

Усовершенствование массовой кадастровой оценки недвижимости с использованием спонтанных данных

Введение

В связи с переходом к расчету налога на недвижимость в Российской Федерации на основе кадастровой стоимости вместо инвентаризационной резко вырос общественный запрос на качество кадастровой оценки. При этом современная система расчета в ее нынешнем виде справедливо вызывает множество нареканий. Систематические проблемы сбора информации об объектах кадастрового учета, потери данных при переходе от одних носителей информации к другим и слабая в методическом отношении расчетная часть, неспособная к учету влияния на ценообразование сложных общественных систем современного города, привели к возникновению большого разрыва между кадастровой стоимостью объектов недвижимости и существующим рыночным уровнем цен.

Сократить этот разрыв может использование так называемых спонтанных данных, генерируемых обществом как побочный продукт жизнедеятельности. Использование этих данных позволяет принять во внимание нематериальную ценность места, выражающуюся через его образ, проходящие человеческие потоки, сложившиеся стереотипы и многие другие факторы, которые играют большую роль при оценке стоимости недвижимости.

Помимо этого, благодаря более точной координатной привязке, появляется возможность перейти к оценке объектов не поквартально, а пообъектно, а тот факт, что сбор спонтанных данных, как правило, не требует разворачивания специальной инфраструктуры, позволяет сократить штат, необходимый для проведения кадастровой оценки объектов недвижимости а также производить переоценку с большей периодичностью, что позволит оперативно реагировать на изменения в городской среде. Таким образом, создание системы расчета, включающей индикаторы, построенные на основе спонтанных данных, выгодно отличается от производимых сейчас расчетов.

Международный контекст

Основные ограничения текущих методов оценки недвижимости связаны с недоиспользованием внеэкономических факторов, таких как значимость места для сообщества, а также отсутствием пространственной и временной детализации, в частности при контроле за функциональной нагрузкой строений.

Спонтанные данные могут быть полезны при использовании гедонических регрессий и индексов для оценки недвижимости, то есть для оценки определенных характеристик или попросту поправки на качество. Гедоническое ценообразование наиболее оправдано именно в пространственном континууме. Например, высокий уровень преступности в районе приводит к снижению цен на недвижимость, так как люди платят не только за сам объект, но и за присущие ему и пространству вокруг него характеристики.

Наиболее значимым результатом анализа спонтанных данных является оценка восприятия пространства сообществом. Качественные показатели могут наравне с количественными быть использованы в разных масштабах - от здания, до города или даже страны.

По мнению некоторых исследователей, в современном мире субъективная оценка качества жизни (well-being) может сказать об обществе больше, чем экономические

индикаторы¹. Американские психологи отмечают, что “уровень финансового благосостояния был важен на начальном этапе экономического развития человечества, когда его главной заботой были базовые потребности. Однако со временем самочувствие, социальные связи и удовлетворение работой вышли на первый план”. Поэтому известны случаи, когда, несмотря на рост доходов, удовлетворенность жизнью у населения сокращалась.

Другое исследование показало, что компьютерный анализ тональности текста (sentiment analysis) сильно коррелирует с оценкой удовлетворенности людей². Основываясь на этом выводе, ученые из университета Орегона предложили рассчитать уровень “валового национального счастья” на основе активности 100 миллионов пользователей Facebook. В Британии смогли увеличить масштаб подобных разработок и выявили корреляцию между тональностями постов в Twitter и экономическими показателями различных районов Лондона³.

Американский экономист Чан Куанг предложил использовать популярность сервисов в том или ином районе на Yelp как поправку на качество, которая может стать переменной при расчете стоимости недвижимости. Он обнаружил связь не только количества, но и качества инфраструктуры потребления (в частности, ресторанов) с ценами на жилье, основываясь на рыночных данных о транзакциях⁴.

В Институте Новой Англии проанализировали посты в Twitter на предмет эмоций, которые вызывают у людей те или иные места. Так, самые позитивные переживания связаны с парками, самые негативные - с транспортно-пересадочными узлами. Кроме того, были выявлены значительные эмоциональные колебания в медицинских центрах, на кладбищах и возле очистных сооружений⁵. Традиционно близость к подобным территориям влияет на стоимость недвижимости.

