Пространственная неоднородность относительных характеристик детализации карты как функция от степени хозяйственного освоения территории

# Постановка задачи

**Задача:** Разработка метрик и программного обеспечения для оценки детализации множества пространственных объектов, а также детализации карты (базы пространственных данных) в целом на основе учета геометрических, семантических и символьных (применяемых при отображении) параметров. Оценка устойчивости и сопоставимости рассчитанных значений метрик для фрагментов данных, извлеченных из топографических карт одного и разных масштабов.

В рамках работ 2019 г были проведены эксперименты с целью выработки понимания того, каким образом можно оценить детализацию карты в числовой форме. Детализация карт регламентируется только для топографических продуктов, однако явной числовой шкалы уровней детализации пока что не существует. С одной стороны, для каждого масштаба разработаны цензы отбора, устанавливающие по геометрическим и семантическим параметрам минимальный порог вхождения объектов на карту и модель их пространственной локализации (точка, линия, полигоны). С другой стороны, устанавливаются нормы отбора, регламентирующие максимально допустимое количество объектов на единицу площади. Наконец, ограничения визуального восприятия устанавливают графическую точность изображения порядка 0,1 мм, что в свою очередь регламентирует минимальный размер деталей (изгибов) в составе изображаемых линий. Эти три компоненты — цензы, нормы и графическая точность — в совокупности оказывают ключевое влияние на детализацию представления географической среды на карте.

Проведенные в 2018 г. эксперименты показали, что существуют индивидуальные характеристики объектов, такие как средняя площадь изгибов, которые систематически изменяются при переходе с одного масштаба (уровня детализации) на другой. Однако при переходе к оценке детализации однородных множеств пространственных объектов (слоёв), таких как объекты гидрографии или транспортные коммуникации, а также оценке детализации карты как совокупности этих множеств, необходимо учитывать также количество объектов на единицу площади и их топологические отношения. При уменьшении масштаба карты в связи с отбором количество объектов на единицу площади поврехности Земли уменьшается, что по логике должно свидетельствовать об уменьшении детализации представления. Соответствующим образом должно уменьшаться и количество пересечений между объектами внутри одного слоя и между слоями. Это утверждение, однако, справедливо при сравнении карт, покрывающих одну и ту же территорию в разных масштабах. Возникает вопрос, а можно ли на основе этих характеристик надежно определить уровень детализации карты? Исследованию этой надежности был посвящен настоящий раздел работ 2019 года.

# Анализ геометрических характеристик детализации карты

Для анализа возможных характеристик детализации карты мы воспользовались плотностными показателями, такими как как плотность объектов, плотность вершин, плотность пересечений между объектами, а также показателями размера, такими как средняя ширина и высота изгиба линии (применима как к линейным объектам, так и к площадным), а также минимальный средний размер объекта. Для выполнения эксперимента мы выбрали 9 фрагментов цифровых топографических карт масштабов 1:200 000, 1:500 000 и 1:1 000 000 с охватом порядка в окрестностях городов Воронеж, Грозный, Москва, Петрозаводск, Ростов-на-Дону, Салехард, Сургут, Уфа и Шарья.

Использовались данные по социально-экономическим объектам (населенные пункты и их структура, транспорт, инженерные коммуникации, административно-территориальное деление). Результаты по трем относительным (плотностным) геометрическим характеристикам представлены на Рис. 1.

