

Managing quality and consistency of geocodes in the Buildings and Dwellings Register in Statistics Austria

As the Buildings & Dwellings Register (BDR) and its address and building geocodes respectively form the basis for further point-based statistics, the quality of the register is of outmost importance. The following aspects are particularly crucial and need special attendance when linking other registers and data:

- Address-ID and building-ID should be unique and unambiguous.
- Address-ID and building-ID should be “active” for the required time-stamp
- Address-ID and building-ID should correspond; it should be clear which parcel address a building address lies on.
- Address-ID and building-ID should have valid geocodes

To safeguard consistent linkage between geospatial data and registers usually the building-ID is the common field. In the case where the address information is only available on parcel level, the parcel address and its geocode is used. So as long as data is provided with the correct and valid building-ID or at least address-ID the linkage with the geocodes should not be a problem.

Due to possible changes in administrative boundaries between the timestamp of the data to be linked and the timestamp of the Buildings and Dwellings register the dependency between building-IDs and municipality codes in the data has to be checked and possibly updated.

Experience also shows that sometimes records cannot be linked due to mistakes in the Buildings and Dwellings register or due to inactive or out-dated building-IDs in some registers. There is no definite procedure to overcome this, but miss-matches between statistical registers/records and geospatial address data/building data are checked both by the staff responsible for the statistics as well as the geoinformation staff. If it turns out to be a mistake in the Buildings and Dwellings Register it is reported to the Buildings and Dwellings Register hotline, which then contact the municipality and if possible, mistakes are corrected right away. If the mistake is found to be due to inactive or out-dated building-IDs in the data to be linked currently the solution is to accept inactive or out-dated building-IDs and use their geocodes or that of neighbouring addresses/buildings for further analysis.

If records are missing a correct building-ID they cannot be linked with the Geodatabase and hence have no geocode to be used for point-based statistics. This is the reason why statistical aggregates from points (grids and point-based analysis) may not amount to 100 percent of the population. Having been aware of the problems this causes to the world of spatial analysis the departments now try to reduce the number of missing building-IDs by looking for the addresses also in the building tables for other dates of validity.

By using the geocodes for spatial analysis or the aggregation of grids and other derived products, internally or by customers, some mistakes in the geocodes have been discovered in the past. They were mainly due to historic entries into the system, which were still done by hand (miss typing), so before the Geoclient was installed. These were reported to the Buildings and Dwellings Register hotline, which then contacted the municipality and the mistakes were corrected.

Examples:

- Geocodes from the Buildings and Dwellings Register are also used for traffic navigation by emergency suppliers (ambulance, fire brigade...), so it is a living register. Errors can be fatal and get reported, so the quality gets constantly improved.
- When calculating the commuting distance based on the geocodes from the Buildings and Dwellings Register the result for the commuting distance for all pupils of a certain primary school in Vienna was more than 15km. This was clearly implausible and it was soon clear that the geocode of that school was wrong! The geocode was corrected in the Buildings and Dwellings Register and the routes recalculated.
- Built-up grid cells in the middle of a lake (no, there is no island there) made incorrect building geocodes visible. The problem was that some buildings at the shore of the lake only had a

geocode for the parcel address and this was placed in the centre of the parcel (piece of "land"), which reached far into the lake. Again these building geocodes were corrected.