Соответственно, можно предположить, что спонтанные данные из социальных сетей являются репрезентативными для реальных процессов на рынке недвижимости.

Спонтанные данные могут быть использованы не только для оценки существующих районов, но и для проектирования будущих. Для контроля и дизайна городской среды необходимо знание о функциональном использовании земли и идентификация значимых объектов. Для сбора подобной информации традиционно используют метод непосредственного (прямого) наблюдения либо анкетирование жителей. Данные техники имеют ограничения ввиду дороговизны, а также нежелания жителей предоставлять информацию. Ряд исследователей предлагает использовать социальные сети с геопривязкой для определения вида землепользования. Подобные исследования проводились с использованием данных социальной сети Foursquare, в которой существуют predefined категории мест (ресторан, кинотеатр и пр)⁶. Похожее исследование на основе фотографий из Flickr позволило выявить достопримечательности как места с наибольшей концентрацией фотографий⁷. Для оценки правильности этого метода были проведены прямые наблюдения на местности.

¹ Diener, E., Seligman, M. E. P. 2004. “Beyond money: Toward an Economy of Well-Being”

http://internal.psychology.illinois.edu/~ediener/Documents/Diener-Seligman_2004.pdf

² Kramer A. D. I. 2010. “An Unobtrusive Behavioral Model of “Gross National Happiness”

<http://www.uvm.edu/~pdodds/files/papers/others/2010/kramer2010a.pdf>

³ Quercia, D., Ellis, J., Capra, L., Crowcroft, J. 2012. “Tracking “Gross Community Happiness” from Tweets”.

<https://www.cl.cam.ac.uk/research/srg/netos/papers/p965-quercia.pdf>

⁴ Chun Kuang. 2015. “Does Quality Matter in Local Consumption Amenities? An Empirical Investigation with Yelp”

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2606775

⁵ Bertrand K.Z., Bialik M., Virdee K., Andreas Gros, Bar-Yam Y. 2013. “Sentiment in New York City: A High

Resolution Spatial and Temporal View” <http://necsi.edu/research/social/newyork/newyorksentiment.pdf>

⁶ Noulas A., Scellato S., Mascolo C., Pontil M. 2011. “Exploiting semantic annotations for clustering geographic

areas and users in location-based social networks” <https://www.cl.cam.ac.uk/~cm542/papers/SMW11.pdf>

⁷ Crandall D. J., Backstrom L., Huttenlocher D., Kleinberg J. 2009. “Mapping the world’s photos”

<https://www.cs.cornell.edu/home/kleinber/www09-photos.pdf>

Соотнесение этих результатов подтвердило валидность использования спонтанных данных.

Схожее исследование, основанное на постах в Twitter, выявило аналогичный список⁸, что говорит об оправданности использования этих данных. Авторы предлагают использовать данную технику для определения новых точек притяжения и появляющихся достопримечательностях в режиме реального времени. Вторая часть этого исследования посвящена определению типа использования той или иной территории. На основе информации о пространстве и времени постов были идентифицированы 4 типа использования: жилое, коммерческое, парки и развлечения, а также ночная жизнь. Первые 3 кластера совпали с официальным зонированием на 62, 77 и 79% соответственно. Четвертый тип - "ночная жизнь" - отсутствует в городском регламенте, хотя влияет на показатели шума, безопасности и чистоты улиц. Таким образом, использование спонтанных данных является эффективным дополнением к информации о градостроительном зонировании.

Моделирование кадастровой стоимости коммерческой недвижимости в российских городах

В рамках проекта в первую очередь был определен спектр спонтанных данных, которые можно использовать для массовой оценки данного вида объектов недвижимости. Среди прочих, были отобраны данные Google Routes, позволяющие косвенно оценить параметры мобильности населения по данным индекса пробок, данные социальной сети Вконтакте, которые выступили как прокси объема и динамики активности людей в определенных местах, оценки отношения к местам через тональность высказываний в постах, а также как часть общей характеристики онлайн-предложения недвижимости и спроса на нее через посты с предложениями об аренде.

Данные Google Places и Open Street Map были использованы для получения информации о количестве объектов обслуживания в определенной местности, дистанциях между объектами и о связывающей их транспортной сети, графике их работы, уровне цен и отзывах о них посетителей.

