4.3 Data privacy and protection

Privacy protection is one of the most critical aspects when CDR data are used for producing statistics for policy purposes (Jansen et al., 2021). Unit-level CDR data de-identified before the data are processed where no-individually identifiable information is included. The data are processed in the premise of the MNO or regulator, and are kept strictly confidential. Statistical products are shared for informing decision only in aggregated forms. Displacement and disaster statistics introduced in this Handbook are based on CDR aggregates. Aggregating CDR data is part of the pre-process, which enables the data producer to maintain privacy (Buckee & Enge-Monsen, 2016; Flowminder mobility data products, 2020). It is important to aggregate the data at certain geographic levels and ensure that the aggregates include enough population in groups (Sweeney, 2000), moreover, combining multiple levels of aggregation in space and time can also provide further information (Flowminder mobility data products, 2020). In addition to aggregation, dropping small counts also enables privacy protection. In general, grid cells including up to 20 people should not be shared (Maas et al., 2019).