Data and methods in the environment-migration nexus

a scale perspective

Eklund, Lina; Romankiewicz, Clemens; Brandt, Martin Stefan; Doevenspeck, Martin; Samimi, Cyrus

Published in:
Die Erde – Journal of the Geographical Society of Berlin

DOI:
10.12854/erde-147-10

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Abstract
The relationship between environment and migration has gained increased attention since the 1990s when the Intergovernmental Panel on Climate Change projected climate change to become a major driver of human migration. Evaluations of this relationship include both quantitative and qualitative assessments. This review article introduces the concept of scale to environment-migration research as an important methodological issue for the reliability of conclusions drawn. The review of case studies shows that scale issues are highly present in environment-migration research but rarely discussed. Several case studies base their results on data at very coarse resolutions that have undergone strong modifications and generalizations. We argue that scale-related shortcomings must be considered in all stages of environment-migration research.

Keywords  Environment-migration nexus, spatial & temporal extent, spatial & temporal resolution, scale

1. Introduction
Intensively discussed since the early 1990s, the climate change debate has increased the attention to the relationship between environment and migration, leading to an upsurge in empirical studies (Laczko and Aghazarm 2009, Warner et al. 2010, Black et al. 2011, McLeman 2013, Fussell et al. 2014).
Data and methods in the environment-migration nexus: a scale perspective

In the course of climate change, livelihoods and habitats are expected to be compromised by both long-term changes and increased frequency of extreme weather events, such as droughts, storms and floods. Yet, there is still little agreement about the relevance of the environmental dimension in population movements and its potential to shape future migration (McLeman and Smit 2006, Afifi 2011).

Guided by different research paradigms and resulting in ambiguous conclusions, this disagreement is reflected by a multitude of applied methodological approaches. Whereas some researchers seek to understand the environment-migration nexus through qualitative methods, others use environmental data merged with socio-economic and migration data for statistical analyses of the relationship (Piguet 2010, Bilsborrow and Henry 2012, McLeman 2013, Fussell et al. 2014). A major challenge in this research is the acquisition of appropriate data, which has resulted in a lack of generalizable empirical studies of how environmental change affects migration (Bilsborrow and Henry 2012).

While Laczko and Aghazarm (2009) sweepingly call for ‘better’ data to analyze the environment-migration nexus, we argue that this can only be achieved through an awareness of limitations in the data and methods applied in empirical research thus far. We also argue for a recognition of complications related to scale, in particular.

Several recent methodological reviews of data collection and methods in the field of environment and migration offer a discussion of censuses, surveys, and ethnographies for the representation of migration, and satellite data, surveys and meteorological data for the representation of the environment (Piguet 2010, Bilsborrow and Henry 2012, McLeman 2013, Fussell et al. 2014). However, scale issues have only cursorily been acknowledged as important for the integration of these data.

The neglect of scale is all the more striking if one considers its relevance with a simple example: The West African Sahel is described as affected by desertification, water shortage and hunger, particularly since the great droughts of the 1970s and 1980s. Furthermore, the region’s population is particularly mobile, and migration beyond Sub-Saharan Africa continues to be of increasing importance. At this scale, it seems reasonable to associate increasing migration with deteriorating climatic and environmental conditions (see, e.g., Boko et al. 2007, Hammer 2004). On the other hand, investigations in the same region at the local level illustrate the complex, cumulative and dynamic character of population movements. A micro-level perspective shows that local environmental conditions vary significantly and are shaped by human activities. Furthermore, people perceive and assess environmental change very differently, potentially shaping migration responses and/or leading to a variety of strategies other than migration (Brandt et al. 2014, Romankiewicz and Doevenspeck 2015). These contrasting images of the environment-migration association are due, in part, to scale.

The aim of this paper is therefore to introduce scale as an important issue to be considered in environment-migration research and to show the potential associated shortcomings inherent in the data and chosen analytical dimensions. We do this by examining the most common data and methods used in case studies concerned with the environment-migration nexus, and by providing examples of scale issues. We analyzed a selection of 27 English-language, peer-reviewed empirical case studies focused on the nexus between migration and the slow onset of climatic and environmental change found among the country reports and case studies listed in the bibliography “People on the move in a changing climate” (International Organization for Migration (IOM) 2012). The choice of articles was guided by our aim to cover a variety of methods as identified by Piguet (2010), Bilsborrow and Henry (2012), McLeman (2013) and Fussell et al. (2014). A number of articles published after 2012 were added to the review by searching for the key words “environment” and “migration” in academic databases.

