In recent years, information extracted from geosocial media feeds like Twitter and Flickr have been increasingly used to investigate the everyday lives of people. This is often done by applying established spatial analysis methods, predating the emergence of these kinds of data. From an analytical viewpoint, it remains unclear whether or not these methods are applicable to these new kinds of data. This thesis challenges this often implicitly made assumption and investigates to what extent we can use established geospatial methodology to investigate user-generated geographic information.

The research conducted focuses on two kinds of spatial-statistical methods: hot-spot techniques for disclosing spatial accumulations of outliers, and spatial autocorrelation estimators for characterising how observations are driven by geography. Spatial heterogeneity and spatial overlap are thereby assumed, because people may communicate any content from often similar locations. The results obtained indicate that hot-spot measures are prone to propagating small-scale effects to larger scales. This mixes up different phenomena, turning measures into functions of scale. Inferences drawn are then biased, and driven by events taking place at scales beyond the analysis scope. Similarly, characterising spatial patterns through autocorrelation estimators is affected strongly by the mix of different, overlapping phenomena in terms of disclosure of spurious patterns. False spatial relations are made between potentially unrelated phenomena, invalidating findings indicative of a mostly meaningless mixture.

This dissertation also makes methodological contributions. One method investigated is an adapted hot-spot technique based on the idea of stratifying places. Stratification is thereby done in terms of scales, which allows investigating each of them separately. A comparison with an established technique has revealed superior results with respect to both statistical aspects and interpretation. Another method proposed is a hypothesis test about spatial heterogeneity. The underlying idea is to find out to what extent geography influences the variance within observations. This test makes it possible to better understand the spatial overlap found in user-generated geographic information, and therefore digital representations of places on geosocial media.

The conclusions of this thesis are of theoretical and practical relevance. Geography, regional science and related disciplines are supported by a profound understanding of empirical results obtained from user-generated geographic information. Statistics benefits in terms of a thorough characterisation of the statistical characteristics of relations between geometrically overlapping point sets generated from different processes. The most important contribution made is to GIScience, by providing a step towards the statistical analysis of subjective, place-based information, one of the topics currently gaining momentum.