Базы данных компании AIRSTATE (индекс загрязненности воздуха, учитывающий концентрации SO₂, CO, пыли и т.п) и тепловизионные спутниковые снимки USGS EROS позволили составить картину о метеорологических условиях и экологических загрязнениях.

Для проведения расчетов, связанных с кадастровой стоимостью объектов недвижимости, были задействованы данные агентств недвижимости.

При построении гедонистической регрессии (зависимости цен от факторов) необходимо обеспечить независимость факторов друг от друга, чтобы избежать проблемы «кажущейся регрессии» и смещения полученных оценок коэффициентов, что снизит качество оценки. Например, оценка платежеспособности потока людей исходя из сочетания данных о мобильности и информации о финансовых транзакциях, совершенных в определенном месте.

После определения спектра данных была построена модель стоимости объектов недвижимости, в которой в качестве зависимых переменных использовать, помимо традиционных, показатели, рассчитанные на основе спонтанных данных.

Точность модели оценивалась на основе сравнения результатов расчета с реальными рыночными ценами. Модель была протестирована на пяти российских городах (Нижний Новгород, Уфа, Хабаровск, Казань, Ярославль), и так как тестирование

⁸ Frias-Martinez V., Soto V., Hohwald H., Frias-Martinez E. 2012. "Characterizing Urban Landscapes using Geolocated Tweets" <https://pdfs.semanticscholar.org/5062/feceb81c8f8e250bfbe67224dde2e434bfd1.pdf>

показало, что вес показателя варьируется от города к городу, конечные модели имели разные наборы показателей.

Полученные расчеты показали, что информация, получаемая с использованием Интернет-ресурсов, таких как социальные сети и навигационные приложения, выгодно отличается от традиционной статистики.

Так, расчеты показали, что в Хабаровске кадастровая стоимость нежилых объектов недвижимости занижена по всему городу. Результаты моделирования с использованием спонтанных данных позволили получить сходную по точности с традиционным кадастром оценку стоимости нежилых объектов недвижимости.