This paper starts by introducing the scale concept and the related scale issues in geographical research. Then we review the dimensions of different data types used in studies of the environment-migration nexus and discuss related issues. After that, we review the scales at which the analyses are carried out and identify and discuss the main scale issues in environment-migration research.

2. Scale in geographical research

Gibson et al. (2000: 219) define scale as the “spatial, temporal, quantitative, or analytical dimensions used to measure and study any phenomenon”. The essential scale parameters are extent (boundaries) and resolution (detail). Scale levels share similar extents and resolutions and refer to locations along a scale, for example micro, meso and macro levels on the spatial scale, or
Data and methods in the environment-migration nexus: a scale perspective

Table 1  A conceptual framework of scale based on Keshkamat et al. (2012), showing the temporal and spatial dimensions of model and data scale and respective examples from environment-migration research

<table>
<thead>
<tr>
<th>Model Scale</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial dimension</td>
<td>Temporal dimension</td>
</tr>
<tr>
<td>Extent</td>
<td>Resolution</td>
</tr>
<tr>
<td>The area of interest covered in the analysis</td>
<td>The smallest spatial unit of analysis</td>
</tr>
</tbody>
</table>

Examples of modeling environment-migration linkage

Quantitative: Environmental degradation contributes to emigration and should be reflected in a measurable correlation between outmigration and vegetation change for the same district, during the same decade, in a given country.

Qualitative: A peasant expresses his concern about the late onset of the rainy season and decides to work temporarily in the capital, because his family’s income largely depends on harvest output. Another family member sends extra money from abroad.

<table>
<thead>
<tr>
<th>Data Scale</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial dimension</td>
<td>Temporal dimension</td>
</tr>
<tr>
<td>Extent</td>
<td>Resolution</td>
</tr>
<tr>
<td>The area covered or spanned in data collection</td>
<td>The smallest spatial unit of data collection</td>
</tr>
</tbody>
</table>

Example of quantitative data sources used to assess the link between migration and environmental change: here, outmigration rate and change of NDVI (Normalized Difference Vegetation Index)

- Census data of a country
- Individual residence change crossing an administrative boundary
- Census round every 10 years
- Residence change 5 years ago

- A satellite image scene (200 x 200 km)
- A pixel of 250 m²
- 14 years of recording
- One image every 16 days

Example of qualitative data to assess the link between migration and environmental change: here, migrants’ biographies and perceived environmental change as potential migration motives

- Transnational migrant network
- A village
- Migration history of a village/family
- Seasonal moves of individual family members

- Interpretation of environmental change at multiple places in the country of origin
- Interpretation of environmental change around the village of origin
- Age of an individual
- A season

long-term versus short-term on a temporal scale. The extent and resolution chosen for a case study are crucial because they affect the ability to identify and explain patterns of various phenomena (Gibson et al. 2000).

According to different research paradigms there exist varying perspectives and approaches to scale. Physical geographers mostly rely on mathematically measurable, deterministic representations of space and time in their investigations of natural phenomena, which they consider to be hierarchically nested at different levels along the spatial and temporal scale (Sheppard and McMaster 2008). This means that biophysical processes operate at distinct levels of time and space (operational scale). On the other hand, human geographers doubt the relevance of predefined geographic scales for (analyzing) social processes and emphasize that space, and therefore scale, are social, political or economic constructs (Sheppard and McMaster 2004). The distance between two places for example appears more significant in terms of social proximity and quality of social interactions rather than to be measured in kilometers or hours to travel. The individual’s identity, location and perception shape the relative importance of scale over space and time.
Keshkamat et al. (2012) illustrate how different perspectives of scale are manifest by describing a conceptual framework which is set up by three “axes”: reality scale, model scale and data scale (Table 1). These three scale types are divided into spatial and temporal dimensions, which are explained by extent (area or period covered) and resolution (level of detail). Several interconnected phenomena and processes in a spatiotemporal continuum, at different extents and resolutions, make up the reality scale. The model scale represents intangible realities in a practicable way depending on political or scientific objectives, which also determine the extent and resolution at which the analysis takes place. In doing so, models translate subjective realities and the complexity of interwoven processes at multiple scales into simplified causal relationships through the use of the data scale (see example in Table 1). The data’s extent and resolution are influenced by analytical needs, but are also determined by the nature of observable attributes and the methods used to measure them. Here, a major problem is the mismatch of the operational scales of ecological processes (e.g. catchment level) and social phenomena, such as migration networks stretching across regions and nations (Bruyninckx 2009).