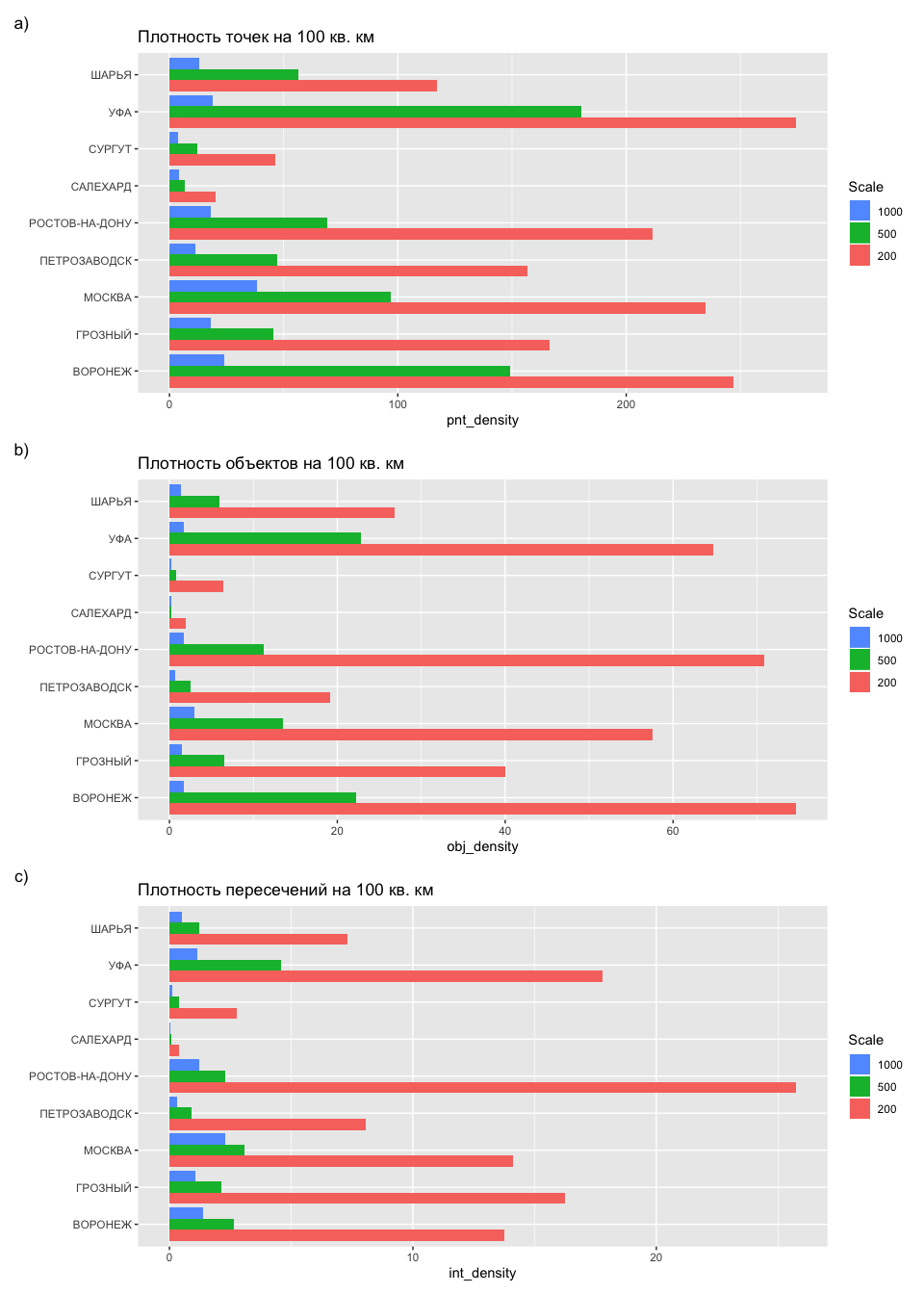


Рис. 1: Относительные характеристики геометрической детализации карты (плотность)

Анализ полученных результатов показывает, что плотностные характеристики не являются надежными индикаторами уровня детализации. Даже в пределах одного масштаба плотность размещения объектов может отличаться на порядки. Например, для участка Воронеж в масштабе 1:200 000 плотность объектов составляет на , в то время как для участка Салехард всего лишь . Подобный разброс наблюдается во всех масштабах и для всех исследуемых величин (плотность точек, объектов, пересечений), хотя следует отметить что с уменьшением масштаба разброс становится менее радикальным, что наглядно демонстрируют диаграммы размаха (Рис. 2).

Выполнение теста Стьюдента на отличие средних между масштабными группами показывает приемлемую статистическую надежность этих отличий (Таблица 1), однако огромный разброс самих значений в пределах каждого масштаба вызывает сомнения в надежности использования данных характеристик.

