# The Potsdam Housing Market: A GIS-based Spatial Analysis using FOS

*Harald Schernthanner, Hartmut Asche*

(Harald Schernthanner, University of Potsdam, Department of Geography, Karl-Liebknecht-Straße 24/25,14476 Potsdam, Germany, harald.schernthanner@uni-potsdam.de)
(Hartmut Asche, University of Potsdam, Department of Geography, Karl-Liebknecht-Straße 24/25,14476 Potsdam, Germany, gislab@uni-potsdam.de)

## 1 ABSTRACT

Housing in Potsdam varies from flats in redeveloped prefabricated high-rise buildings to apartments in historical townhouses to condominiums in Germany's first gated community. Increasing demographic development and a stagnant public housing sector generate potential for spatial conflicts. For the time being in-depth GIS-based spatial analysis of the housing market lacks. This article analyses spatial trends and distribution patterns of the Potsdam housing market, using geostatistical methods implemented in free open-source geographic information systems (FOS GIS). To assemble a spatially differentiated picture of the housing market, methods such as spatial interpolation techniques and spatial declustering are applied. The analysis presented here is based on a representative sample of recent housing market data from 2009. The study provides a basis for discussion of a generic approach to housing market analysis based on free open-source geoinformation systems.

## 2 INTRODUCTION

Today the benefits of GIS for the real estate sector are undisputed. Application of GIS in this sector is, however, far from being state-of-the-art. Until a few years ago not even major players in the real estate sector applied GIS analysis in their business (Borchert 2006). Frequently the application of GIS in real estate portals[1], e.g., is still limited to illustrate real estates traded. Currently, a trend can be observed to link real estate portals with real estate online mapping services (as provided in google maps, bing maps etc.) to so called geo mashups.This development can be considered a clear indication ofthe rising acceptance of GIS in the real estate sector. At the start of this new decade one finds that GIS in real estate is applied in research as well as in the evaluation of sales and for marketing purposes (Borchert 2006). However, the use of GIS in the real estate sector, especially in real estate management, often does not reach beyond the management and visualisation of real property. The existing potential of geographic information systems for GIS-based real-estate analysis, whether proprietary or FOS based, is thus not made full use of.

This paper presents a FOS GIS based housing analysis of the city of Potsdam. It proposes a process chain of a FOS-GIS based real estate analysis focusing on the housing market (section 3). In the second part of the analysis the drafted proposal is implemented (section 4); preliminary results of the analysis are presented in section 5. . Potsdam, a city of some 153,000 inhabitants, is the regional capital of the federal state of Brandenburg and the former residence of the Prussian kings. Neighbouring the national capital Berlin in the southwest, Potsdam is part of the Greater Metropolitan Area of Berlin. Located in an environment of great touristic value (UNESCO world heritage), the city benefits from its favourble location, resulting in positive demographic development which, in turn, stimulates the housing market.

## 3 FOS GIS-BASED PROCESS CHAIN FOR HOUSING ANALYSIS

When it comes to usability and functional range free open-source geographic information systems have reached a degree of maturity in the last decade that makes them a competitive alternative to proprietary GIS-systems. Similar software capabilities and usability of FOS GIS, no software license fees and direct access to the source code compared against high acquisition and maintainace costs of commercial GIS-systems make FOS-GIS systems an obvious choice for GIS-based real estate analysis. Steiniger and Bocher (2009) provide an overview over existing open source GIS developments. The FOS-GIS based process chain suggested here consists of three modular process components. In all three components the base rent per square meter[2] is used as the most important indicator for the detection of housing market trends. As will be seen a range of FOS GIS has been employed in the study discussed here. This has been done to use the best functionality for the solution of a particular problem available in the FOS GIS under review. To be able to do so is another

[1] Relational database system based online marketplace for the trade of real estate property
[2] High variable service charges and heating costs are not considered in the suggested process flow

major advantage of FOS GIS over proprietary GIS the cost of which generally limit software use to one product.

### 3.1 Component 1: Visual interpretation and descriptive statistics

This component focuses on existing trends and tendencies in the housing market. Trends are analysed by geovisual analysis, i.e. localizing, identifying and interpreting the spatial distribution of the housing markets supply and demand situation in GIS maps. Housing sites usually are presented as individual points (geocoded by x and y co-ordinates). Visualisation of the spatial distribution presents housing sites by classifying the base rent per square metre and the calculation of a simple histogram to identify the distribution of base rent is suggested. Fast calculated descriptive statistics provide a first overview of the status quo in the housing market.This component facilitates as a quick assessment of present housing market conditions. It thus provides a starting point for further in-depth GIS-based analysis.