Nevertheless, data are often processed to fit the chosen model scale and adjusted to the desired resolution (rescaling) to yield comparability across different data types. Down- and upscaling techniques include averaging, smoothing, extrapolating and interpolating, but can have severe consequences on the data accuracy, especially when data are missing (Atkinson and Tate 2000).

The Modifiable Area Unit Problem (MAUP) refers to how the units used in spatial analyses can take many different shapes and sizes, which in turn affects the outcomes of statistical analyses and thus the reliability of results (Dark and Bram 2007, Sheppard and McMaster 2004). The first sub-problem of the MAUP, the zonation effect, refers to how data are grouped into arbitrary (spatial or temporal) units that could easily be changed, yielding different analysis results (Openshaw and Taylor 1979). The second sub-problem, the scale effect, refers to the number and size of units into which the data are divided. Coarser units decrease the variance in the data and hence neglect that the socio-economic properties of individuals, or their spatial distribution, might not be homogeneous within area boundaries chosen for analysis (Dark and Bram 2007). This means that inferences about relationships between aggregated data do not necessarily hold true for non-aggregated data in the area of analysis, which is the geographical version of “ecological fallacy”.

The MAUP highlights that scale is a construction and not ontologically pre-given. Therefore the choice of analytical dimensions has political aspects (Delaney and Leitner 1997, Brown and Purcell 2005, Moore 2008, Rangan and Kull 2009). Brown and Purcell (2005) describe a “scalar trap” as a problem where researchers assume that e.g. the link between policies or actions, and the social and environmental effects of these actions, are stronger at local levels, while this may not necessarily be the case. The chosen analytical dimensions in environment-migration nexus studies may therefore also be determined by assumptions of the spatial and temporal nature of these processes.

3. Review of case studies

3.1 Data scale

3.1.1 Environmental data

Quantitative environmental data are included in many of the case studies that seek to determine the relevance of environmental factors in migration patterns. In these studies, environmental and climate conditions are principally represented by ground-measured climate data, global climatological datasets and/or satellite based vegetation data.

Station-measured climate data represent the meteorology of a point location, and therefore have limited spatial extent. The locations and density of stations are crucial for their ability to represent general meteorology, and all upscaling efforts are subject to the MAUP. Measured climate data can be limited in availability and quality (Bilsborrow and Henry 2012). In some countries, there is a long record of climate data and a high density of stations, like in the U.S., Burkina Faso and Mexico (Mitchell and Jones 2005, Méndez González et al. 2008, Lodoun et al. 2013). For case studies using station data from these countries see Henry et al. (2003), McLeman and Smit (2006), Feng et al. (2010) and Nawrotzki et al. (2013). In other countries, stations are sparse and data quality is not sufficient for spatial or temporal analyses.

Here, global gridded climatological datasets can serve as alternatives or complements. Many case studies use data originating from the Climate Research Unit (CRU) at the University of East Anglia (among them the stud-
ies by Henry et al. 2004a, Henry et al. 2004b, Barrios et al. 2006, van der Geest et al. 2010, van der Geest 2011, Marchiori et al. 2012). These data are available as grids with different temporal extents depending on version (Table 2). The grids are based on the upscaling of climate point data from different sources stored in a ground stations database and the interpolation of missing values based on anomalies (Mitchell and Jones 2005).

One of the main problems with constructing a global reference database of climate data is the spatially and temporally varying quality and availability of reference data. As an example, the number of precipitation stations included in the global database increased from approximately 3,000 to nearly 10,000 between 1901 and 1970, but decreased to less than 2,000 in the year 2000 (Mitchell and Jones 2005). Figure 1 illustrates the uneven distribution of the stations and the fact that their number has decreased further. In the CRU version 3.21, less than 100 stations report rainfall data for Sub-Saharan Africa in January 2008. Data gaps are interpolated in the preparation of the dataset, but might not reflect the actual climatological situation in the region.

The spatial dimensions of the CRU dataset is another limitation that affects its usability in analyses at finer resolutions. A cell size of 0.5° (approximately 50 km) neglects spatial variations and is much coarser than the satellite-based vegetation data available. An alternative to CRU is the Worldclim dataset (used by Gray 2009) that has a spatial resolution of 1 km and has also been interpolated based on a global database of climate stations (Hijmans et al. 2005). However, the data sources are similar to CRU and thus a finer spatial resolution does not necessarily mean an improved quality.