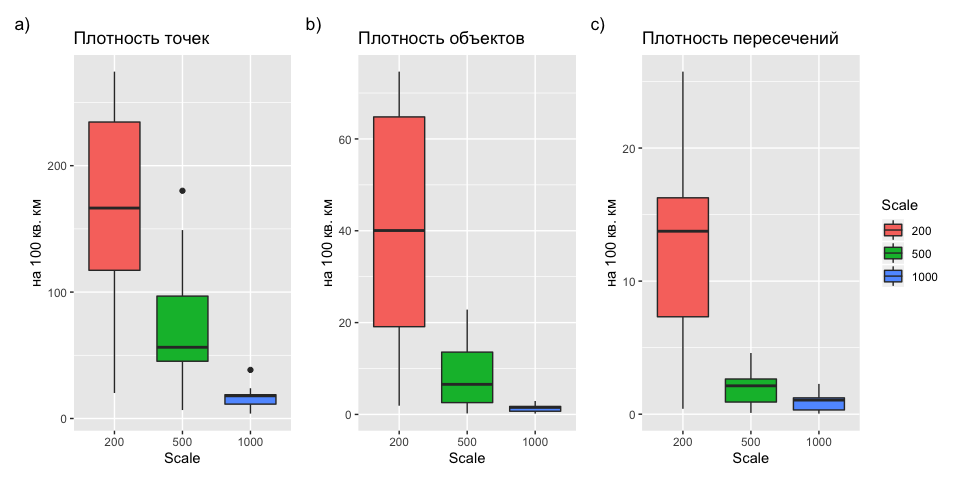


Рис. 2: Разброс относительных характеристик геометрической детализации карты (плотность)

Таблица 1: p-значения теста Стьюдента на равенство средних для относительных характеристик детализации

|  |  |  |  |
| --- | --- | --- | --- |
| param | scale1 | scale2 | p-value |
| pnt\_density | 200 | 500 | 0.0235970 |
| pnt\_density | 500 | 1000 | 0.0198256 |
| obj\_density | 200 | 500 | 0.0110676 |
| obj\_density | 500 | 1000 | 0.0212643 |
| int\_density | 200 | 500 | 0.0056675 |
| int\_density | 500 | 1000 | 0.0785035 |

Абсолютные характеристики детализации, представленные на Рис. 3, демонстрируют более устойчивое поведение, однако и в этом случае диапазон разброса значений может отличаться в несколько раз: для тех же участков Салехард и Воронеж в масштабе 1:200 000 отношения средних ширин и высот составляет и соответственно. Наиболее стабильный характер имеет минимальная средняя площадь полигонального объекта (Рис. 3c), вычисленная как минимум из средних площадей объектов каждого слоя. Тем не менее, и по данной характеристике могут наблюдаться выбросы. Так, например, в масштабе 1:1 000 000 для участка Салехард минимальная площадь полигонального объекта в 136 раз больше следующего по размеру минимальной средней площади участка Сургут.

