives to measure well-being have focussed on international comparisons of country-level indicators (Eurostat, 2020; OECD, 2020). More recently, this perspective has been complemented by regional studies, both in research and official statistics. For example, some National Statistical Institutes have started to report well-being at a local level (see Office for National Statistics, 2019 and Istituto Nazionale di Statistica – Istat, 2019). Eurostat, National Statistical Institutes and the European Commission cooperate in a voluntary data collection exercise to build a database for measuring life quality in European cities (a project previously known as Urban Audit; see Eurostat, 2017). Eurofound (2020), the European Agency for the improvement of living and working conditions, recently examined the quality of life in European capitals in comparison to the rest of the country based on its own European Quality of Life Surveys (EQLS). Moretti *et al.* (2019) employ small area estimation techniques to estimate composite well-being indicators on a regional level. More detailed, they employ factor analysis to reduce the dimensionality of complex indicator systems and integrate this approach into multivariate small area statistics to gain estimates at municipality level.

Generally, the availability of data is a challenge for each initiative to measure well-being at a regional level. Data commonly used to assess the quality of life come from survey data such as the European Union Statistics on Income and Living Conditions (EU-SILC). This data cannot reliably be evaluated at a local level. Small area estimation techniques, as employed by Moretti *et al.* (2019), are a possible solution. Additionally, a feasible strategy is to exploit further data sources such as administrative data from local authorities (see *e.g.* Istituto Nazionale di Statistica, 2019). Integrating administrative and other data, however, may suffer from different degrees of granularity. Shuvo Bakar *et al.* (2020) provide a Bayesian approach to model predictions that help compensating overlapping geographical areas. Alternative methods from geostatistics are known as spatial resampling methods. Against this backdrop, we explore the opportunity of combining high-resolution remote sensing data² with official data as an efficient strategy to conduct analyses of well-being at local level.

2 In the field of remote sensing, high resolution data is used to describe raw images. In this paper, we use the term to describe products derived from satellite images or other georeferenced sources. These derived products are referred to as remote sensing data products in the field of remote sensing.

Since the late 1950s, geospatial data have been incorporated in the analysis of urban sociology. In his pioneering work, Green (1957) relates aerial photographic interpretation information with socio-economic data of Birmingham, Alabama and finds that photographic interpretation information can supplement and substitute other socio-economic data sources. Satellite data have since been used to estimate well-being at local level. Lo and Faber (1997) complement census data with satellite data and assess the quality of life in the Athens-Clarke County of Georgia with an environmental perspective. Ghosh *et al.* (2013) evaluate well-being using night-time light data and Engstrom *et al.* (2017) estimate economic well-being by extracting object and texture features from satellite images of Sri Lanka.

In this article, we analyse the potential of combining high-resolution remote sensing data and official data for small-scale estimation (*e.g.* block and district level) using the example of well-being in the city of Cologne. We consider data with a resolution of 100 metres or more as high-resolution data. The analysis of well-being is conducted at 100 metre grid cell and city district level. Since the data come from different sources and have different scales, we introduce scaling techniques from the field of geostatistics as well as small-area estimation and investigate methodological (*e.g.* different scales) and technical (*e.g.* confidentiality requirements) challenges of combining methods from both disciplines. The main focus is on methodological challenges, especially on how to deal with different scales. Special emphasis is put on different upscaling and downscaling methods and their impact on the composite indicator of well-being for the city of Cologne. By means of a sensitivity analysis, various uncertainty factors in the construction steps are investigated and quantified.

2. Data

This article combines INSPIRE conform³ Census 2011 grid cell data at 100 metre resolution, OpenStreetMap data and Pan-European High-Resolution Layers (HRL). The impact of scaling methods is investigated using the example of the composite indicator of well-being for the city of Cologne, since the City of Cologne (2017, 2014) provides georeferenced and socio-demographic data.

The composite indicator comprises socio-demographic and place-related data. Socio-demographic data are obtained from the Federal Statistical Office and the statistical offices of the Länder (2018) and the City of Cologne (2017). Data on single parents are available at 100 metre grid cell level from disaggregated census statistics. The 100 metre grid cell data do not contain values smaller than three due to disclosure control. In this study, empty cells are treated as zero, resulting in deviations of 2 percent (%) at city level for single parents. The City of Cologne (2017) offers a variety of statistical data at district and municipality level, including information on unemployment with time reference to December 2017 and single parents as of 31 December 2017.

Georeferenced data on schools, museums, play- and sports grounds, hospitals and libraries are published by the City of Cologne (2014) in their open data portal. These data together with OpenStreetMap data are used to conduct network analyses using the QNEAT3 (distance matrices) QGIS plugin (see Figure 1). Residential buildings are taken from OpenStreetMap and include all houses and apartments which are tagged as residential. In order to include distances from each address to primary schools, museums, hospitals, libraries and play- and sports grounds in the analysis of well-being, network analyses are conducted based on the shortest distance using streets tagged as highway⁴ from OpenStreetMap (OpenStreetMap contributors, 2019). The OpenStreetMap data are taken from the QuickOSM Plugin in QGIS. Figure 1 shows the distance to primary schools in 250 metre intervals for a part of Cologne⁵.

3 INSPIRE conformity means that spatial data are harmonised across Europe and comply with international geomatics standards (European Commission, 2019).

4 Highways include any type of road, street or path.

5 The distances are determined using the QNEAT3 (Iso-Areas) QGIS plugin, which only accepts a projected coordinate system. Therefore, the street network is taken from Geofabrik GmbH and OpenStreetMap contributors (2018).

Figure 1 - Distances to primary schools for a part of Cologne

Source: Own illustration based on data from OpenStreetMap contributors (2019), Geofabrik GmbH and OpenStreetMap contributors (2018) and the City of Cologne (2014)

Remote sensing data are also regarded as place-related data. HRL are obtained from satellite imagery by applying automatic processing and interactive rule based classification. Currently, HRL provide information on tree cover density and forest types, grasslands, wetness and water, small woody features and imperviousness (Copernicus, 2019a). In this study we integrate imperviousness data into the analysis of well-being at local level. The imperviousness product gives the percentage of impervious surfaces. Imperviousness data is available in the original 20 metre and 100 metre pixel size for the years 2006, 2009, 2012 and 2015 (Copernicus, 2019b). Furthermore, we consider data on vacant dwellings as place-related data in the broader sense. Results on vacant dwellings⁶ at 100 metre resolution are available from Federal Statistical Office and the statistical offices of the

6 The Census 2011 defines an apartment as vacant if it is neither rented out nor used by the owner on the date of the survey and if it is not a holiday and leisure apartment, diplomatic apartment, apartment of foreign armed forces and commercially used apartment.

Länder (2018). The treatment of empty cells as zero leads to a deviation of 11.5 % at city level. Census 2011 results for the city of Cologne are published by Federal Statistical Office and the statistical offices of the Länder (2018). In this article, we analyse well-being at 100 metre grid cell and city district level.

3. Theoretical framework for the construction of a composite indicator for well-being

3.1 Definition of composite indicators

Composite indicators are used to aggregate indicator information to a lower dimension. We refer to Münnich and Seger (2014) for a formal derivation. In the following, we restrict ourselves to a linearly weighted aggregation. Well-being at local level is assessed by a composite indicator, comprising $q = 1, \dots, Q$ sub-indicators I , which is calculated as

$$CI_d = \sum_{q=1}^Q w_q \cdot I_{qd}, \quad (1)$$

with w denoting the weights and d the area of interest. Our composite indicator at district level includes twelve sub-indicators: (1) single parent households (%), (2) unemployment (%), (3) youth unemployment (%), (4) vacant dwellings (%), the average distance to (5) primary schools, (6) libraries, (7) museums, (8) play- and sports grounds and (9) hospitals, respectively, (10) parks, green areas and sport fields (%), (11) forest areas (%) and (12) water areas (%). At grid cell level the composite indicator comprises 10 sub-indicators. Sub-indicators (1) to (9) are the same as at district level but as absolute numbers and sub-indicators (10) to (12) are summarised as natural areas approximated by the mirror image of impervious surfaces⁷.