The issue of heterogeneous station data is addressed in the Global Precipitation and Climatology Project (GPCP) dataset, available globally since 1979 and used in Gray and Mueller (2012a), and Gray and Mueller (2012b). Here, rain gauges are blended with satellite-based precipitation estimates (Adler et al. 2003). However, despite its high reliability, the extremely coarse resolution of 2.5° (approximately 250 km) makes GPCP data inappropriate for detailed analyses.

The Inter-Governmental Panel on Climate Change (IPCC) dataset, used by Barrios et al. (2006) and Marchiori et al. (2012), is a rescaled version of the CRU dataset, where country level data have been calculated from grid cells, using weighted averages (Mitchell et al. 2002). Mitchell et al. (2002) stress that these data are not intended to be used at subnational level, and

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Variable</th>
<th>Temporal extent (version)</th>
<th>Spatial resolution</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Research Unit Time Series (CRU TS) Version 1.0 – 3.2</td>
<td>Precipitation, temperature</td>
<td>1901-1995 (1.0) 1901-2000 (2.0) 1901-2002 (2.1) 1901-2012 (3.21)</td>
<td>0.5°</td>
<td>New et al. 2000 Mitchell et al. 2004 Mitchell and Jones 2005 Harris et al. 2014</td>
</tr>
<tr>
<td>Worldclim</td>
<td>Precipitation</td>
<td>1950-2000</td>
<td>1 km</td>
<td>Hijmans et al. 2005</td>
</tr>
<tr>
<td>Global Precipitation and Climatology Project (GPCP)</td>
<td>Precipitation</td>
<td>1979-present</td>
<td>2.5°</td>
<td>Adler et al. 2003</td>
</tr>
<tr>
<td>Global Inventory Modeling and Mapping Studies (GIMMS)</td>
<td>NDVI</td>
<td>1981-2006</td>
<td>8 km</td>
<td>Tucker et al. 2004</td>
</tr>
<tr>
<td>Moderate Resolution Imaging Spectroradiometer (MODIS)</td>
<td>NDVI</td>
<td>2000-present</td>
<td>250 m</td>
<td>Solano et al. 2010</td>
</tr>
</tbody>
</table>

Table 2 Global continuous environmental data used in the reviewed studies and their spatial and temporal characteristics
cannot represent conditions at a certain point. While aggregating data over political areas (as opposed to grids or climate zones) may seem desirable for global or regional comparisons, the data are generalized and thereby so limited that the added value of using the IPCC dataset rather than a grid-based dataset appears questionable. Needless to say climate is not spatially homogeneous within a country, and neither are climate anomalies (see the example of Mali in Figure 1).

Vegetation data from satellites are a common alternative to directly measured or interpolated climate data, as recommended by Bilsborrow and Henry (2012). With data over time, these sources can reflect vegetation response to climate and other factors. A decreasing vegetation cover might thus be interpreted as a sign of land degradation and a reduction of ecosystem services (Henry et al. 2004a, Van der Geest et al. 2010). Continuous satellite data provide global images and are available from e.g. the Moderate Resolution Imaging Spectroradiometer (MODIS) (used in Leyk et al. 2012), Pathfinder or Global Inventory Modeling and Mapping Studies (GIMMS) datasets (used in Henry et al. 2004a, van der Geest et al. 2010 and van der Geest 2011). These datasets include Normalized Difference Vegetation Index (NDVI) data that represent the greenness of the earth surface.

For all satellite data there is a tradeoff between temporal and spatial resolutions. Large extent and coarse resolution satellites have a high frequency of data collection, while the fine resolution satellites cover smaller areas, and less frequently. Pathfinder and GIMMS have a resolution of 8 km and cover the periods 1982-1995 (Pathfinder) and 1981-2006 (GIMMS) (James and Kalluri 1994, Tucker et al. 2004). For each grid cell, a 1.1 km sample of AVHRR (Advanced Very High Resolution Radiometer) satel-

---

**Fig. 1** Massive decline of stations in Africa in CRU data (v. 3.2) and strong regional differences in average annual rainfall in Mali, based on Tropical Rainfall Measuring Mission (TRMM, 0.25° spatial resolution)
Data and methods in the environment-migration nexus: a scale perspective

In this context, the analytical power of remote sensing methods is bounded and limited. More specifically, the temporal consistency of satellite data is largely biased by orbital drifts, having severe impacts on trend analysis (Tian et al. 2015).