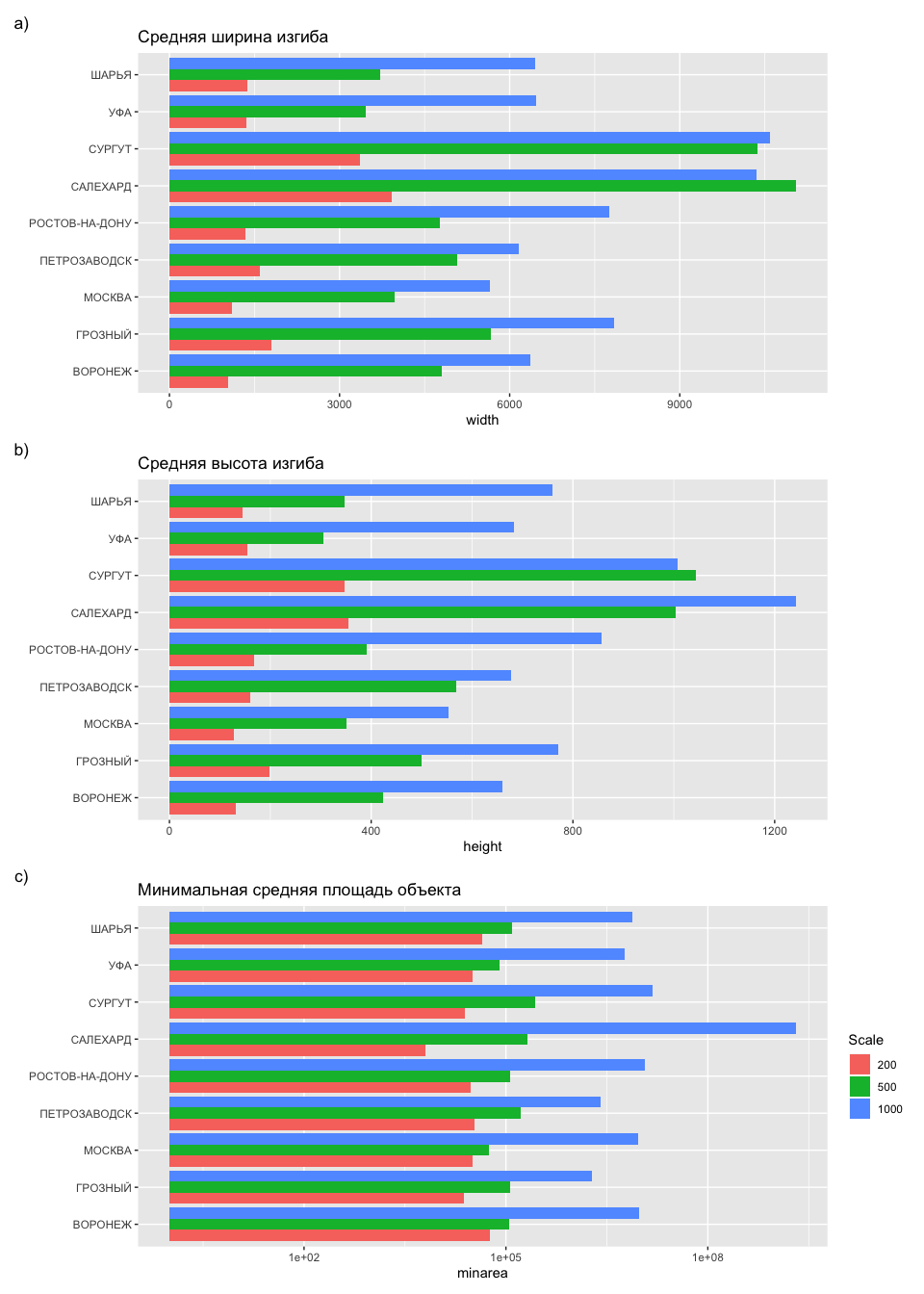


Рис. 3: Абсолютные характеристики геометрической детализации карты (размеры)

В целом более качественная разделимость уровней детализации по абсолютным характеристикам в сравнении с относительными подтверждается диаграммами размаха, представленными на Рис. 4. Тем не менее, статистическая проверка разделимости уровней детализации по данным характеристикам показывает, что и здесь не во всех случаях наблюдаются значимые отличия в средних (Таблица 1).

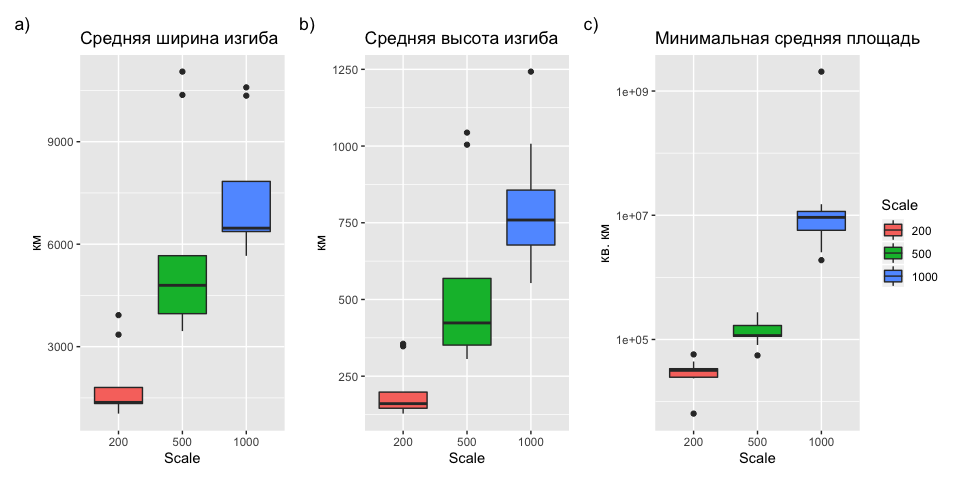


Рис. 4: Разброс абсолютных характеристик геометрической детализации карты (плотность)

Таблица 2: p-значения теста Стьюдента на равенство средних для абсолютных характеристик детализации

|  |  |  |  |
| --- | --- | --- | --- |
| param | scale1 | scale2 | p-value |
| width | 200 | 500 | 0.0025734 |
| width | 500 | 1000 | 0.1654631 |
| height | 200 | 500 | 0.0056187 |
| height | 500 | 1000 | 0.0478674 |
| minarea | 200 | 500 | 0.0000326 |
| minarea | 500 | 1000 | 0.0001091 |

# Анализ зависимости плотностных характеристик детализации от уровня хозяйственной освоенности территории

Значительный разброс относительных характеристик геометрической детализации обусловлен тем, что территории характеризуются неодинаковой хозяйственной освоенностью и сложностью ландшафта. Если первая особенность влияет на плотность изображения социально-экономических объектов (населённые пункты, транспорт, инженерные коммуникации, промышленные сооружения и т.д.) на топографических картах, то вторая компонента обуславливает плотность изображения природных элементов (рельеф, гидрография, растительность, грунты).

