

12.4 Description of data analysis methods

Data analysis is, according to Bryman (2012: 13), “fundamentally about *data reduction* – that is, it is concerned with reducing the large corpus of information that the researcher has gathered so that he or she can make sense of it”. It translates raw data into empirical findings (Patton 2002: 432) and is therefore a vital prerequisite to interpret the empirical data and draw conclusions from it. Data analysis involves two major steps: Firstly, data preparation and secondly, the analysis of the data. The following section will describe how qualitative, quantitative, and spatial data sets were prepared and analysed.

12.4.1 Preparation and analysis of qualitative data

The qualitative data of this research project stem from exploratory in-depth and semi-structured interviews as well as participatory focus group discussions with authorities and households. The analysis was guided by profound methodological literature (Flick 2007; Bernard 2006; Taylor-Powell, Renner 2003; Patton 2002; Mayring 2000; Miles, Huberman 1994) and expert advice.

In a first step, the field notes taken throughout the interviews and discussions were transcribed by the researcher and by the survey assistants. There were no voice recordings taken so that no word-by-word transcriptions were undertaken. A second step was the data reduction, i.e. “the process of selecting, focusing, simplifying, abstracting and transforming the data that appear in written-up field notes” (Miles, Huberman 1994: 10). This process of data condensation followed the guiding principles of a directed qualitative content analysis (Hsieh, Shannon 2005). It aimed to summarise data material without losing substantive content, explaining ambiguous text sequences with complementary material and structuring the data purposive (Mayring 2000: 115). Therefore, field notes were processed by a systematic classification in MaxQDA, a software program for computer-assisted content analysis. Three different code systems were developed for the analysis: One code system for all interviews which was key for developing the conceptual components and linkages of the theoretical framework; a second code system which was used to structure migration-centred interviews according to general migration-based indicators, and a third system which classified strategy-centred authority interviews to analyse the individual evaluation of governmental strategies.

The process of coding and identifying patterns was directed by the research questions and by conceptual and theoretical considerations. Three guiding conceptual components, i.e. social-ecological system prerequisites, individual risk perception and strategy characterisation and evaluation were used as initial super-ordinate codes for the first code system; subcodes were then deductively associated with these codes. Throughout the data analysis, new codes were added and previous codes were transformed or merged where necessary. These analytical choices were based on an inductive process in the vein of ‘Grounded Theory’ (see Glaser, Strauss 1967). Moreover, each interview document was characterised according to quantitative variables or attributes, e.g. production type, actor group or major hazard exposure. In that way, an integration of qualitative and quantitative data was possible. It allowed, for instance, quantifying how many households migrated to Ho Chi Minh City and how these migrants evaluated a certain strategy compared with the others.

12.4.2 Preparation and analysis of quantitative data

In the present research context, there are two main sources of quantitative data, the household survey and production-cost-benefit interviews. The analysis process followed guidelines provided by expert researchers and relevant secondary literature (see e.g. Bryman 2012; Raab-Steiner, Benesch 2008; Atteslander 2006; Trochim 2006; UN 2005).

The analysis of the quantitative survey data required, firstly, an appropriate format for a reliable data entry. Data logging of questionnaires was undertaken by the two leaders of the enumerator groups. They did not only know the survey design well but also cross-checked each of the questionnaires conducted during the survey. They were, however, only acquainted with Microsoft Office and not SPSS. For this reason, the data were first entered in an Excel template and then imported into SPSS. In the data entry program, questions and answer options were rationally and simply coded, missing values and the level of measurement were defined, and automatic error reports were programmed so that entry mistakes could be reduced. The data entry assistants received one day of training and conducted first data entries together with the researcher. Throughout the data entry period, spot checks on a random basis and feedback rounds with data entry assistants were conducted. After having conducted data quality checks, some of the variables were transformed, i.e. categorised, recoded or calculated into new variables or answer sets. Having checked and cleaned the data set, statistical analyses guided by the overall research questions were undertaken. Descriptive analyses gave a basic characterisation and summary about sample and measures; inferential statistical methods, particularly correlation analyses, were applied to make inferences from the present sample to the whole population.