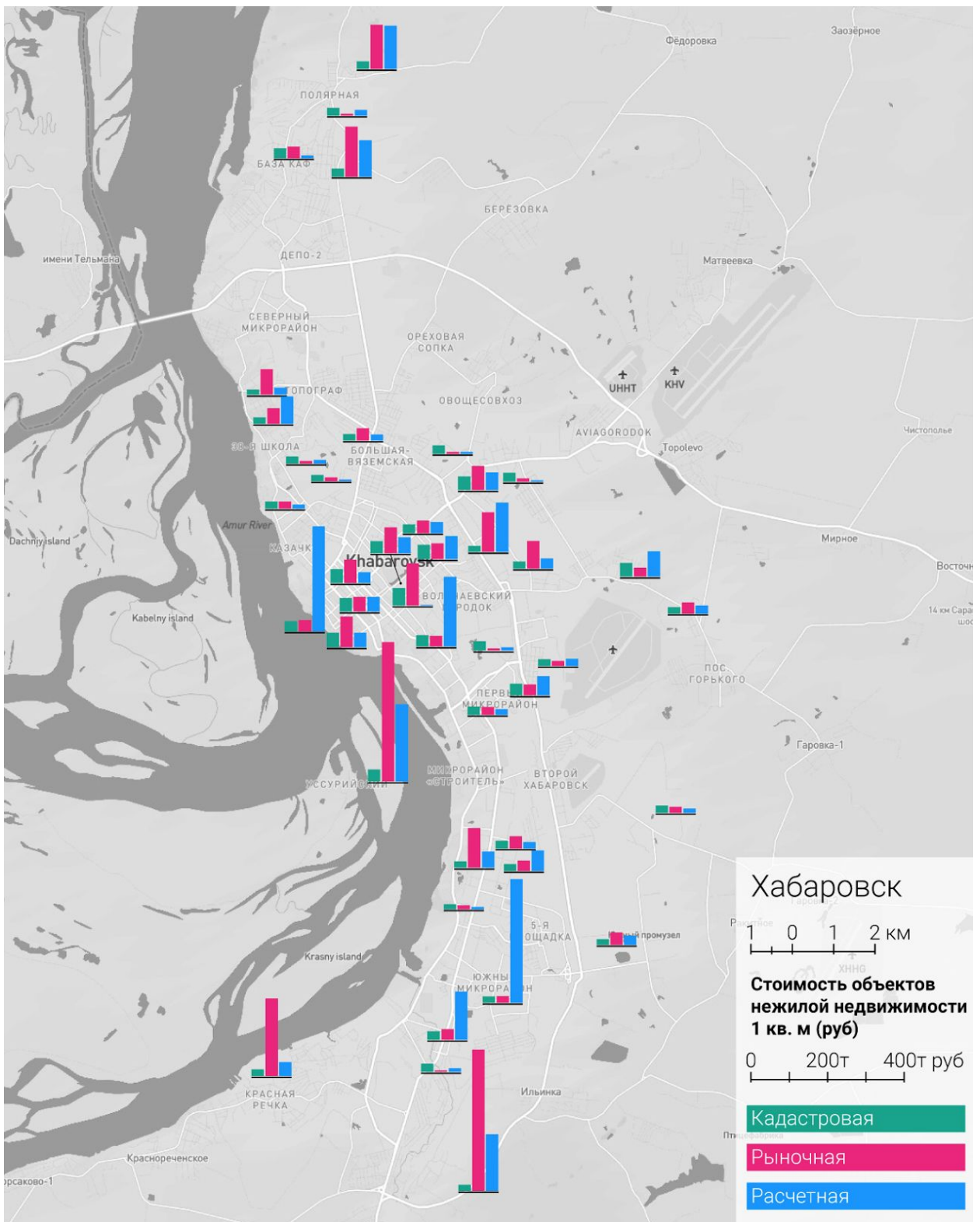


Рис 1. Стоимость 1 кв. м объектов нежилй недвижимости (руб.), г. Хабаровск

В Нижнем Новгороде существующие кадастровые цены также занижены относительно рыночного уровня. При этом, если в целом по городу расчетная цена была намного ближе к рыночной, чем кадастровая, в историческом центре она ее сильно превысила, что может говорить о недооцененности существующим рынком наиболее выгодных с точки зрения размещения объектов города.