Заметим, что плотность объектов на карте влияет на надёжность оценок детализации, получаемых на основе анализа абсолютных характеристик детализации: при малом количестве и разнообразии объектов оценки размеров будут несостоятельными. Так, например, значительный выброс оценки минимальной средней площади для участка Салехард в масштабе 1:1 000 000 связан с тем, что на этом участке наименее крупными полигональными социально-экономическими объектами являются муниципальные районы, имеющие огромную площадь.

В данной части исследования мы провели эксперимент направленный на выяснение зависимости между уровнем хозяйственной освоенности территории и плотностью размещения социально-экономических объектов на картах. Для определения уровня хозяйственной освоенности мы использовали независимый источник информации — растровый набор данных Copernicus Global Land Cover (<https://land.copernicus.eu/global/products/lc>), имеющий глобальное покрытие и размер ячейки 100 м. Для каждой ячейки этого растра известна доля одного из 22 типов земельного покрова, из которых два — урбанизированные территории и с/х земли — можно отнести к элементам хозяйственного освоения территории.

Карты земельного покрова для исследованных участков показаны на Рис. 5 — 13

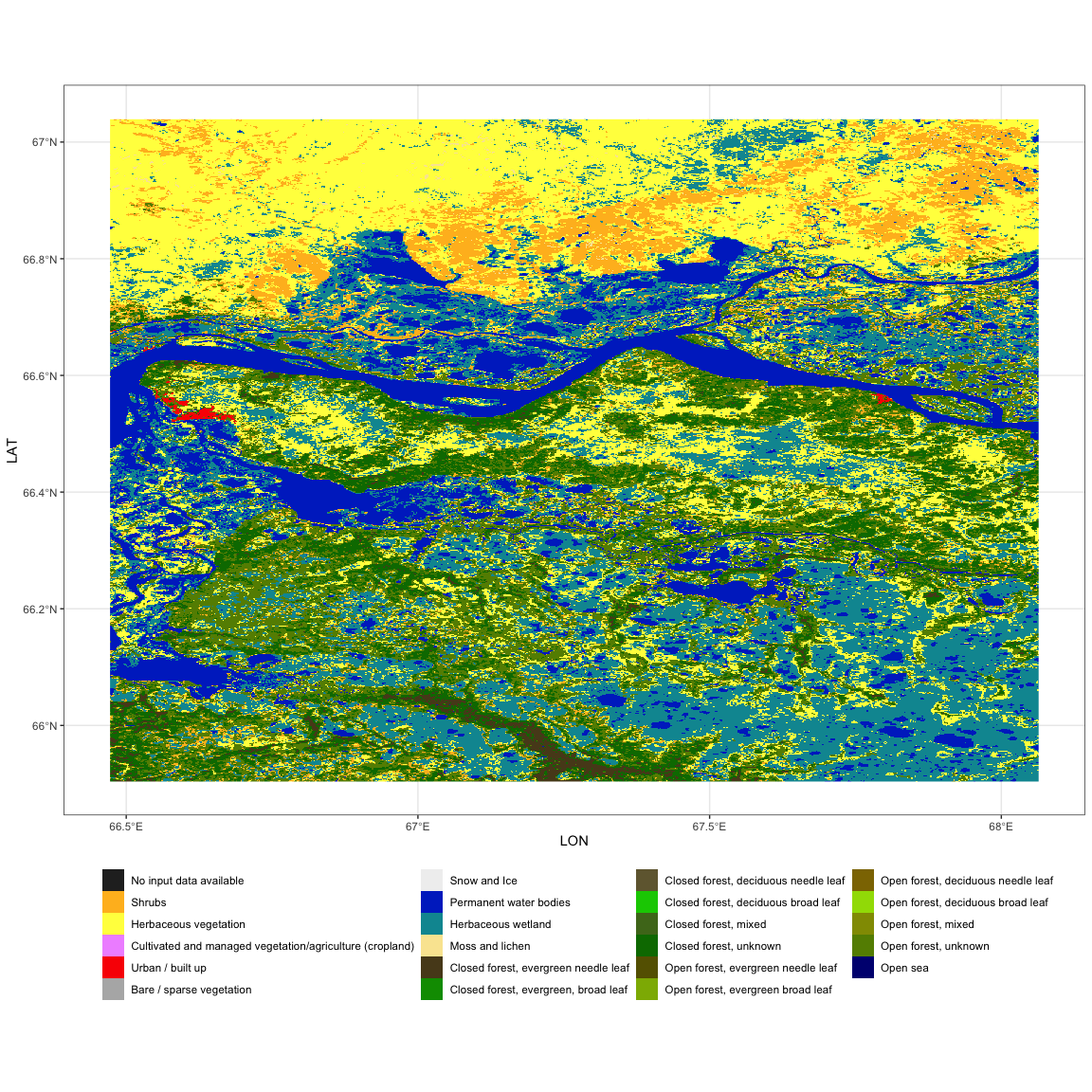


Рис. 5: Типы земельного покрова по данным Copernicus Global Land Cover (участок Салехард)