The indicator is constructed using spatial and official data. In our case, unemployment data are only available at city district level and information about the number of vacant dwellings at 100 metre grid cell level. Both unemployment and vacancy data have to be re-scaled in order to analyse well-being at grid cell and district level, respectively. Therefore, methods for changing the scale are required. In addition, the construction of composite indicators requires normalisation and weighting of sub-indicators. In the following, different scaling methods from spatial research and small area estimation are presented in brief, following Rao and Molina (2015) and Zhang *et al.* (2014). Moreover, various normalisation and weighting methods

⁷ The imperviousness raster data are summarised within the 100 metre census grid cells using the zonal toolset from the spatial analysis toolbox of ArcGIS Pro.

according to OECD *et al.* (2008) are introduced. In particular, possible sources of uncertainty, which arise from scaling methods, selection of sub-indicators, data normalisation and weighting choices are considered. These sources of uncertainty are regarded as construction steps of composite indicators with each step offering several selection choices, also called triggers. Each possible combination results in a different composite indicator (see equation (1)).

3.2 Scaling of data

As described above, the data come from different data sources. The disaggregated census data on vacant dwellings and single parents are available at 100 metre grid cell level, HLR imperviousness data at 20 metre resolution and socio-demographic data at city district level. In addition, we have georeferenced data on schools, museums, play- and sports grounds, hospitals and libraries. In order to conduct analyses at 100 metre grid cell level and city district level, the scales have to be harmonised using up- or downscaling techniques. In the following, several up- and downscaling methods from the fields of geostatistics and small area estimation are introduced.

3.2.1 Upscaling methods

Upscaling refers to the aggregation of fine-resolution input data to coarse-resolution output data. Information on vacant dwellings is only available at 100 metre grid cell level. In order to conduct an analysis of well-being at city district level, this information needs to be upscaled. The selection of the method is determined by characteristics of the input data. In the following, different upscaling methods are introduced.

Upscaling methods include aggregation using the (weighted mean), block-kriging, the methods of random selection, median rule, mid-point rule, majority rule and reclassification of coarsened images are needed (Yang and Merchant, 1997; Zhang *et al.*, 2014, pp. 219ff.). Random selection assigns randomly a value from the fine-resolution grid to the aggregated coarse-resolution cell. The data at 100 metre census grid cells level are interpreted as point data, which are randomly assigned to represent the district. The selection probability of a fine-resolution grid cell to represent the aggregated coarse-resolution cell is proportional to its occurrence. Thus, the method of random selection is more

likely to preserve the structure than the majority rule, widely which is used. The majority rule chooses the most frequent value from the fine-resolution grid cells within the coarse-resolution area and qualifies the coarse-resolution output accordingly. If two or more classes occur with the same frequency, it is drawn randomly. The median rule attaches the (weighted) median value of the fine-resolution grids within the coarse-resolution area as aggregated value to the district. The degradation-reclassification approach applies an averaging degradation process and reclassifies the resulting images to obtain a coarse-resolution image with the same characteristics as the input images except for the resolution (Yang and Merchant, 1997; Zhang *et al.*, 2014, pp. 219ff.).

In this study the average vacancy rate is determined by calculating the weighted mean and applying block-kriging. In the first method, the vacancy rate is calculated by weighting the number of vacant dwelling and total dwelling in grid cells which are intersected by district boundaries with the high-resolution impervious data. Assuming that the number of dwellings correlates with the impervious surface and is evenly distributed, the dwellings are distributed proportionally to the impervious surface at 100 metre grid cell. The second method is based on Zhang *et al.* (2014, pp. 109ff.) and Zhang and Yao (2008). Block-kriging (or point-to-area kriging) utilises additional information, *e.g.* spatial dependence in the underlying distribution to estimate the mean value of a variable Z for a predefined large area, *e.g.* districts. Block kriging assumes that the mean value of a random variable over a block v_x centred at location x is defined as the average of all n_p random variable points $Z(x_\beta)$ which discretise the block

$$Z(v_x) = \frac{1}{n_p} \sum_{\beta=1}^{n_p} Z(x_\beta).$$

The simplest form of kriging derives kriging weights based on the criteria of unbiasedness and minimum variance of the estimator. Stationarity of the mean and covariance of the problem domain is assumed. The estimator for Z over a block v_x is a linear combination

$$\hat{z}(v_x) = m_Z + \sum_{\beta=1}^{n_p} \lambda_\beta [z(x_\beta) - m_Z] \quad (2)$$

holds, with m_Z being the known stationary mean and λ_β denoting kriging weights (see Zhang and Yao, 2008). The average prediction error $\hat{z}(v_x) - z(v_x)$ is set to be zero

$$E[\hat{z}(v_x) - z(v_x)] = m_Z + \sum_{\beta=1}^{n_p} \lambda_{\beta} \{E[z(x_{\beta})] - m_Z\} - E[z(v_x)] = 0.$$

The variance of prediction is given as

$$\begin{aligned} \sigma_Z^2(v_x) &= \text{var} [\hat{z}(v_x) - z(v_x)] \\ &= \text{var} \left\{ \sum_{\beta=1}^{n_p} \lambda_{\beta} [z(x_{\beta}) - m_Z] - [z(v_x) - m_Z] \right\}. \end{aligned}$$

Defining

$$\begin{aligned} \boldsymbol{\lambda}^* &= \begin{bmatrix} \boldsymbol{\lambda} \\ -1 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n_p} \\ -1 \end{bmatrix} \\ \mathbf{Z}^* &= \begin{bmatrix} \mathbf{Z}(\mathbf{x}_{\beta}) - \mathbf{m}_Z \\ z(v_x) - m_Z \end{bmatrix} = \begin{bmatrix} z(x_1) - m_Z \\ \vdots \\ z(x_{n_p}) - m_Z \\ z(v_x) - m_Z \end{bmatrix}, \end{aligned}$$

the reduced form can be written as

$$\begin{aligned} \text{var}[\boldsymbol{\lambda}^{*T} \mathbf{Z}^*] &= [\boldsymbol{\lambda}^T \quad -1] \begin{bmatrix} \mathbf{cov}(\mathbf{x}_{\beta}) & \mathbf{cov}(\mathbf{x}_{\beta}, v_x) \\ \mathbf{cov}(\mathbf{x}_{\beta}, v_x)^T & \text{var}[z(v_x)] \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ -1 \end{bmatrix} \\ &= \boldsymbol{\lambda}^T \mathbf{cov}(\mathbf{x}_{\beta}) \boldsymbol{\lambda} - 2\boldsymbol{\lambda}^T \mathbf{cov}(\mathbf{x}_{\beta}, v_x) + \text{var}[z(v_x)] \end{aligned}$$

with $\mathbf{cov}(\mathbf{X}_{\beta}, v_x)$ being the point-to-block covariance vector. The weight vector

$$\boldsymbol{\lambda} = \mathbf{cov}(\mathbf{x}_{\beta})^{-1} \mathbf{cov}(\mathbf{x}_{\beta}, v_x)$$

minimises the prediction variance and solves equation (2). Thus, the block kriging estimator is given as

$$\begin{aligned} \hat{z}(v_x) &= m_Z + \sum_{\beta=1}^{n_p} \lambda_{\beta} (z(x_{\beta}) - m_Z) \\ &= m_Z + \boldsymbol{\lambda}^T (\mathbf{Z}(\mathbf{x}_{\beta}) - \mathbf{m}_Z) \end{aligned}$$

and the kriging variance as

$$\sigma_Z^2(x) = \text{var}[z(v_x)] - \boldsymbol{\lambda}^T \mathbf{cov}(\mathbf{x}_{\beta}, v_x),$$

where $\text{var}[z(v_x)]$ denotes the average covariance within the block being predicted (Zhang and Yao, 2008, p. 183; Zhang *et al.*, 2014, pp. 109ff.).