Still, coarse vegetation data offer a unique assessment on long-term trends: human activities and environmental processes, however, are rarely uniform, shaping a heterogeneous landscape. Hence, the visibility of changes in vegetation cover is scale-dependent and often obscured or neutralized by adjacent areas and merged into single mixed pixels (Brandt et al. 2014).

Despite these drawbacks in spatial (GIMMS) and temporal (MODIS) aspects, satellite data are considered the most accurate assessment of land degradation (Bai et al. 2008, Nachtergaele et al. 2011). An older approach by Oldeman et al. (1990) produced a global map based on regional expert judgments, the GLASOD (Global Assessment of human-induced Soil Degradation) map. Although this rather qualitative approach has been severely criticized (Sonneveld and Dent 2009, Nachtergaele et al. 2011) and found unreliable even at a national scale, it has been used as data input in environment-migration models (Henry et al. 2003).

### 3.1.2 Migration and socio-economic data

Relying on most definitions, the phenomenon of migration as a particular form of population movement is a human construct and itself subject to scale issues. Both the crossing of an administrative boundary and the shift of the usual residence for a given period of time (Standing 1984) are politically constructed dimensions of temporal and spatial scale and define which movements become relevant for investigation. To reliably capture the multitude and diversity of migratory movements is an unfeasible attempt. The scope of migration and socio-economic information considered in case studies is therefore restricted by the respective method(s) and focus of data collection. Quantitative data about socio-economic characteristics and migration patterns commonly used in regression analyses are generated by censuses, population registers or (large n) sample surveys (Fussell et al. 2014). Qualitative data are collected through ethnographic research methods and intensive fieldwork or come from historical records and archives.


However, censuses do not capture detailed information on population mobility, and few of these studies reflect upon the methods and quality of the data used. This is particularly important because many of the reviewed empirical investigations are set in developing countries where the reliability of census information is uncertain (Pedersen 1995).

Examining how census data are produced reveals several scale-related limitations. Generally, migration information is obtained by looking at an individual’s place of birth and/or place of usual residence at a specific date in the past (e.g. one or five years ago), and the current place of usual residence. Moreover, census surveys usually take place every ten years (Chudinovskikh et al. 2008). Thereby, census data cannot specify the exact timing of a migration and do not allow the identification of movements and residence changes that took place “in-between”. This particularly applies to birthplace-related migration data (van der Geest et al. 2010), i.e. solely comparing place of birth and current residence and reducing all potential migration steps of a lifetime to only one of unknown timing. It is therefore not possible to identify specific forms of movements such as seasonal migration (Henry et al. 2003) or return and circular migration. Fussell et al. (2014) stress that the time lags between migration information from censuses and environmental events pose the greatest difficulties to comprehensibly associate environmental parameters and migration from a specific area, let alone derive causal linkages for households or individuals.

Another scale-related problem is the fact that census reports generally aggregate the given information to larger administrative levels. Van der Geest et al. (2010), for example, had to downscale regional census migration data to enable an analysis with district level NDVI data. Moreover, a longitudinal analysis is difficult when, due to a previous restructuring of
Data and methods in the environment-migration nexus: a scale perspective

administrative boundaries or bad census quality, long-term data are not comparable over time (cf. Pedersen 1995, van der Geest et al. 2010). Administrative units and, above all, national borders, i.e. politically constructed spatial scales, set the respective frame for data collection and analysis of national census information. Even though the majority of migration flows typically takes place within countries, non-negligible numbers of migrants cross international borders and are missed by census surveys and respective analyses. While censuses are able to grasp the immigrant population of a country, to reliably register emigration and migration of a country, to reliably register emigration and internationals residing abroad is hardly possible and international migration cannot be derived from census data alone (Chudinovskikh et al. 2008, United Nations Department of Economic and Social Affairs 2008).


Scale issues of sample surveys already take effect during data collection. Cross-sectional surveys are typically carried out once and gather information about past migration retrospectively. A panel survey is a costly and more reliable approach to conduct repeated interviews over a period of time to disclose individual courses of migration, as Massey et al. (2010) did during 36 months in Nepal. Asking individuals retrospective and time-specific questions about the temporal order of environmental and migration events, and/or migration motives (Doevenspeck 2011, Findley 1994) may give indications of causality in migration decisions at the individual or household level (Fussell et al. 2014). However, the comparability and representativeness of survey results are limited, considering the varying size of samples and the fact that not only the time span defining a migration event but also the types of movements applied in data collection and analysis vary. Massey et al. (2010), for example, defined an absence of at least one month as migration, for others it is three months (Henry et al. 2004b) or six months (Gray 2009). In this regard, survey results demand a careful case-specific interpretation and cannot be simply aggregated and compared at the regional or national level (Meze-Hausken 2000). The empirical work of Massey et al. (2010) demonstrates that analyses of data from a larger, longitudinal sample yielded, in part, contradictory results to a study conducted earlier in the same area using the same questionnaire and survey method but using cross-sectional data (Shrestha and Bhandari 2007).