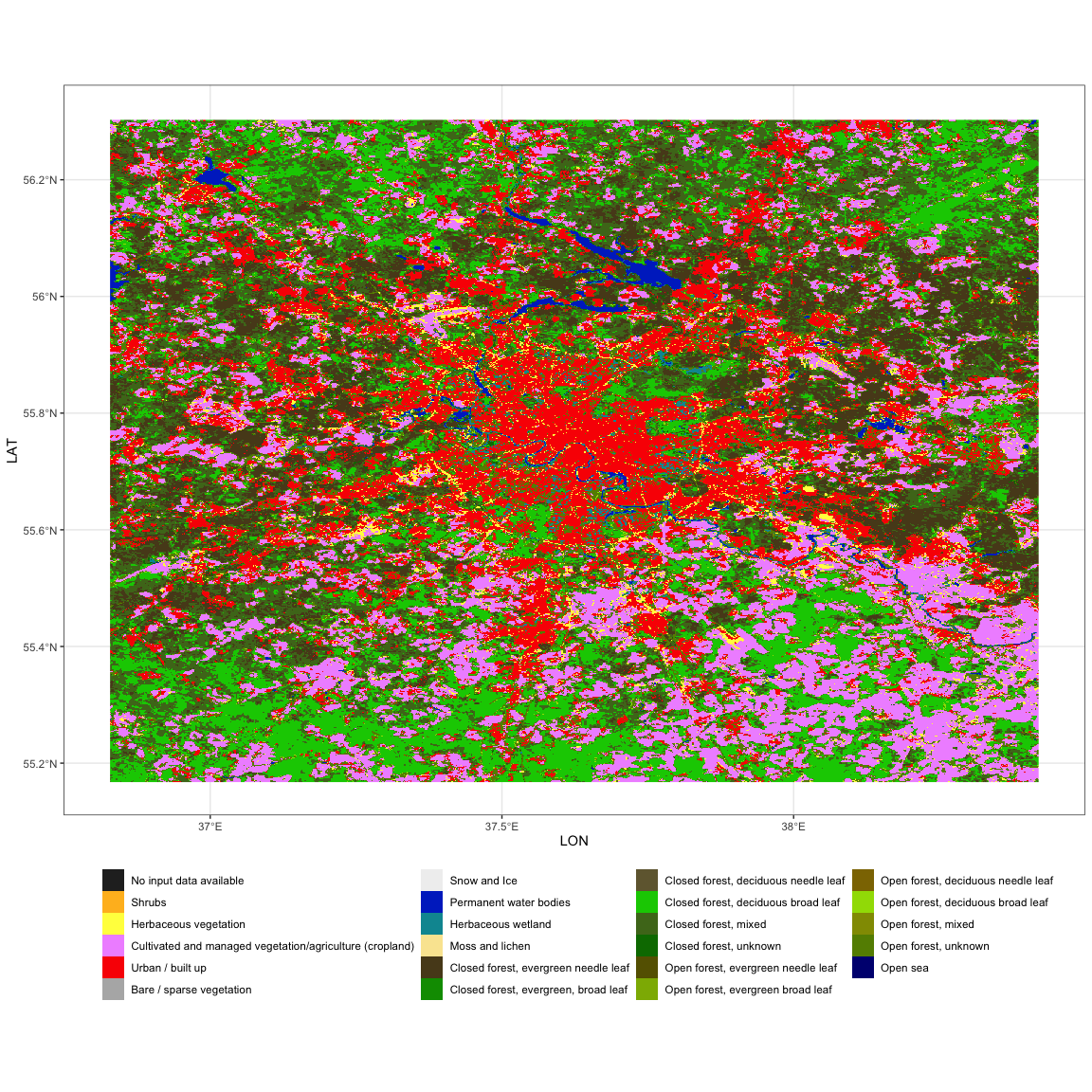


Рис. 6: Типы земельного покрова по данным Copernicus Global Land Cover (участок Москва)



Рис. 7: Типы земельного покрова по данным