3.2.2 Downscaling methods

Downscaling is the inverse of upscaling and refers to the conversion of coarse-resolution data to fine-resolution data. In our case, data on unemployment and youth unemployment are available at district level. In order to conduct an analysis of well-being at 100 metre grid cell level, this information needs to be downscaled. Several downscaling methods from the small area literature and geospatial research are available, which are presented briefly in the following.

The method of area-to-point kriging is in contrast to block kriging as explained above. Assuming that the grid cell is represented by its centroid, point values for each grid cell are estimated based on available data at district level. For the derivation of area-to-point kriging, we refer to Zhang *et al.* (2014, pp. 120ff.) and Kyriakidis (2004). Area-to-point kriging is not used in this study as the quality of the results at 100 metre grid cell level is distorted as very coarse information is converted to very fine information.

Further examples of downscaling methods include geographic centroid assignment, areal weighting, dasymetric mapping and regression methods. Geographic centroid assignment assigns a representative value for the district to its centroid. The values for the large area are then assigned to the 100 metre grid cell centroids using the distances between the large area centroid and small area centroids as weights. Grid cells that are uninhabited are not included in the analysis. Whether a grid cell is uninhabited is determined on the basis of the number of houses classified by OpenStreetMaps. Areal weighting approaches are based on cartographic techniques. Simple area-weighted interpolation determines weights based on the percentage of the overlapping large area and small area assuming that the socio-demographic variable of interest (*e.g.* population, unemployment) is evenly distributed within the large area. Thus, areal weighting methods rely on the assumption of homogeneity within each large area (see *e.g.* Goodchild and Lam, 1980). Dasymetric mapping uses auxiliary data such as remote sensing data (*e.g.* high-resolution imperviousness data). High-resolution imperviousness data

are used to better depict the distribution of the socio-economic variable of interest in the large area and, thus, accounts for the fact that some parts of the area of interest might not be populated assuming that impervious surfaces approximate population and unemployment is evenly distributed across the population (see *e.g.* Eicher and Brewer, 2001). The downscaling methods described above rely on interpolation. Alternatively, the conversion of coarse-resolution data to fine-resolution data can be achieved by regression methods, which establish a relationship between different scales (see *e.g.* Fernandes *et al.*, 2004; Martinez *et al.*, 2009; Wu and Li, 2009).

Alternatively to these methods from the geoscientific research, simple approaches from the discipline of small area estimation can be applied as downscaling techniques. Small area estimation (SAE) generally deals with situations in which survey data is to be evaluated on a highly disaggregated level. In this case, the sample size in some or many areas is typically so small, that traditional direct estimators that only rely on the sample data in a specific area lack accuracy. The strategy then is to use indirect estimation methods, that borrow information from other areas to stabilise estimation.

There is a broad range of different approaches to small area estimation. Which method is suitable, crucially depends on the availability of data, both with respect to the variable of interest and auxiliary information, and the type of the target information. An introduction into the field and a comprehensive overview can be found in the monograph of Rao and Molina (2015). Recent developments are also presented in Pfeiffermann (2013). An introductory overview in German language is given by Münnich *et al.* (2013). We focus on approaches that might be of relevance in the context of measuring well-being on a low aggregation level. We first look at some SAE approaches with minimal data requirements on the target resolution level and then present some relevant methods from the broader range of SAE techniques, that opens up if additional information is available. We finally focus on the SPREE-estimator and related extensions because this might prove to be a relevant approach in the context of well-being indicators, which in many cases rely on categorical variables.

Before presenting the selected small area methods, note that small area estimation as a part of survey statistics generally deals with data that was obtained in a random sample in order to obtain reliable statistics for a

larger population from which the sample was taken. Most procedures and expressions presented below can however also be applied in cases where data for the entire target population, *i.e.* register data, is available for a larger area, for example obtained from administrative sources. In these cases the synthetic approaches (while not being *estimators* in the classical sense then) can still be suitable ways to deduce statistics on the level of smaller areas (see *e.g.* Rao and Molina, 2015, Chapter 3.2.3).

If no sample information for the variable of interest is available at the targeted fine-resolution level, the range of feasible approaches is largely restricted and only some very simple synthetic estimators can be considered. Generally, a synthetic estimator uses a reliable estimator for a larger area to derive an indirect estimator for smaller areas within this larger area, relying on the assumption that the small areas share the characteristics of the larger area (Rao and Molina, 2015). If no area information is available, a very simple naive synthetic estimator for the mean (or proportion) \hat{Y}_d in a small area d is given by

$$\hat{Y}_d^{syn} = \hat{Y} \quad d = 1, \dots, D,$$

where \hat{Y} is the direct estimator of the large area. This naive synthetic estimator relies on the implicit assumption that the small area means are equal to the large area mean. Obviously, it is highly inadequate when this strong assumption is inappropriate. If a suitable auxiliary variable x is available, the ratio-synthetic estimator for the domain total Y_d can be obtained as:

$$\hat{Y}_d^{rs} = X_d \frac{\hat{Y}}{\hat{X}} \quad d = 1, \dots, D,$$

where X_d is the known area-level total and \hat{X} is the population or larger-area total estimated from the same sample as \hat{Y} . This estimator relies on the assumption that the rate $R_d = Y_d/X_d$ is approximately equal to the overall ratio $R = Y/X$ and the bias might be large if this assumption is not fulfilled (see Rao and Molina, 2015, Section 3.2). We use the ratio-synthetic estimator to obtain unemployment numbers at 100 metre grid cell. The known area-level totals X_d , here inhabitants, come from the 100 metre census grid cells and Y and X , inhabitants and unemployed people on the district level, from the City of Cologne (2017). It has to be noted that some grid cells can not be allocated

to one district only. In some cases, the 100 metre grid cells are intersected by district borders. We assign the grid cells to the district with the largest intersection.

So far, we presented methods with minimal data requirement on the targeted fine-resolution level. Typically in SAE problems, a sub-sample (albeit small) is available in at least most of the areas. The best-known and regularly applied methods use this information. Synthetic approaches that use sample information on the target level are the regression synthetic estimator and the GREG-synthetic estimator (Rao and Molina, 2015, Section 3.2). More importantly, the most common small area models, which as special cases of a General Linear Mixed Model employ an explicit statistical model to obtain small area estimates, become feasible. A large part of Rao and Molina (2015) is dedicated to these models.

A special problem of SAE is that of estimating cell counts (or proportions) of a categorical variable. Assume that the counts are arranged in a two-way table, where each row contains a vector of frequencies for the p categories of the variable of interest in a given area. Following Hernandez, we call this arrangement a population composition and denote it by Y (Hernandez, 2016). Assume that a sample of the target population is available that – while yielding reliable estimates for the margins of the compositions – is too small to obtain accurate estimates of cell frequencies. Further, some proxy composition X is available. This can, for example, be the result from a previous census that needs updating.