### 3.2 Component 2: Generation of spatially interpolated surfaces

This component deals with the generation of interpolated surfaces and the rating of administrative levels in respect of their relevance to the housing market. In addition, techniques of spatial declustering are applied in order to produce a spatially differentiated picture of the prevailing trends in the housing market. A main goal of the second process component is the definition of independent (disaggregated of administrative boundaries) real estate sub-markets.

### 3.3 Component 3: Intersection with ancillary data

This component facilitates the augmentation of housing data with ancillary data, such as socio-economic or demographic data, zoning or actual land use data. Combing these data in various ways allows for a progression from an analysis of trends to a GIS-based oberservation of the existing housing space.

## 4 PROCESSING REAL ESTATE DATA

To assess its validity the above workflow has been applied to the housing market in Potsdam. For this purpose large data sets of the nationwide German real estate portal Immobilienscout 24 from 2009 were used. The workflow was implemented with the FOS-GIS systems SAGA GIS,OpenJUMP, Quantum GIS and GRASS GIS.

### 4.1 Data source, data preprocessing and spatial site selection

It is a truism that GIS-based housing analysis requires data about the housing sector. Nevertheless, only a few data sources to analyse the housing market exist (BBSR 2009) -despite the significant social and economic relevance of the housing sector in Germany. About a decade ago real estate portals started to record spatial information on objects for sale and rent in the housing and industrial estate market.If accessible, this data is a valuable statistic basis for GIS based housing analysis. In the present analysis a database extract of the above-mentioned real-estate portal Immobilienscout 24 was used as base data. The dataset consists of approximately 1.4 million housings on offer. Database excerpts of “real estate” databases usually consist of geocoded data (point data with x and y coordinates) that can be processed in GIS systems. Each entry to the database consists of 63 attributes, such location (x and y coordinates), base rent per apartment with and without service costs, living space in square metre, number of rooms, year of construction, purchase price etc.

| GEOX | GEOY | address | living space ($m^2$) | rooms | construction |
|---|---|---|---|---|---|
| **207310** | 2493456 | Am Springbruch | 77,36 | 4 | 1890 |
| **207243** | 2493458 | Am Springbruch | 77,36 | 4 | 1982 |

Fig. 1: Spatial Database (Clip of database table)

A first inspection of the dataset reveals some inconsistencies (missing entries, wrong field types) in the data. Thus the data had to be preprocessed which included data cleaning and updating. This has been done by using the open source database managment program Open office base..Additionally field types etc. had to be adopted for further processing in GIS. To perform this operation the FOS GIS OpenJUMP is employed, as this is the only FOS GIS software allowing to change field types. Existing point data are then geocoded in Lambert Conformal Conic (WIGeoEU) which had to be recoded from to Universal Transverse Mercator (Datum WGS 84, Zone 33 North) for this study. The gdal library implemented in GRASS GIS 6.3 was used for this data transformation. The original database includes records for 4000 housing offers in Potsdam. Preprocessing eventually has reduced this database to a datasetof 1600 individual points in Potsdam which can be used for further analysis. This study database is then subjected to spatial site selection. Only districts with 20 or more individual points were used in the following analysis. Randomized sampling of the data was done, using the Quantum GIS plugin Ftools. Spatial correlation was tested by calculating the semivariance in SAGA GIS.

## 4.2 Visual interpretation, descriptive statistics and spatial declustering

In this phase of the analysis techniques discussed in section 3.1are applied to the Potsdam housing data. First the data is visualised in Quantum GIS using 4 classes (quantiles) of base rent per square meter, next basic descriptive statistic was build to get a first impression of the trends in the Potsdam housing market.