Copernicus Global Land Cover (участок Ростов-на-Дону)

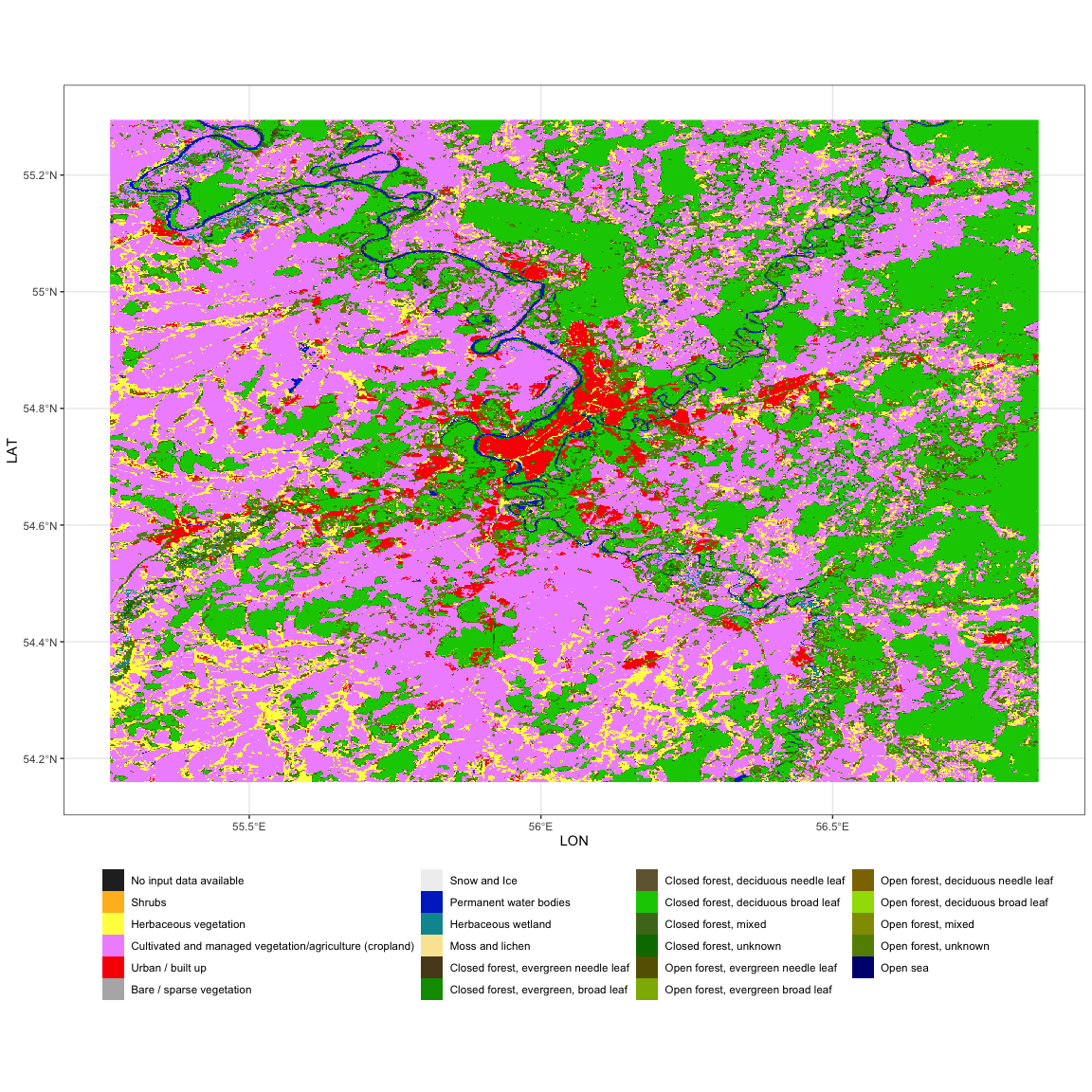


Рис. 8: Типы земельного покрова по данным Copernicus Global Land Cover (участок Уфа)