Generally, structure preserving estimators provide estimates of the cell frequencies by adjusting them to the known margins while at the same time in some way preserving the association structure, *i.e.* the relationship between rows and columns, observed in the proxy composition X . In this adjustment, several assumptions on the relationship between the association structure in X and Y can be used. For estimation, typically, the method of iterative proportional fitting (IPF) (Deming and Stephan, 1940) is employed. Note that calibration to known margins in these approaches is an inherent feature of the estimation process.

The basic structure preserving estimator (SPREE) was introduced by Purcell and Kish (1980). It makes the simple assumption that the association structure of X and Y is equal. Let $Y_{d,a}$ denote the count for area d , $d = 1, \dots, D$

and category a , $a = 1, \dots, p$. Further, we use Y_{d+} and Y_{+a} to denote the known row- and column-margins, respectively. Assume that a proxy composition X with the same dimensions as Y is available. The aim is, to obtain estimates $\hat{Y}_{d,a}$ that minimise the distance between the cell counts and fitted values under constraints implied by the margins. As this optimisation problem cannot be solved in closed form, an estimate for Y is obtained iteratively by IPF, applying the following procedure (see Hernandez, 2016; Agresti, 2013, Chapter 9.7.2):

$$\text{Step 1:} \quad \hat{Y}_{d,a}^{(1)} = X_{d,a} \frac{Y_{+a}}{X_{+a}} \quad (3)$$

$$\text{Step 2:} \quad \hat{Y}_{d,a}^{(2)} = \hat{Y}_{d,a}^{(1)} \frac{Y_{d+}}{\hat{Y}_{d+}^{(1)}} \quad (4)$$

$$\text{Step 3:} \quad \hat{Y}_{d,a}^{(3)} = \hat{Y}_{d,a}^{(2)} \frac{Y_{+a}}{\hat{Y}_{+a}^{(2)}} \quad (5)$$

Step 2 and 3 are repeated until convergence. The algorithm converges to the optimal solution, *i.e.* it minimises the distance between cell counts and fitted values according to the Kulback-Leibler discrimination information measure $\sum_d \sum_a Y_{ia} \log \frac{Y_{d,a}}{\hat{Y}_{d,a}}$ (Ireland and Kullback, 1968).

A closely related approach is presented by Dostal *et al.* (2016), who suggest an extension of the SPREE with an alternative distance function. This has proven to be a suitable approach in the case of very small domains (Dostal *et al.*, 2016). If sample estimates for the inner cells are available, more elaborate methods become feasible. Zhang and Chambers (2004) propose a Generalised SPREE (GSPREE), that assumes a proportional relationship between the association structure of the target compositions and its proxy. They further present a version of this approach that allows for cell-specific random effects. Hernandez (2016) presents a Multivariate SPREE (MSPREE), an extension of the GSPREE that allows for further flexibility regarding the structural assumption.

Due to restrictions in data availability on the very fine resolution level of districts and grid cells, the more elaborate approaches presented here, cannot be applied in the study at hand. As they open up the opportunity to account for the socio-demographic structure in downscaling, we think that it is worth to pursue them further. This will require close cooperation with official statistics.

3.3 Construction steps of composite indicators

The data need to be normalised prior to any aggregation. Different methods, such as normalised ranking, standardisation or min-max methods, are available. Normalised ranking is the simplest normalisation method. It evaluates the performance of the area of interest in the subsequent dimension according to its relative position

$$I_{qd} = \frac{\text{Rank}(x_{qd})}{D},$$

where x_{qd} is the value of sub-indicator q of area d and D the total number of areas.

Standardisation ensures that the sub-indicators have zero mean and a standard deviation equal to one, *i.e.*

$$I_{qd} = \frac{x_{qd} - x_{qd=\bar{d}}}{\sigma_{qd=\bar{d}}},$$

where $x_{qd=\bar{d}}$ and $\sigma_{qd=\bar{d}}$ denote the average and standard deviation across countries, respectively. Sometimes the composite indicator is adjusted, *e.g.* by weights, as outliers distort the composite indicator.

The min-max methods subtracts the minimum value and divides the difference by the range of the sub-indicator values

$$I_{qd} = \frac{x_{qd} - \min_d(x_q)}{\max_d(x_q) - \min_d(x_q)}$$

and, thus, sub-indicators range between zero and one (OECD *et al.*, 2008).

The relative importance of the sub-indicators is determined by the attached weights. The most frequently used weighting technique is equal weighting. It assigns the same weight to all sub-indicators, implying that all sub-indicators are equally important. However, due to potential correlation between sub-indicators, equal weighting does not guarantee an equal contribution of the sub-indicators to the composite indicator. Alternatively, weights based on statistical models such as the principal component analysis (PCA) or on expert opinions can be applied. The goal of PCA is to determine how different variables change in relation to each other and how these variables are associated. Correlated variables are converted into a new set of uncorrelated

variables using a covariance matrix or correlation matrix. PCA involves finding the eigenvalues of this covariance matrix (OECD *et al.*, 2008; Nardo *et al.*, 2005). Weights based on PCA are constructed following Nicoletti *et al.* (2000). The indicators with the highest squared factor loading are grouped into intermediate composites with the squared factor loadings summed to unity as weight. These intermediate indicators are aggregated by assigning the proportion of explained variance as weights to the intermediate composites. An exemplary application of the construction of weights based on PCA can be found in OECD *et al.* (2008, p. 90f.). The PCA is conducted applying the `prcomp` function in R.

4. Implementation of the sensitivity analysis

4.1 Theoretical framework for sensitivity analyses

In order to assess the robustness of the resulting composite indicator with respect to the scaling schemes, the normalization method and the choice of weights, a sensitivity analysis is conducted. Sensitivity analyses aim at determining the effect of a change in the input factors on the variable of interest. In this article, the output variation in the composite indicator is caused by the choices in the construction steps. The sensitivity analysis is based on a variance decomposition method as described by Saltelli *et al.* (2000, 2008) as an extension of the original approach proposed by Sobol (1993) and Homma and Saltelli (1996).

The composite indicator for well-being is calculated according to equation (1). The different construction steps (triggers), *i.e.* scaling techniques, normalisation methods and the choice of weights, have a direct impact on the output of the composite indicator. Variance decomposition within the scope of sensitivity analyses indicates how much each construction step contributes to the total variance in the result. Following Saltelli *et al.* (2008), the composite indicator CI is understood as a function of the k uncertain construction alternatives T :

$$CI = f(T_1, \dots, T_k).$$

For mutually independent input factors T , the total or unconditional variance of the output, $V(CI)$, can be decomposed into

$$V(CI) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \dots + V_{12\dots k},$$

where

$$V_i = V(f_i(T_i)) = V[E(CI|T_i)]$$

and

$$V_{ij} = V(f_{ij}(T_i, T_j)) = V[E(Y|T_i, T_j)] - V_i - V_j.$$

V_i is the first-order or main effect of T_i on CI , i.e. the individual contribution of T_i to the variance of the output. The second-order effect V_{ij} represents the joint effect of T_i and T_j . Setting these effects in relation to the total variance, a first-order sensitivity index of T_i on CI is obtained:

$$S_i = \frac{V[E(CI|T_i)]}{V(CI)}.$$

S_i is in the interval $[0, 1]$ with values close to 1 indicating input factors with a large effect on the output. Correspondingly, higher order sensitivity indices, e.g. the second order effect S_{ij} , are derived by setting the higher order effects in relation to the total variance. Additionally, one might be interested in the total contribution of a specific input factor T_i to the output, i.e. the sum of its main effect and all relevant higher-order effects. For T_1 and three insecure input factors a corresponding total sensitivity index is for example given by (Münnich and Seger, 2014; Saltelli *et al.*, 2008)