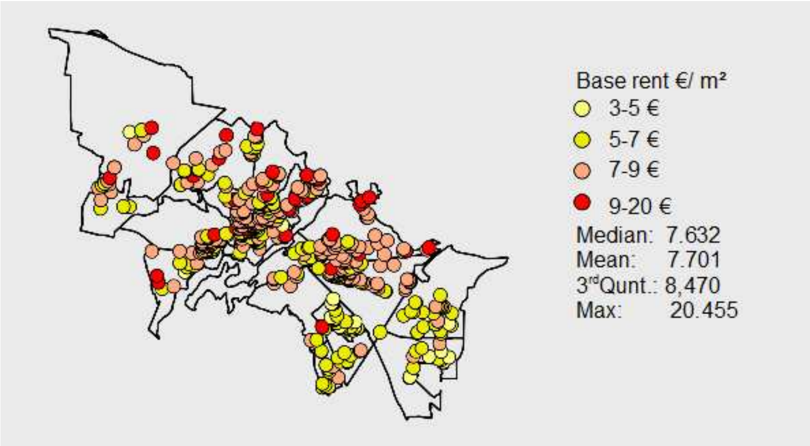


Fig. 2: Potsdam built-up area: base rent of individual points of housing offers

Measurement of statistical parameters, such as mean value, standard diviation or regression, are not reliable a priori, as existing housing data might not be representative of the whole housing market. Spatial declustering techniques help to remove the effects of clustering. Thear are different approaches to declustering, de Smith et al. (2006) provide a useful survey of the existing techniques. In this study we applied a global estimation of the mean by polygonal declustering. Each data point is weighted according to the weight of the respective Voronoi polygon. The influence of clustered points on the global estimation of the arithmetical mean is reduced according to their lower spatial relevance. Using the preprocessed data points, Voronoi Polygons have been calculated using the GRASS GIS command v.voronoi.area. The value (base rent per square meter) of each data point is then weighted by the size of the corresponding Voronoi polygon. The global mean of the base rent per square metre calculated by descriptive statistics is 7.7 €/m², the estimation of the mean by polygonal declusteringis slightly lower and averages 7.2 €/m².

## 4.3 Generating housing market trend surfaces

Based on the randomized data points interpolated trend surfaces of the base rent per square meter have been generated in a consecutive step. To do so various interpolation techniques of the FOS GIS GRASS GIS and SAGA GIS have been applied. Spatial interpolation techniques estimate values from a limited number of sampling points and create continous surfaces from point data following Tobler´s first law of geography (Tobler 1970) which says that everything is related to everything, but objects closer in space are also closer with respect to their attribute values. A multitude of interpolation methods exist, see Albrecht (2007) and de Smith et. al. (2006) for an overview of existing methods. In this analysis the deterministic approaches of IDW (Inverse Distance weighted) and Spline together with the stochastic, geostatistic method of Ordinary Kriging have been applied. Both SAGA GIS and GRASS FOS-GIS provide functionalities for spatial interpolation. While Kriging methods are available in SAGA GIS, GRASS GIS provides spline interpolation functions. IDW is the method of choice when a quick generation of surfaces is required. One major disadvantage of IDW, however is the potential occurence of bulls eye effects that cause the data not to be extrapolated (de Smith et.al. 2006) Spline interpolation is computationally intensive, but results are more accurate and the surfaces generated are smoother. To generate a trend surface with maximum flexibilty and accuracy the use of Ordinary Kriging is recommended. In Ordinary Kriging the relationship of all data points and the influence of each point on each other point is calculated (Albrecht, 2007). Ordinary Kriging has applied using SAGA GIS. SAGA GIS allows the construction of a variogram. Unsampled point values are estimated by calculating the linear weighted average of neighboring points.Weights are determined by a fitted variogram (de Smith et. al. 2006). In the present analysis linear regression regression has been used as variogram model.