Contrary to ecological fallacy, missing context information in analyzing individual migration data may result in atomistic fallacy, i.e. mistakenly inferring that a correlation between two variables at the individual level holds true at aggregated levels too. Instead, recording socio-economic and migration data together with survey questions on perceptions of (changing) environmental conditions (e.g. rainfall, crop yields or the availability of natural resources) for the same household may contribute to more harmonized data (Bilsborrow 2009, Massey et al. 2010). Moreover, such perceptions refer to relevant scales of environmental change that have material impacts on households (Manson 2008). As such, perception data may be more relevant when drawing conclusions about causal linkages to migration decisions.

Qualitative methods, in particular participatory observations, in-depth and biographical interviews with migrants and migrant households, can reduce shortcomings of quantitative methods. Such a context-sensitive approach provides better insights into social constructions of temporal and spatial scales of alleged “facts”, such as environmental degradation and droughts as well as migration decision-making (Meze-Hausken 2000, Mcleman and Smit 2006, Carr 2005, Gilbert and Mcleman 2010). Scale aspects are identifiable at the point of data collection since data are often collected among fewer respondents and in fewer places. These detailed data, often collected over longer periods of time, represent a prerequisite for yielding qualitative knowledge of complex social processes and the interplay between environment and migration. Carr (2005), for example, conducted qualitative interviews with 30 individuals during 13 months of ethnographic fieldwork over three years. Other scholars relied on historical records from national archives (Pedersen 1995, Mcleman and Smit 2006). Mcleman and Smit (2006), for example, used reports and administrative records of migrant camps, oral histories and autobiographies from the 1930s and 1940s to investigate
Data and methods in the environment-migration nexus: a scale perspective

3.2 Model scale

Scale issues related to data and to methodological approaches are closely connected. In this section, we discuss examples of scale issues as related to analytical approaches.

3.2.1 Rescaling

Data on migration, environment and socio-economic factors are of different spatial and temporal resolutions. Therefore, some data are often rescaled to a certain model scale, typically human-created administrative boundaries, to facilitate the analysis. In Henry et al. (2003) and Henry et al. (2004b) for example, CRU data were aggregated from grid cells of 0.5° to department boundaries (Bilsborrow and Henry 2012). Van der Geest et al. (2010) up-scaled GIMMS NDVI data to district level by a weighted-average approach and downscaled migration data from regional to district level using a regression model. Weighting the data according to the share of pixels within district boundaries allows for a more realistic aggregation. Instead of using administrative boundaries as model resolution, Leyk et al. (2012) focused on household level data and aggregated NDVI data into two-kilometer buffer zones surrounding the household residence, based on the distance commonly traveled for the collection of natural resources. This approach takes into account the location of settlements and its proximate vegetative resources instead of aggregating into administrative units.

3.2.2 Modifiable area unit problem and ecological fallacy

The chosen model resolution is often an administrative unit that does not necessarily reflect the spatial distribution, or socioeconomic characteristics, of the population, nor local environmental features. Alternative zonation approaches could be based on catchment areas, buffer zones around settlement areas or grids. The size of the units in which the data are aggregated has a major influence on the results of the analysis, as the variance decreases with increased unit size. Piguet exemplifies this issue of ecological fallacy when he states that “[N]othing guarantees that the very people who emigrated and contributed to a negative migration balance in an area under environmental stress, for example, are the same individuals who experienced that environmental stress and took a decision to migrate accordingly” (Piguet 2010: 518). Therefore, the choice of analytical scale should be based upon an understanding of the scale at which migration and the environmental parameters of interest operate, rather than the scale of data available.