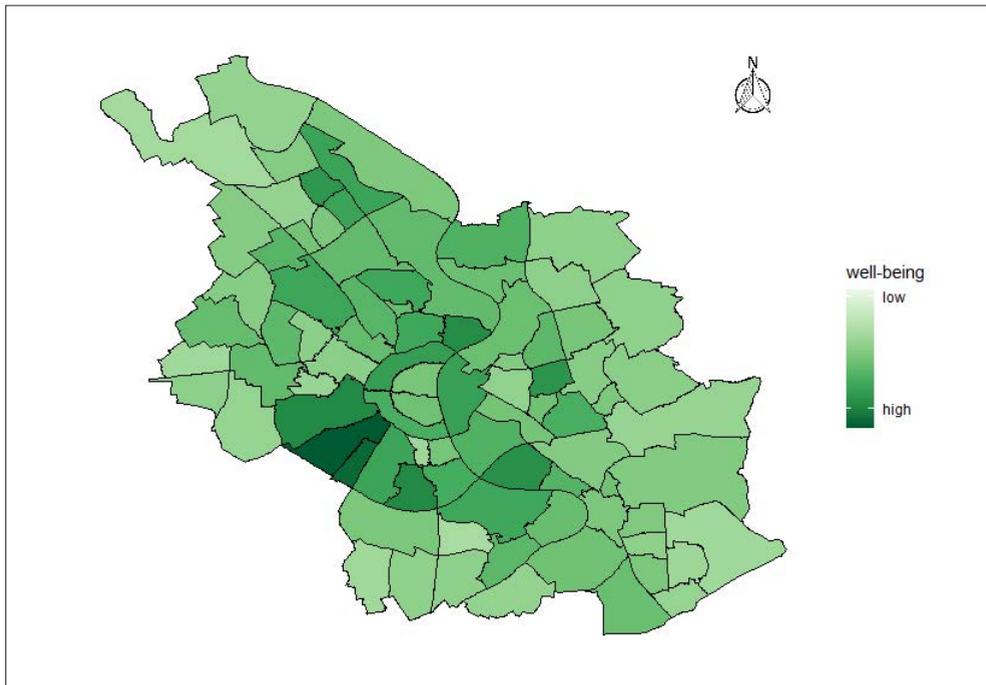
$$S_{tot_1} = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3}.$$

For detailed explanations we refer to Saltelli *et al.* (2008, pp. 20-21 and 155-174). The sensitivity analysis is performed using the R-package multisensi (see Bidot *et al.*, 2018).

4.2 Composite indicator of well-being at district level

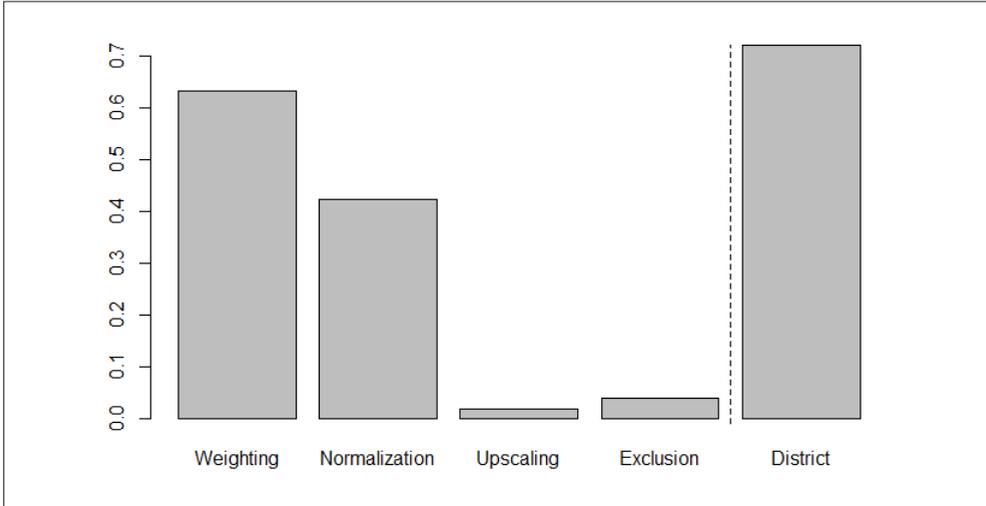
As explained above, the construction of composite indicators comprises scaling of the data, selection and normalisation of sub-indicators and the choice of weights. In the sensitivity analysis, we consider two upscaling methods (weighted mean, block-kriging), three normalisation schemes (min-max method, ranking and standardisation), two weighting possibilities (based on PCA and equal weighting) and exclusion of one indicator. Each of the twelve indicators considered at district level is left out once and further all indicators are taken into account, resulting in 13 exclusion options. This results in 156 possible combinations per district.

Figure 2 illustrates the composite well-being indicator for the City of Cologne at district level based on all twelve indicators and using the weighted mean as upscaling method, the min-max method for normalisation and weights resulting from PCA.

Figure 2 - Well-being indicator values at district level

Source: Own illustration

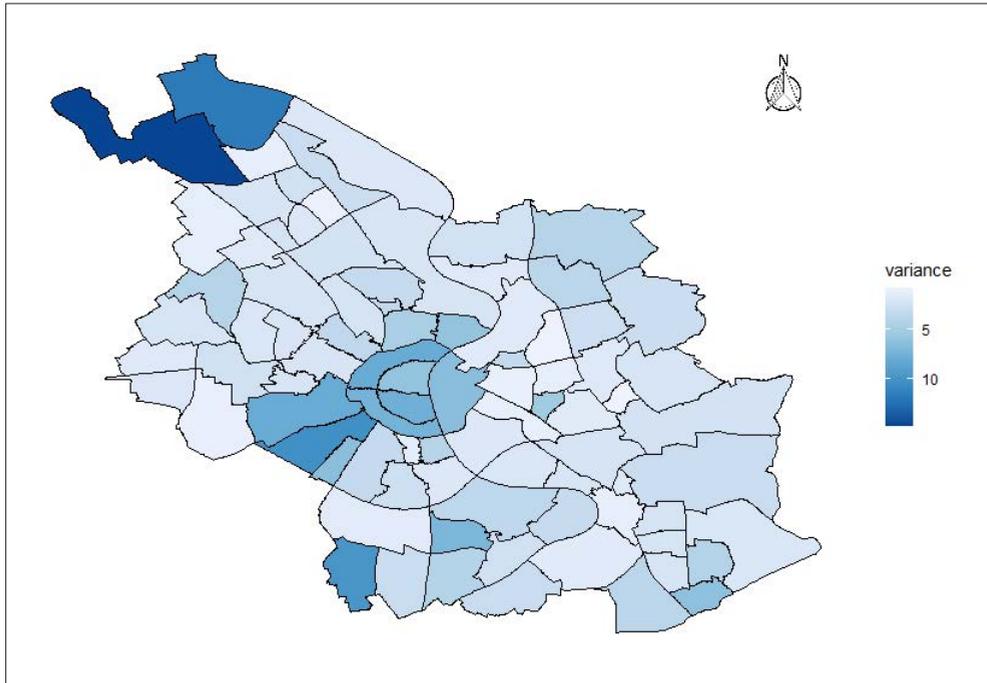
The uncertainty in the construction of the composite indicator for well-being is assessed by a sensitivity analysis. Sensitivity analyses study how much each source of uncertainty contributes to the output variance. Figure 3 presents the total-order effect with scaling, selection of sub-indicators, normalisation and weighting as input factors. To set these results into relation, we also depicted the variation of indicator results between districts as a reference.