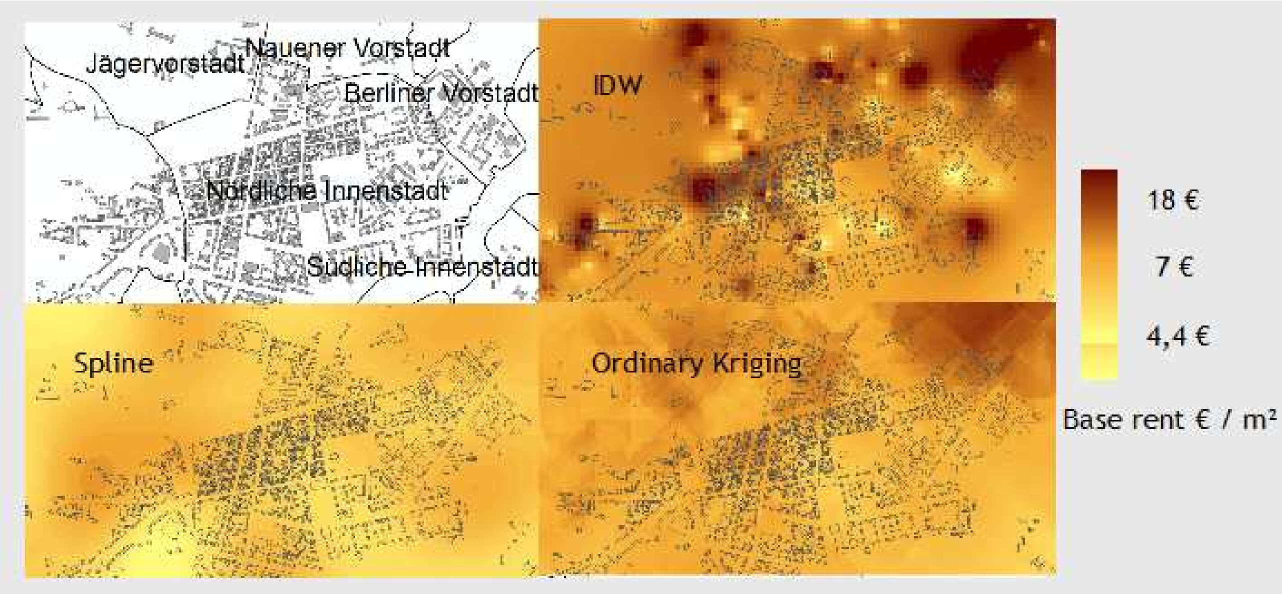


Fig. 3: Subset of the investigation area and comparison of 3 interpolation methods. Strong bull eye effects can be observed in the IDW example.

As has been mentioned above (3.3) ancillary data was introduced into the analysis. The types of built up area inPotsdam has been mapped according to the Brandenburg zoning plan. A total of 8 classes of building zones (residential area, commercial and industrial area, mixed area, touristic sites, public infrastructure) have been mapped. Mapping was performed by interpreting satellite and orthophoto imagery from 2006 -2009. The resulting data were then combined with building blocks extracted from the OpenStreetMap (OSM) project which provides free geographic data. In a subsequent step the class residential area has been subdivided into further subclasses of housing types and combined with socioeconomic data in order to get from a trend analysis to a GIS based oberservation of the housing space (see section 3.3). Finally, the interpolation results of Kriging have been merged with the OSM building blocks to mask all geo-objects with no relevance to the housing market.

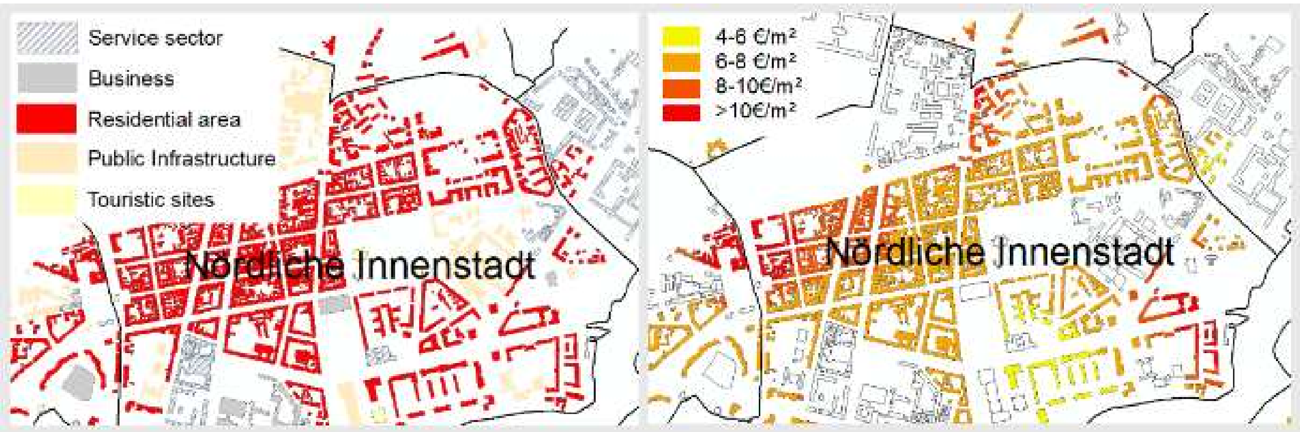


Fig. 4: Settlement structure and interpolation results intersected with identified with building blocks identified as housing.

## 5 TRENDS IN THE POTSDAM HOUSING MARKET: FIRST RESULTS

### 5.1 Investigation area Potsdam

The need for a GIS-based housing market analysis, as presented in this contribution, is evident when considering Potsdam´s position as the second most expensive place to live in East Germany after Berlin (TLG Properties 2009).