We identified several coarse-scale case studies, of which three engaged data at the national level (Barrios et al. 2006, Reuveny and Moore 2009, Marchiori et al. 2012). Barrios et al. (2006) compared Sub-Saharan Africa (SSA) with non-SSA countries with regard to correlations between precipitation and urbanization rates. The authors acknowledged that there might be variations in rainfall within countries, but refer to lack of data to explain the neglect of sub-national variation. The different spatial extents between SSA and non-SSA group could explain the varying strengths of the rainfall-urbanization relationship, which are expected to increase with increasing extent. Using the same spatial dimensions and geographical focus, Marchiori et al. (2012) used regression models to assess the number of people, who had been “displaced” by weather anomalies between 1960 and 2000, and to estimate future displacement due to climate change, based on population and climate projections. Such estimates, based on coarse and limited data, resemble the (in)famous projections made by Myers (2002), which have been widely cited and later harshly criticized (Brown 2008).

In a study of environment and migration in South Africa, Maclaurin et al. (2015) examined the implications of scale effects, focusing on aggregation. They use NDVI data from the MODIS instrument to represent greenness/natural capital, and migration and socio-economic data from a rural household survey. Migration and socio-economic data, originally collected at household level, were aggregated and summarized into nine levels, where level 1 included 2-4 households and level 9 included 10-20 households. Using both global and local regression models to investigate the relationship
to migration data, the findings revealed that whereas the greenness measure was a rather stable predictor of out-migration at all aggregation levels, the socio-economic variables’ ability to explain out-migration was reduced with increasing levels of aggregation. Mac-laurin et al. (2015) conclude that empirical research on the migration-environment nexus is best studied at the household level where decisions are made, and that higher levels of aggregation may miss important associations. It should be noted that the coarsest aggregation level of 10-20 households in this study is still far from the coarseness of studies that conduct analyses at the district, province, or even country level.

3.2.3 Temporal scale

The temporal dimensions of an analysis are just as important as the spatial dimensions for the ability to identify and explain patterns. A temporal dimension is necessary to denote environmental change or stress, but some studies focus solely on the spatial differences in environmental characteristics (e.g. Amacher et al. 1998).

Studies using a time series or event history approach (e.g. Gray and Mueller 2012a) assume that patterns of migration should partly follow the evolution of e.g. rainfall patterns or vegetation density over the period of time under review (Piguet 2010). If a migration event is temporally associated with an environmental change, it is important to study the different parameters at an appropriate temporal resolution, for example to make sure that the strong increase in migration during a certain year of drought actually happened after the onset of the drought. Furthermore, many natural disasters happen over long time periods, such as land degradation and drought, which calls for a wide temporal extent of analyses. Several of the studies have a rather narrow temporal extent and focus only on a single period of environmental stress (e.g. Findley 1994, Ezra and Kiros 2001, Henry et al. 2003, Leyk et al. 2012). This highlights the challenges of designing spatiotemporal analyses based on data with limiting spatial and temporal dimensions.

3.2.4 Multilevel or multi-scale approaches

Some case studies have adopted a multilevel approach, where the data on migration and environment are analyzed at different resolutions or extents (Ezra and Kiros 2001, Henry et al. 2004b, Gray and Mueller 2012a, 2012b, Leyk et al. 2012). This method gives more precise results by disentangling the complexity of scale effects and has the potential to provide insights about the reliability of the results, by showing how statistical relationships vary across analysis levels. At the same time the approach requires, besides detailed environmental data, a costly and extensive collection of data through individual (longitudinal) and community surveys. In Henry et al. (2004b)’s analysis, multilevel models use socio-economic data collected at two or more levels, although the environmental data were collected at a single administrative level. The statistical analysis was then conducted at the individual level, with community variables assigned from community level data, and rainfall variables assigned from department level data. Thus, multi-level analysis did in this case not mean multiple analytical levels, but rather multi-level data downscaled to one analysis level. Leyk et al. (2012) instead attempted to overcome some scale issues associated with regression models by conducting the analysis at three “nested spatial scales” (extents), while retaining the analysis resolution. They found that the strength of the relationship varied with the analysis extent, and a stronger relationship between migration and natural resources was found when focusing on a smaller area. The results show the effects of variation within samples on statistical relationships between variables, as explained by Wiens (1989).