Figure 3 - Total-order effects of the sensitivity analysis at district level

Source: Own illustration

From Figure 3 we observe that the scaling methods have the lowest impact on the output variance and that the construction decision with the largest effect is the choice of the weighting method. It can also be taken from this plot that most of the variability in results is still due to district identity, showing that there is an actual regional heterogeneity of well-being which is not offset by construction decisions.

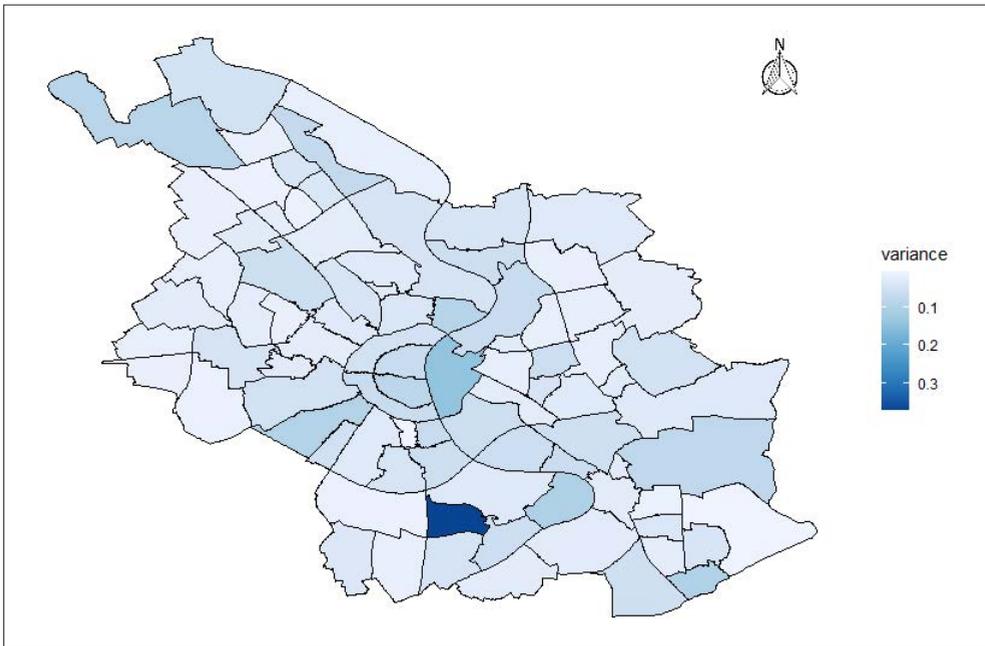
The variance of the composite indicator depending on the construction is depicted in Figure 4.

Figure 4 - Conditional variance of the composite indicators at district level

Source: Own illustration

The results of the sensitivity analysis (Figure 3) suggest that the variance within the districts is largely caused by the weighting method. For this reason, the variance conditional on the weighting method (here: PCA) is shown in Figure 5.

Figure 5 - Conditional variance of the composite indicators using weights resulting from PCA at district level



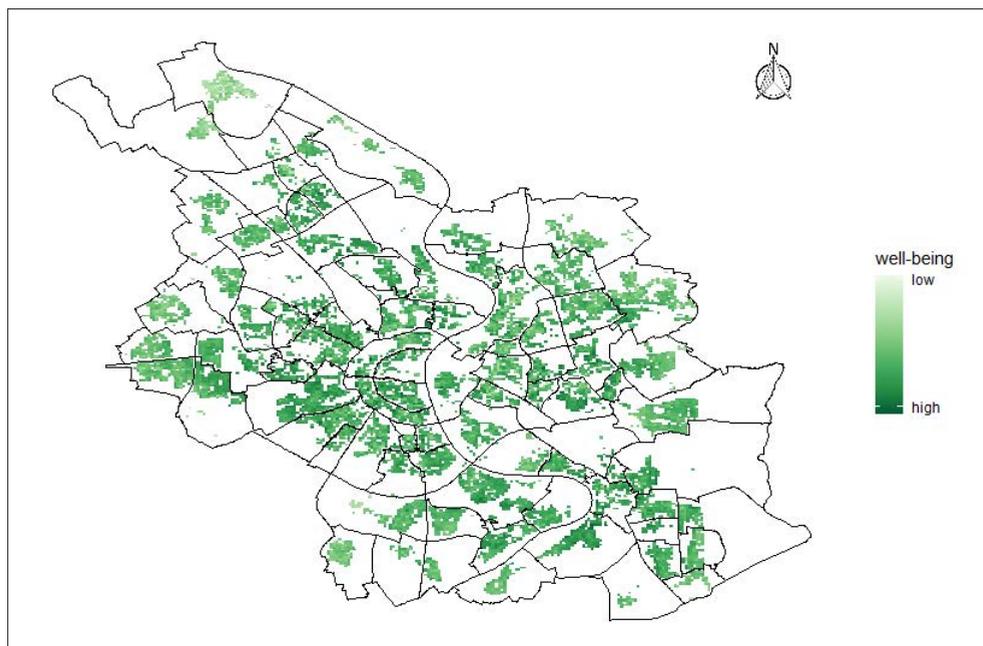
Source: Own illustration

As mentioned above, if various sub-indicator form a group, equal weighting might effectively result in unequal weights. For example, distances to primary schools, museums, hospitals, libraries and play- and sports grounds could be grouped into a single sub-indicator infrastructure. Assigning a weight of $w_q = 1/12$ to each of the q sub-indicators, results in the infrastructure related sub-indicators having a weight of $5/12$ in total. This might explain why the variance is particularly high in the north-western and central districts. The infrastructure in the city centre is usually better developed than in the rest of the city. The central distances have the shortest distances and the north-western districts have the highest distances. It should be noted that the results might also be distorted as closer hospitals etc. can be located outside Cologne. However, in this study we concentrate exclusively on the urban area of Cologne. It would be desirable for future investigations to group the sub-indicators before weighting them. Furthermore, it would be sensible to include influential factors beyond the border of the region of interest in the analysis of well-being.

4.3 Composite indicator of well-being at grid cell level

The composite indicator at 100 metre grid cell level is illustrated in Figure 6 based on all sub-indicators normalised by standardisation and weights resulting from PCA. The unemployment numbers are downscaled by dasymetric mapping.

Figure 6 - Well-being indicator values at grid cell level

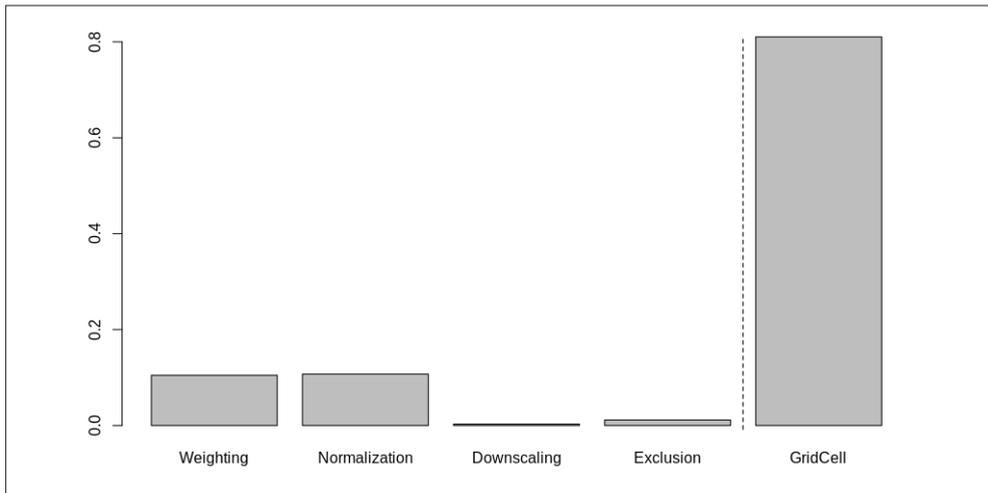


Source: Own illustration

The composite indicator at 100 metre grid cell level has 264 different input combinations per grid cell. We consider four downscaling methods (centroid assignment, dasymetric mapping, areal weighting and the ratio-synthetic estimator), three normalisation schemes (min-max method, ranking and standardisation), two weighting possibilities (based on PCA and equal weighting) and the exclusion of one indicator in each step. This results in a decision matrix of size 2.589.048 5 (number of construction steps). Each row has a different composite indicator as result. In order to be able to handle the size, 2.000 construction possibilities were drawn based on LP-Tau quasi-

random numbers on the interval $[0, 1]$ (Münnich and Seger, 2014; Saltelli *et al.*, 2000). Figure 7 quantifies the impact of the normalisation scheme, weighting choice, downscaling method, exclusion of sub-indicators and the grid cell itself on the output variance. Note again that grid cell identity is not a trigger per se. It was included in the analysis to set the effect of construction decisions in relation to the regional heterogeneity in the results.