For the analysis of the quantitative production-cost-benefit data an appropriate data entry design also had to be delineated. The records from semi-structured interviews were logged into an Excel template by the enumerators who conducted the interview. They received a half-day training and were able to express problems and unclarities in feedback rounds afterwards. The data entry form was ordered along the process of production and comprised variables describing the future, present and past costs for each step as well as the benefits generated from production. Having translated the Excel spreadsheet of each interview from Vietnamese into English, the data were checked for consistency and the units of the numerical data were harmonised. Subsequently, descriptive analyses were undertaken to compare different production types, socio-economic groups and time periods.

12.4.3 Preparation and analysis of spatial and visual data

Spatial data, i.e. data which describes the “spatial relationship or spatial interactions between cases” (Haining 2003: 4) were either drawn from secondary sources in form of GIS-files and maps or were produced in interviews with hamlet leaders in form of resource risk maps. Other visual data were created in form of Venn diagrams in participatory household focus group discussions. The analysis was guided by relevant literature (Ormsby; Haining 2003) and supported by colleagues.

Firstly, all maps were translated from Vietnamese to English and Venn diagrams were digitalised. The geo-referenced vector data from the WISDOM server provided a foundation to create a base map of the research communes and the water bodies in Tra Cu district and in the Mekong Delta. This base map provided a foundation for creating several thematic maps based on statistical secondary data. There were no geo-referenced data available which depict the hamlet borders, however. Therefore, the resource risk maps and communal reports were used as templates to create new shape files delineating hamlet borders. This base map provided the basis for creating thematic maps from socio-economic statistical data at the hamlet level, for designing an elevation map based on raster data, and for making a saline intrusion map based on isohaline shape files. The depiction of dominant production types was based on shape files which were created in accordance to a land use map of Tra Cu District and were simplified and abstracted with knowledge received from authority interviews and field observations.

Publication bibliography

- ATTESLANDER, P. (2006): Methoden der empirischen Sozialforschung. Berlin. Erich Schmidt Verlag.
- BERNARD, H. R. (2006): Research methods in anthropology. Oxford, UK. Rowman & Littlefield Publishers.
- BRYMAN, A. (2012): Social research methods. 4th ed. Oxford, New York. Oxford University Press.
- FLICK, U. (2007): Qualitative Sozialforschung - Eine Einführung. Hamburg. Rowohlt Taschenbuch Verlag.
- GLASER, B. G.; STRAUSS, A. (1967): The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago. Aldine.
- HAINING, R. P. (2003): Spatial data analysis: theory and practice. Cambridge, UK. Cambridge University Press.
- HSIEH, H.-F.; SHANNON, S. E. (2005): Three approaches to qualitative content analysis. *Qualitative health research* 15 (9): 1277-1288.
- MAYRING, P. (2000): Qualitative Inhaltsanalyse. *Forum Qualitative Sozialforschung* 1 (2): Art. 20.
- MILES, M. B.; HUBERMAN, A. M. (1994): Qualitative data analysis: An expanded sourcebook. London, Thousand Oaks. Sage Publications.
- ORMSBY, T.: Getting to know ArcGIS desktop: basics of ArcView, ArcEditor and ArcInfo. ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (ESRI). Redlands, USA.
- PATTON, M. Q. (2002): Qualitative evaluation and research methods. Thousand Oaks, London, New Delhi. Sage Publications.
- RAAB-STEINER, E.; BENESCH, M. (2008): Der Fragebogen. Von der Forschungsidee zur SPSS-Auswertung. 3rd ed. Wien, Stuttgart. UTB.
- TAYLOR-POWELL, E.; RENNER, M. (2003): Analyzing Qualitative Data. UNIVERSITY OF WISCONSIN-EXTENSION COOPERATIVE EXTENSION. Madison, USA.
- TROCHIM, W. M. (2006): Research Methods Knowledge Base. WEB CENTER FOR SOCIAL RESEARCH METHODS. Available online at <http://www.socialresearchmethods.net>, checked on 29/07/2013.
- UN (2005): Household Sample Surveys in Developing and Transition Countries. New York *Studies in Methods*, Series F No. 96.