4. Conclusion

This review of scale in the environment-migration nexus highlights the need for understanding the complexity of the processes that shape migration patterns in order to translate them into a comprehensive model. Given that socio-economic, migration and environmental data are very different in measurability and quality, scale mismatches and inadequately detailed data complicate the analysis and inevitably demand the use of various re-scaling techniques to facilitate comparability of the data. The paper furthermore shows that many studies include spatial analyses, but the temporal dimension of the environment-migration nexus is more often neglected. The lack of discussion of scale in the majority of the literature indicates that the choice of model scale is based on preconceived ideas of what scales the environment-migration nexus operates at and on the spatial and temporal dimensions of the available data.
This is a call for awareness of scale implications specific to both quantitative and qualitative research approaches in data collection and analysis for establishing environment-migration linkages. We recommend combining multiple data sources and qualitative and quantitative approaches to reduce the effects of data errors and generalizations, and to allow for multiple perspectives. We emphasize that field experience and contextual knowledge is essential to interpret statistics in a nuanced way. Considering scale aspects in the investigation of environment-migration linkages throughout the research process may help to reflect on the pertinence of research questions and the reliability of research results. It can also motivate scholars and policy-makers to facilitate the production, collection and analysis of relevant qualitative and quantitative data across multiple scales.

The theoretical and terminological complexity of the scale concept admittedly offers further promising fields of studying the importance of scale as an epistemological and ontological category for environment-migration research beside the emphasis on observational scale defined by extent and resolution. On the one hand, it is unavoidable for environment-migration research, as for any other scientific endeavour, to define a model scale as an epistemological reference in order to identify and understand certain structures. On the other hand, this definition is always a conscious act with material and political effects. Being aware of this double structure of related model and reality scale involves a next step of exploring the importance of scale for environment-migration research, putting emphasis on the social and political importance of scale. Questions of how political agendas and targets of science policy effect scales of analysis and how the making of scales of analysis, operation and intervention are mutually linked are only two aspects of the research gap that must be approached in this field of environment-migration linkages.

Acknowledgements

This research was made possible through a Short Term Scientific Mission funded by COST Action IS1101: Climate change and migration: knowledge, law and policy, and theory; the Middle East in the Contemporary World (MECW) project funded by the Swedish Research Council; and the micle project, funded by the BMBF (German Ministry of Education and Research). The discussions in this manuscript were stimulated by discussions at two workshops on scale, one organized by the CLIMBECO research school in Lund and one organized by COST Action IS1101 in Bonn. We thank Dr. Jonathan Seaquist for comments on an early version of this manuscript, and Professor Petter Pilesjö for comments on a later version.

References


Data and methods in the environment-migration nexus: a scale perspective


Brown, J.C. and M. Purcell 2005: There’s nothing inherent about scale; political ecology, the local trap, and the politics of development in the Brazilian Amazon. – Geoforum 36 (5): 607-624

Brown, O. 2008: The numbers game. – Forced Migration Review 31: 8-9

Bruynincks, H. 2009: Environmental evaluation practices and the issue of scale. – New Directions for Evaluation 122: 31-39


Doevenspeck, M. 2011: The thin line between choice and flight: environment and migration in rural Benin. – International Migration 49, Supplement 1: e50-e68


Gray, C.L. and V. Mueller 2012a: Natural disasters and population mobility in Bangladesh. – PNAS. Proceedings of the National Academy of Sciences of the United States of America 109 (16): 6000-6005


Laczko, F. and C. Agahazar (eds.) 2009: Migration, environment and climate change: assessing the evidence. – Geneva
Data and methods in the environment-migration nexus: a scale perspective


McLeman, R. 2013: Developments in modelling of climate change-related migration. – Climatic Change 117 (3): 599-611


Moore, A. 2008: Rethinking scale as a geographical category: from analysis to practice. – Progress in Human Geography 32 (2): 203-225

Myers, N. 2002: Environmental refugees: a growing phenomenon of the 21st century. – Philosophical Transactions of the Royal Society B: Biological Sciences 357 (1420): 609-613


Romankiewicz, C. and M. Doevenspeck 2015: Climate and mobility in the West African Sahel: conceptualising the local dimensions of the environment and migration nexus. – In: Greschke, H. and J. Tischler (eds.): Grounding global climate change. – Dordrecht et al.: 79-100

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf and M. Ziese 2011: GPCC full data reanalysis version 6.0: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. – Global Precipitation Climatology Centre (GPCC), c/o Deutscher Wetterdienst. – Offenbach, Germany. – Accessed online
Data and methods in the environment-migration nexus: a scale perspective

Tucker, C.J., J.E. Pinzon and M.E. Brown 2004: Global inventory modeling and mapping studies. – Global Land Cover Facility, University of Maryland. – College Park, Maryland United Nations, Department of Economic and Social Affairs, Statistics Division 2008: Principles and recommendations for population and housing censuses. – New York
Warner, K., M. Hamza, A. Oliver-Smith, F. Renaud and A. Julca 2010: Climate change, environmental degradation and migration. – Natural Hazards 55 (3): 689-715