Figure 7 - Total-order effects of the sensitivity analysis at grid cell level



Source: Own illustration

As expected, the sensitivity analysis of the well-being indicator at 100 metre resolution reveals that grid cells have the largest impact on the output variance, followed by the weighting and normalisation methods. However, the results of the sensitivity analysis have to be treated with caution as it is based on 2.000 construction possibilities only.

At both district and grid cell level the differences between the areas of interest play a dominant role in the sensitivity analysis. This is even more pronounced on the very fine resolution level of grid cells. Altogether, the results indicate that there is a relevant heterogeneity of well-being at the regional level and, thus, confirm that efforts to measure this multidimensional concept at the local level are worth-while. A second central conclusion is that, even if composite indicators are sensitive to construction decisions, the effect of these decisions does not “mask” the actual regional differences in indicator results.

5. Concluding remarks

In this article, we explore the potentials of using remote sensing data for local level estimation of well-being. So far, the analysis of well-being is mainly focussed on the country-level. However, differences in central dimensions of well-being, such as material living conditions or the preconditions of social interaction and the quality of leisure time, often exist at street level. Therefore, we combine survey data and remote sensing data to enable an analysis on the very low aggregation level of city districts and 100 metre grid cells. We present different sources of remote sensing data and create infrastructure-related sub-indicators in the composite well-being indicator using tools from the geosciences. Survey data are usually provided at administrative levels, whereas remote sensing data are available at small scale resolutions. Therefore, different upscaling and downscaling methods are introduced. We determine a composite indicator for well-being and quantify the impact of the different scaling techniques and other construction decisions by means of a sensitivity analysis with the scaling techniques, the normalisation scheme, the weighting methods and the exclusion of sub-indicators as uncertainty factors.

At district level, the incorporation of remote-sensing data is very promising. We can show that the upscaling methods only account for a minor proportion of the output variance. Following our application, the weighting scheme is the construction decision with the largest impact on indicator results. This can be attributed to the fact that, among other, equal weights are assigned to each sub-indicator resulting in an actual overweight of the infrastructure-related indicators. For future research, it would, therefore, be desirable to conduct the sensitivity analysis with grouped sub-indicators again.

The analysis at grid cell level is methodologically more challenging. The data at 100 metre resolution contain many empty cells as values lower than three are not published due to confidentiality reasons. Moreover, the change of resolution from district to 100 metre grid cell level is large, which comes at the cost of quality of the estimates at grid cell level. Generally, data availability on this very fine resolution level is largely restricted, so that more complex – and probably better – downscaling techniques could not be applied. For future research, we envisage a close collaboration with official statistics in order to build a data basis that will enable more elaborate approaches from the field of small area statistics. At both district and grid cell

level, the variation of well-being between the areas of interest was included in the analysis as a reference, *i.e.* to set the variability introduced through the construction decisions in relation to the actual heterogeneity between districts. At both levels, this variation between regional entities was the most relevant source of variability. There, thus, is a relevant heterogeneity on these low resolution levels which confirms our motivating notion that the micro-location matters. Further, even if composite indicators are sensitive to construction decisions, the actual differences between regional entities are not offset by these construction decisions. This study can be seen as a feasibility study showing that further research in this area could open up many possibilities. In particular, the downscaling methods have to be further developed. Moreover, alternative remote sensing data, such as land surface temperature, could be included as an environmental variable to describe human heat stress. All in all, the incorporation of remote sensing data has a huge potential for analyses of living conditions at local level and should be further investigated.

Acknowledgments

The research was conducted within the MAKSWELL project (<https://www.makswell.eu>), funded by the EU under Horizon 2020 (H2020-SC6-CO-CREATION-2017; 2nd and 5th author). Further collaboration took place with the MikroSim project (DFG research unit FOR 2559, <https://www.mikrosim.uni-trier.de>), funded by the German Research Foundation (3rd and 5th author), the REMIKIS project, funded by the Nikolaus Koch Stiftung (1st and 5th author), and the Trier Centre of Sustainable Systems, funded by the Rhineland-Palatinate Research Initiative (4th and 6th author).

We thank Professors Maria Giovanna Ranalli and Li-Chun Zhang for the invitation for this paper. Further, we thank Professor Alessandra Petrucci for inviting the 5th author to the ITACOSM conference.

Finally, we thank the editors and reviewers for the very positive and valuable comments that helped improving the readability of the paper.

References

Agresti, A. 2002. *Categorical Data Analysis*. Hoboken, NJ, U.S.: John Wiley & Sons.

Bidot, C., H. Monod, and M.-L. Taupin. 2018. *A Quick Guide to multisensi, an R Package for Multivariate Sensitivity Analyses*. <https://cran.r-project.org/web/packages/multisensi/vignettes/multisensi-vignette.pdf>.

City of Cologne. 2014. Offene Daten Köln. <https://www.offenedaten-koeln.de/dataset> (September 6th 2019).

City of Cologne. 2017. *Statistische Daten - Thematische Karte*. <https://www.stadt-koeln.de/politik-und-verwaltung/statistik/statistische-daten-thematische-karte> (September 6th 2019).

Copernicus. 2019a. *High Resolution Layers*. <https://land.copernicus.eu/pan-european/high-resolution-layers> (October 21st 2019).

Copernicus. 2019b. *Imperviousness*. <https://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness> (October 21st 2019).

Deming, W. E., and F. F. Stephan. 1940. “On a Least Squares Adjustment of a Sampled Frequency

Table When the Expected Marginal Totals Are Known”. *The Annals of Mathematical Statistics* 11(4): 427–444.

Dostal, L., S. Gabler, M. Granninger, and R. Münnich. 2016. “Frame Correction Modelling with Applications to the German Register-Assisted Census 2011”. *Scandinavian Journal of Statistics*, 43(3): 904–920.

Easterlin, R.A. 1974. “Does Economic Growth Improve the Human Lot? Some Empirical Evidence”. In David, P.A., and M.W. Reder (Eds.). *Nations and Households in Economic Growth*: 89–125. Cambridge, MA, U.S.: Academic Press.

Eicher, C.L., and C.A. Brewer. 2001. “Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation”. *Cartography and Geographic Information Science*, 28(2): 125–138.

Engstrom, R., J. Hersh, and D. Newhouse. 2017. “Poverty From Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being”. *Policy Research Working Paper 8284*, World Bank.

European Commission. 2019. *Inspire Data Specifications*. <https://inspire.ec.europa.eu/training/inspire-data-specifications> (November 6th 2019).