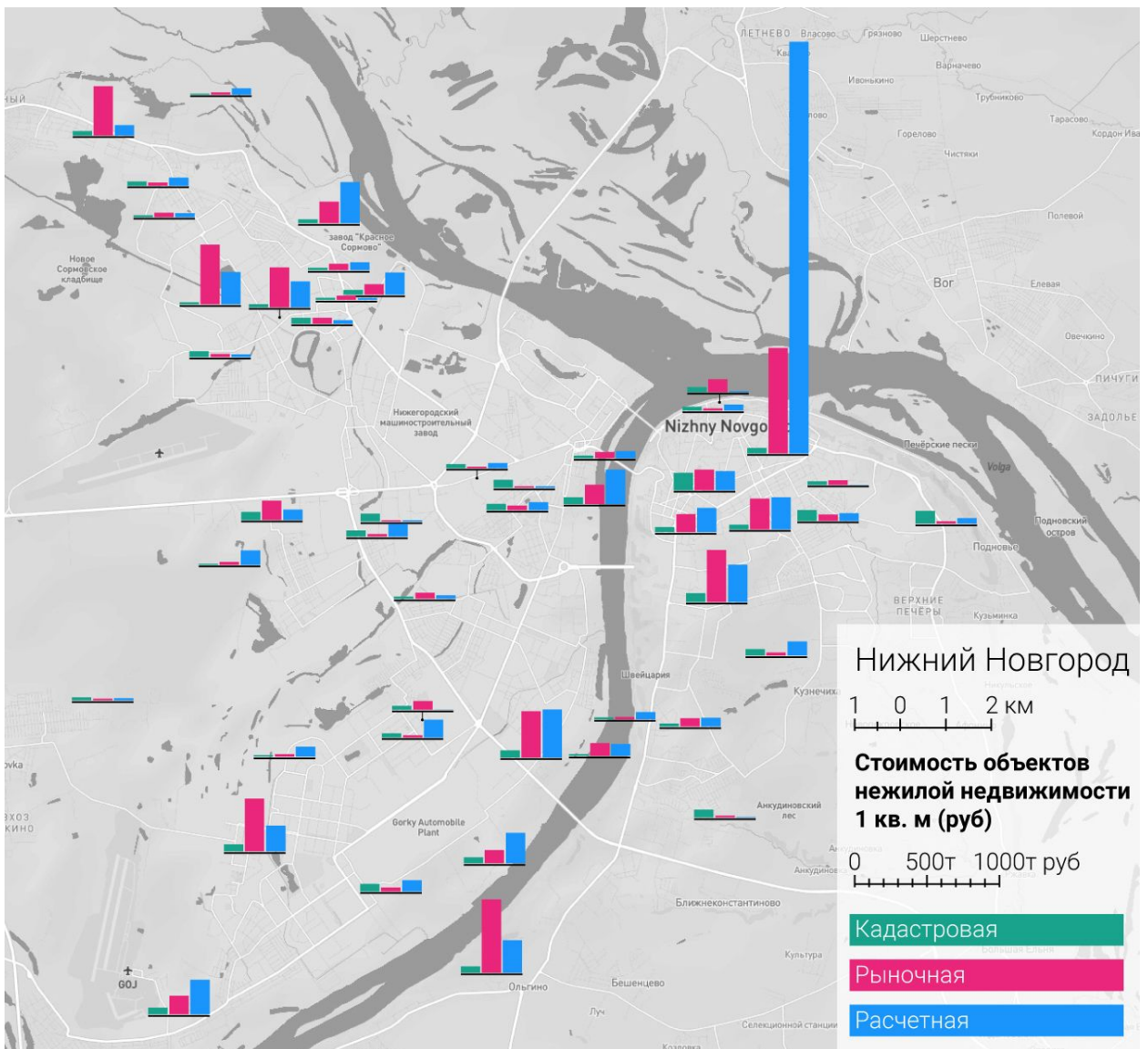


Рис 2. Стоимость 1 кв. м объектов нежилкой недвижимости (руб.), г. Нижний Новгород

Существующая кадастровая оценка недвижимости Ярославля также сильно занижена относительно рыночной. В случае Ярославля наиболее значимыми оказались переменные, маркирующие городскую селитебную периферию (влияют на стоимость коммерческой недвижимости негативно). Помимо этого, в Ярославле любопытны три аномальных точки, где и модельный расчет, и кадастровая стоимость равно сильно отличались от цены сделки. Можно предположить, что эти сделки по какой-то причине были заключены на нерыночных основаниях.



Рис 3. Стоимость 1 кв. м объектов нежилой недвижимости (руб.), г. Ярославль

В Уфе кадастровая стоимость нежилых объектов недвижимости завышена по всему городу, особенно в центральных районах. Интересно, что в Уфе проявилась не встречавшаяся в предыдущих городах метрика центральности – через расстояние до спортивных объектов. Наличие в Уфе сильных и популярных хоккейной и футбольной команд определяет тот факт, что спортивные события действительно перестраивают жизнь города, ориентируя потоки на стадионы и в окружающие их районы.