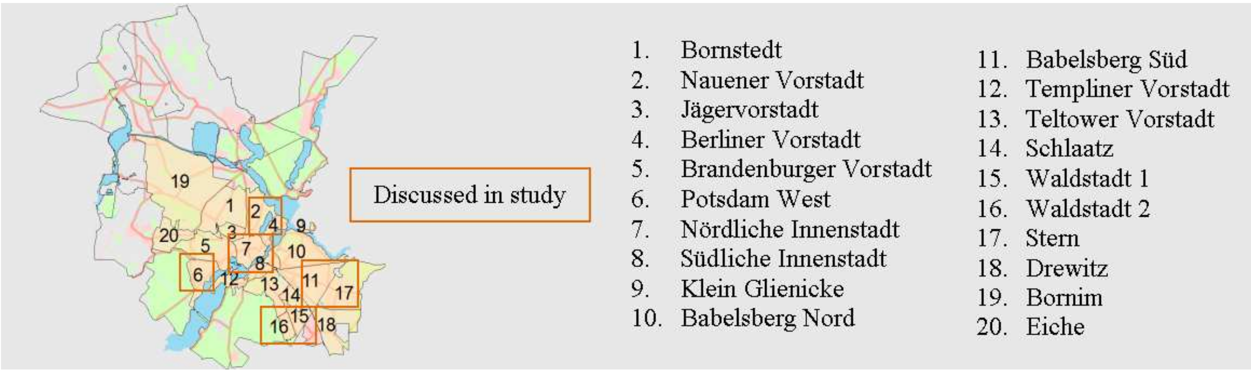


Fig. 5: Administrative division of Potsdam, city districts under investigation (orange rectangle, discussed in section 5.2)

### 5.2 Preliminary results

In Potsdam the range of the interpolated base rent per square meter, using the ordinary kriging method, varies from 4,88 € to14 €/m². Major differences can be observed between the southern districts of Waldstadt, Schlaatz and am Stern on the one hand and high base rents in Babelsberg, in the Innenstadt and the Berliner and Nauener Vorstadt quarters on the other. The average base rent per square meter calculated by using techniques of spatial declustering (see section 4.2) averages 7.2 €/m² . The global mean for a four-room apartment of about 90m² averages 1,076 €/m² , the global mean for a two-room apartment of about 65 m² averages 490 €/m² excluding service costs. Extreme values for four-room apartments can be observed in the Jaegervorstadt and Berliner Vorstadt districts, where apartments of a size of about 160 m² have been offered at prices of exceeding1600 €/m² without service costs. Low base rents of 4 to 6 €/m² can be found in the Potsdam-West district, particularly in Zeppelin Strasse, where the residential area mainly consists of prefab buildings built by East Germany in the 1970s. In the northern part of Potsdam-West base rents per square metre are much higher, up to 9 €/m². Spatial patterns of particular interest can be found in Babelsberg. In recent years in Babelsberg North and South a substantial amount of reconstruction of old buildings can be observed. Very high base rents per square meter have been found in the eastern part of Babelsberg with strong clusters at the border of neighbouring Griebnitzsee district as well as in a corridor on the railway line dividing Babelsberg North and Babelsberg South, i.e. up10 €/m². The highest base rents per square metre in the area under investigation is found in the Innenstadt district and the touristic centre (Durch Quarter), as well as Berliner, and Nauener Vorstadt with up to 11€/m² .

## 6 CONCLUSION

This paper presents a GIS-based housing analysis using FOS GIS. The study demonstrates clearly that such analysis is feasible yielding results that compare to to of commercial GIS. To conduct this FOS GIS analysis, a generic workflow was designed and evaluated by using Potsdam real estate data. Because of its generic nature the processing chain presented here can similarly be employed in spatial planning or municipal administraition and related tasks. The first results of this ongoing project indicate that additional ancillary data and reference data of different observation periods are required to progress from an analysis of current trends in the housing market to a fact-based prognosis of future real estate developments. First steps have already been undertaken to channel the study into this direction.

## 7 REFERENCES

ALBRECHT, J. (2007), Key concepts & Techniques in GIS.

BORCHERT, A. (2006), Geographische Informationssysteme in der Immobilienwirtschaft. In: STANDORT-Zeitschrift für Angewandte Geographie, Vol. 3, pp. 127-131..

DE SMITH, M.J., GOODCHILD, M.F., LONGLEY, A.L. (2007), Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools.

BBSR (Bundesinstitut für Bau-, Stadt- und Raumforschung).Wohnungsmarktbeobachtung. (2010), http://www.bbr.bund.de.
STEINIGER, S., BORCHER, E. (2009); , An overview on current free and open source desktop GIS developments. In: Int. J. of Geographical Information Science 23(10), pp. 1345-1370.
TLG Immobilien. (2009), Immobilienmarkt Ostdeutschland 2009, Marktdaten der kreisfreien Städte und Berlin.
TOBLER, W. (1970), A computer movie simulating urban growth in the Detroit region". In: Economic Geography, 46(2), pp. 234-240.