European Foundation for the Improvement of Living and Working Conditions - Eurofound. 2020. “What Makes Capital Cities the Best Places to Live?”. *European Quality of Life Survey 2016 series*. Luxembourg: Publications Office of the European Union.

Eurostat. 2019. *Quality of Life Indicators - Measuring Quality of Life*. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/gdp-and-beyond/quality-of-life> (August 13th 2019).

Eurostat. 2017. *Methodological Manual on City Statistics*. Luxembourg: Publications Office of the European Union. <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-006> (May 5th 2020).

Eurostat. 2020. *Sustainable Development Indicators. SDG 3: Good health and well-being*. <https://ec.europa.eu/eurostat/de/web/sdi/good-health-and-well-being> (May 6th 2020).

Federal Statistical Office and the Statistical Offices of the Länder. 2018. *Ergebnisse des Zensus 2011 zum Download - erweitert*. <https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html> (September 6th 2019).

Fernandes, R.A., J.R. Miller, J.M. Chen, and I.G. Rubinstein. 2004. “Evaluating Image-Based Estimates of Leaf Area Index in Boreal Conifer Stands Over a Range of Scales Using High-Resolution CASI Imagery”. *Remote Sensing of Environment*, 89(2): 200–216.

Geofabrik GmbH and OpenStreetMap contributors. 2018. *Download OpenStreetMap Data for this Region: Regierungsbezirk Köln*. <http://download.geofabrik.de/europe/germany/nordrhein-westfalen/koeln-regbez.html> (September 6th 2019). License: Open Database License.

Ghosh, T., S.J. Anderson, C.D. Elvidge, and P.C. Sutton. 2013. “Using Night time Satellite Imagery as a Proxy Measure of Human Well-Being”. *Sustainability*, 5: 4988–5019.

Goodchild, M.F., and N.S.-N. Lam. 1980. “Areal Interpolation: A Variant of the Traditional Spatial Problem”. *Geo-Processing*, 1: 297–312.

Green, N.E. 1957. “Aerial Photographic Interpretation and the Social Structure of the City”. *Photogrammetric Engineering*, 23: 89–96.

Hernandez, A.L. 2016, October. *Multivariate Structure Preserving Estimation for Population Compositions*. Ph.D. Thesis, University of Southampton.

Homma, T., and A. Saltelli. 1996. “Importance Measures in Global Sensitivity Analysis of Nonlinear Models”. *Reliability Engineering and System Safety*, 52(1): 1–17.

Ireland, C.T., and S. Kullback. 1968. “Contingency Tables With Given Marginals”. *Biometrika*, 55(1): 179–188.

Italian National Institute of Statistics – Istat. 2019. *BES at Local Level*. Roma, Italy: Istat. <https://www.istat.it/en/well-being-and-sustainability/the-measurement-of-well-being/bes-at-local-level> (May 5th 2020).

Kyriakidis, P.C. 2004. “A Geostatistical Framework for Area-to-Point Spatial Interpolation”. *Geographical Analysis*, 36(3): 259–289.

Lo, C.P., and B.J. Faber. 1997. “Integration of Landsat Thematic Mapper and Census Data for Quality of Life Assessment”. *Remote Sensing of Environment*, 62: 143–157.

Martinez, B., F. Garcia-Haro, and F.C. de Coca. 2009. “Derivation of High-Resolution Leaf Area Index Maps in Support of Validation Activities: Application to the Cropland Barrax Site”. *Agricultural and Forest Meteorology*, 149(1): 130–145.

Moretti, A., N. Shlomo, and J.W. Sakshaug. 2019. “Multivariate Small Area Estimation of Multidimensional Latent Economic Well-being Indicators”. *International Statistical Review*, 88(1): 1–28.

Münnich, R., J.P. Burgard, and M. Vogt. 2013. “Small Area-Statistik: Methoden und Anwendungen”. *AStA Wirtschafts- und Sozialstatistisches Archiv*, 6(3): 149–191.

Münnich, R.T., and J.G. Seger. 2014. “Impact of Survey Quality on Composite Indicators”. *Sustainability Accounting, Management and Policy Journal*, 5(3): 268–291.

Nardo, M., M. Saisana, A. Saltelli, and S. Tarantola. 2005. *Knowledge Economy Indicators. Workpackage 5: Input to Handbook of Good Practices for Composite Indicators' Development*. Roma, Italy: Ispra - Joint Research Centre.

Nicoletti, G., S. Scarpetta, and O. Boyland. 2000. Summary Indicators of Product Market Regulation With an Extension of Employment Protection Legislation. *Economics Department Working Papers No. 226 (ECO/WKP(99)18)*. Paris, France: OECD.

Organisation for Economic Co-operation and Development - OECD. 2020. *How's Life? 2020: Measuring Well-Being*. Paris, France: OECD Publishing.

Organisation for Economic Co-operation and Development - OECD, Statistics Directorate and Directorate for Science, Technology and Industry, and Joint Research Centre - JRC of the European Commission in Ispra, Italy, Applied Statistics and Econometrics Unit. 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris, France: OECD Publishing.

Office for National Statistics - ONS. 2019. "Personal Well-Being in the UK: April 2018 to March 2019". *Statistical bulletin*.

<https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/measuringnationalwellbeing/april2018tomarch2019#personal-well-being-by-local-area> (May 5th 2020).

OpenStreetMap contributors. 2019. QuickOSM Plugin (QGIS). <https://github.com/3liz/QuickOSM>. License: Plugin: License GPL Version 2.

Pfeffermann, D. 2013. "New Important Developments in Small Area Estimation". *Statistical Science*, 28(1): 40–68.

Purcell, N.J., and L. Kish. 1980. "Postcensal Estimates for Local Areas (or Domains)". *International Statistical Review/Revue Internationale de Statistique*, 48(1): 3–18.

Rao, J.N.K., and I. Molina. 2015. *Small Area Estimation*. Hoboken, NJ, U.S.: John Wiley & Sons, *Wiley Series in Survey Methodology*.

Saltelli, A., K. Chan, and E.M. Scott. 2000. *Sensitivity Analysis: Gauging the Worth of Scientific Models*. Chichester: Wiley.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Chichester, UK: John Wiley & Sons Ltd.

Shuvo Bakar, K., N. Biddle, P. Kokic, and H. Jin. 2020. “A Bayesian Spatial Categorical Model for Prediction to Overlapping Geographical Areas in Sample Surveys”. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(2): 535–563.

Sobol, I.M. 1993. “Sensitivity Estimates for Nonlinear Mathematical Models”. *Mathematical and Computational Experiments*, 1(4): 407–414.

Stiglitz, J.E., A. Sen, and J.-P. Fitoussi. 2009. Report by the Commission on the Measurement of Economic and Social Progress. <https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report> (August 13th 2019).

United Nations. 2019. *The Sustainable Development Goals Report*. <https://unstats.un.org/sdgs/report/2019/> (September 20th 2019).

Wu, H., and Z.-L. Li. 2009. “Scale Issues in Remote Sensing: A Review on Analysis, Processing and Modeling”. *Sensors*, 9: 1768–1793.

Yang, W., and J.W. Merchant. 1997. “Impacts of Upscaling Techniques on Land Cover Representation in Nebraska, U.S.A”. *Geocarto International*, 12(1): 27–39.

Zhang, J., P.M. Atkinson, and M.F. Goodchild. 2014. *Scale in Spatial Information and Analysis*. Boca Raton, FL, U.S.: CRC Press.

Zhang, J., and N. Yao. 2008. “The Geostatistical Framework for Spatial Prediction”. *Geo-Spatial Information Science*, 11(3): 180–185.

Zhang, L.-C., and R.L. Chambers. 2004. “Small Area Estimates for Cross-Classifications”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2): 479–496.