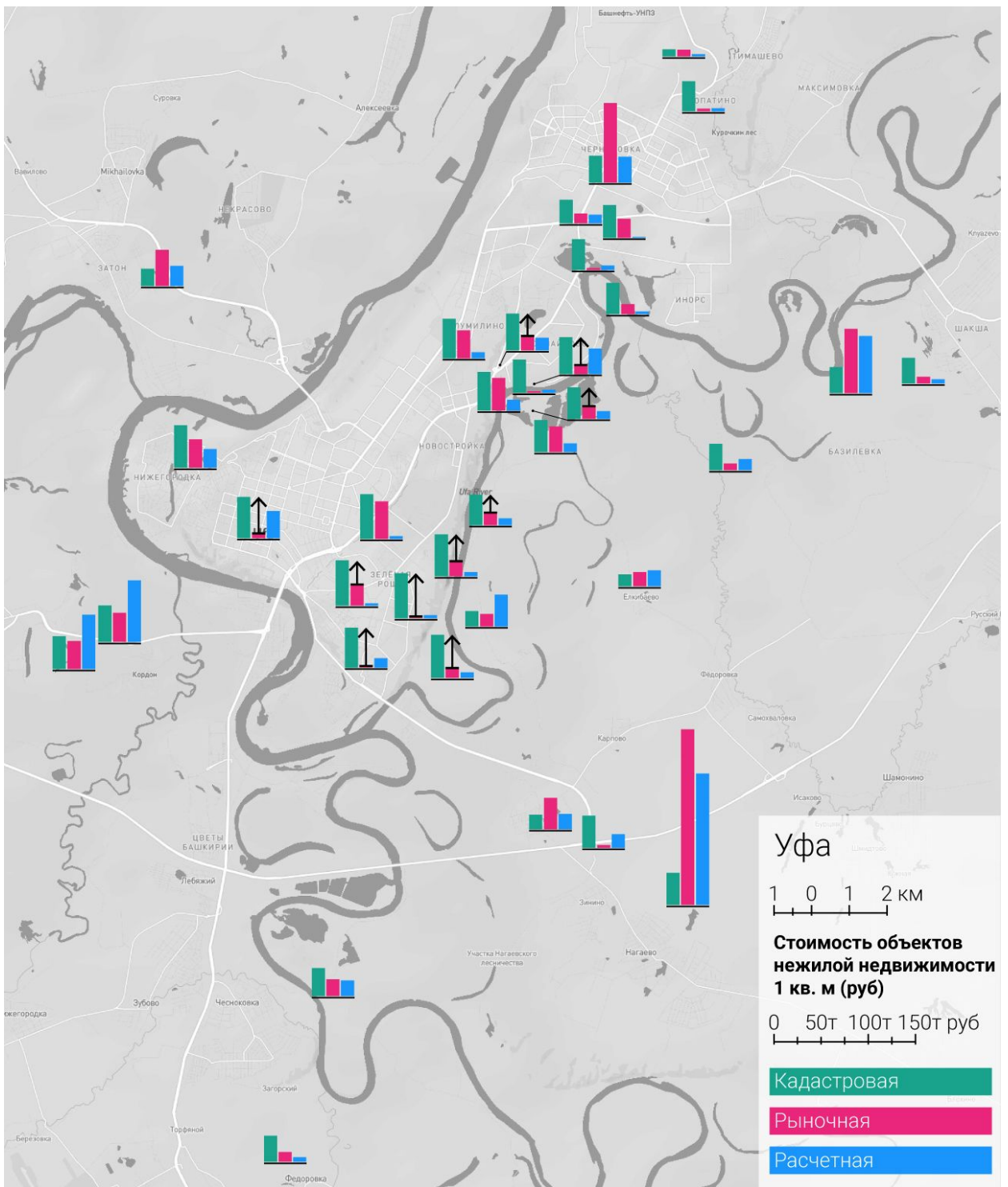


Рис 4. Стоимость 1 кв. м объектов нежилой недвижимости (руб.), г. Уфа

В Казани, где кадастровая стоимость нежилых объектов недвижимости занижена по всему городу, особенно в центральных районах, модель сработала лучше всего в Вахитовском районе, который оказывается наиболее проблемным для традиционного кадастра и в котором сконцентрированы многофункциональные торгово-развлекательные и офисные центры, обладающие сложной структурой ценообразования.

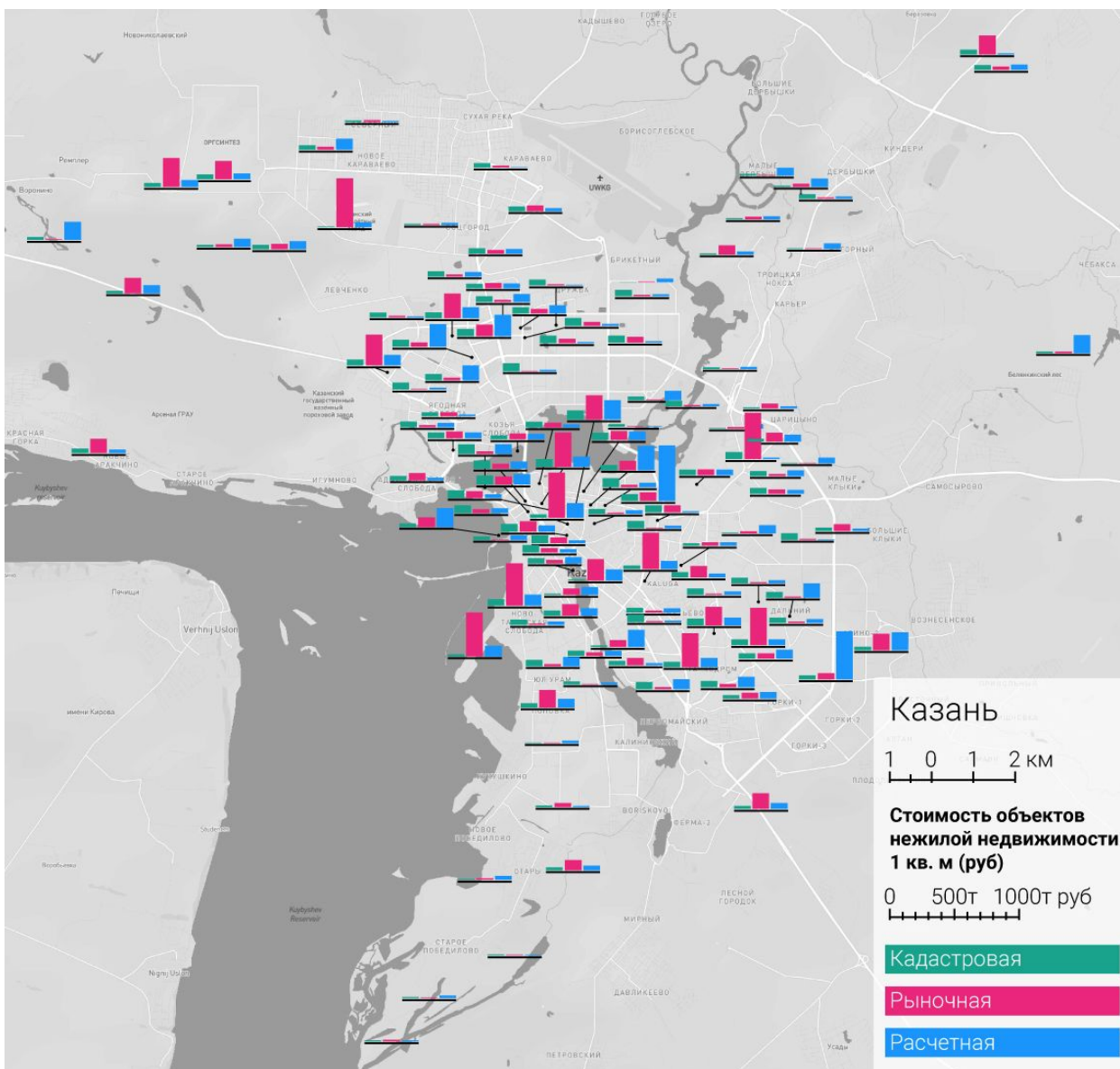


Рис 5. Стоимость 1 кв. м объектов нежилой недвижимости (руб.), г. Казань

Вывод

Исследование подтвердило, что будучи выверенными городским сообществом, им же поддерживаемыми и обновляемыми, спонтанные данные гораздо актуальнее и обладают большим пространственным разрешением, нежели традиционная статистика, что привело к закономерному результату: точность расчетов в большинстве городов повысилась. Расчетные стоимости стали ближе к рыночным показателям.

При этом показатели, указывающие на поведение сообществ, такие как уровень активности в социальных сетях (отражающий востребованные населением места), уровень температуры (свидетельствующий об уровне активности), пробки и данные по транспортной доступности доказали свою значимость в ценообразовании коммерческой недвижимости в исследуемых городах. Приведенные примеры анализа для задач оценки стоимости недвижимости являются частью аналитического пакета Habidatum для городских и региональных планировщиков. Если у вас есть вопросы, пожалуйста, высылайте их по адресу ask@habidatum.com

Массовая оценка недвижимости для целей государственного кадастрового учета является частью методики повышения качества городского управления с использованием

спонтанных данных, разработанной Habidatum совместно с Институтом экономики города по заказу Федерального центра проектного финансирования. Проект также включает в себя анализ эффективности функционирования общественного транспорта и измерение миграционных процессов в городских агломерациях.