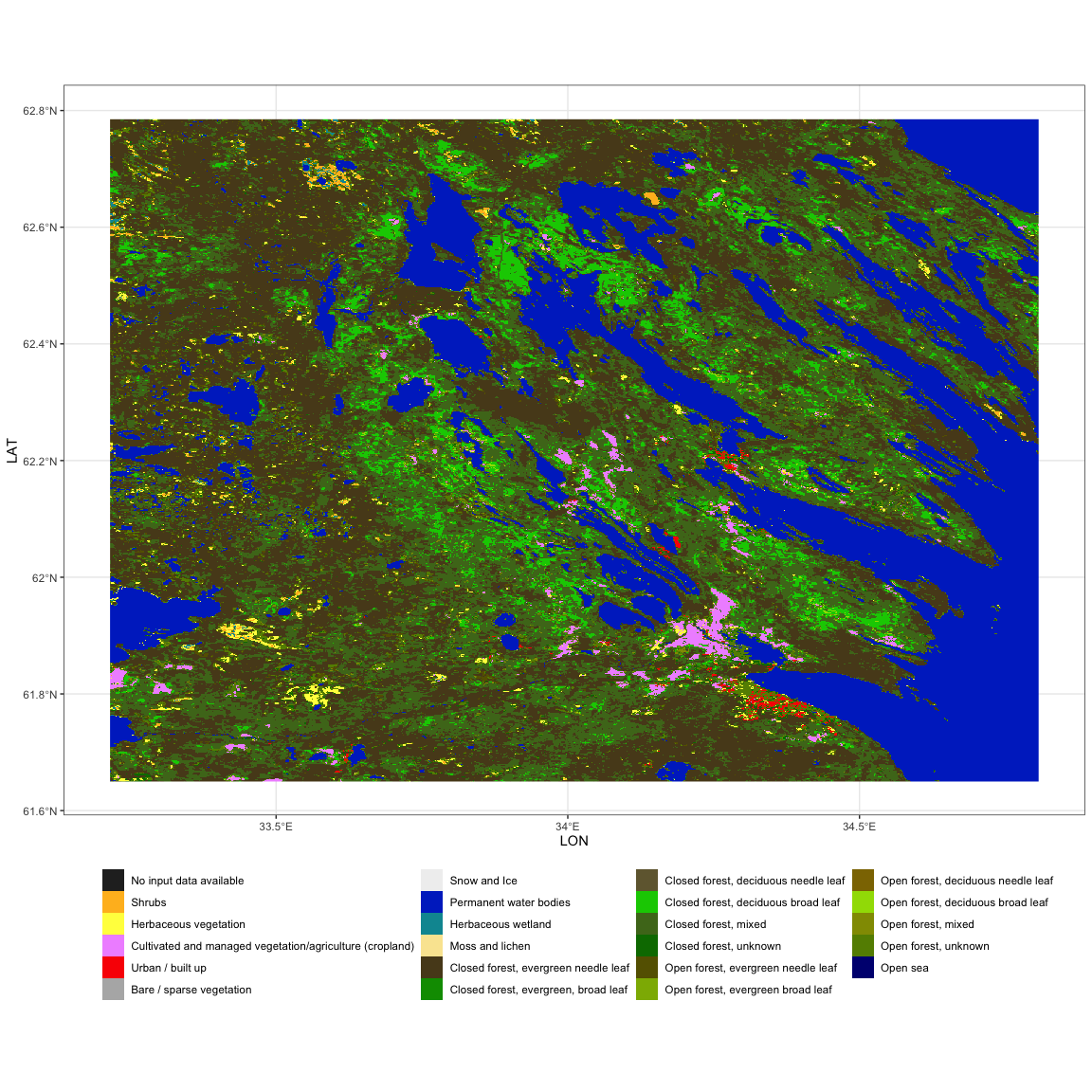


Рис. 9: Типы земельного покрова по данным Copernicus Global Land Cover (участок Петрозаводск)

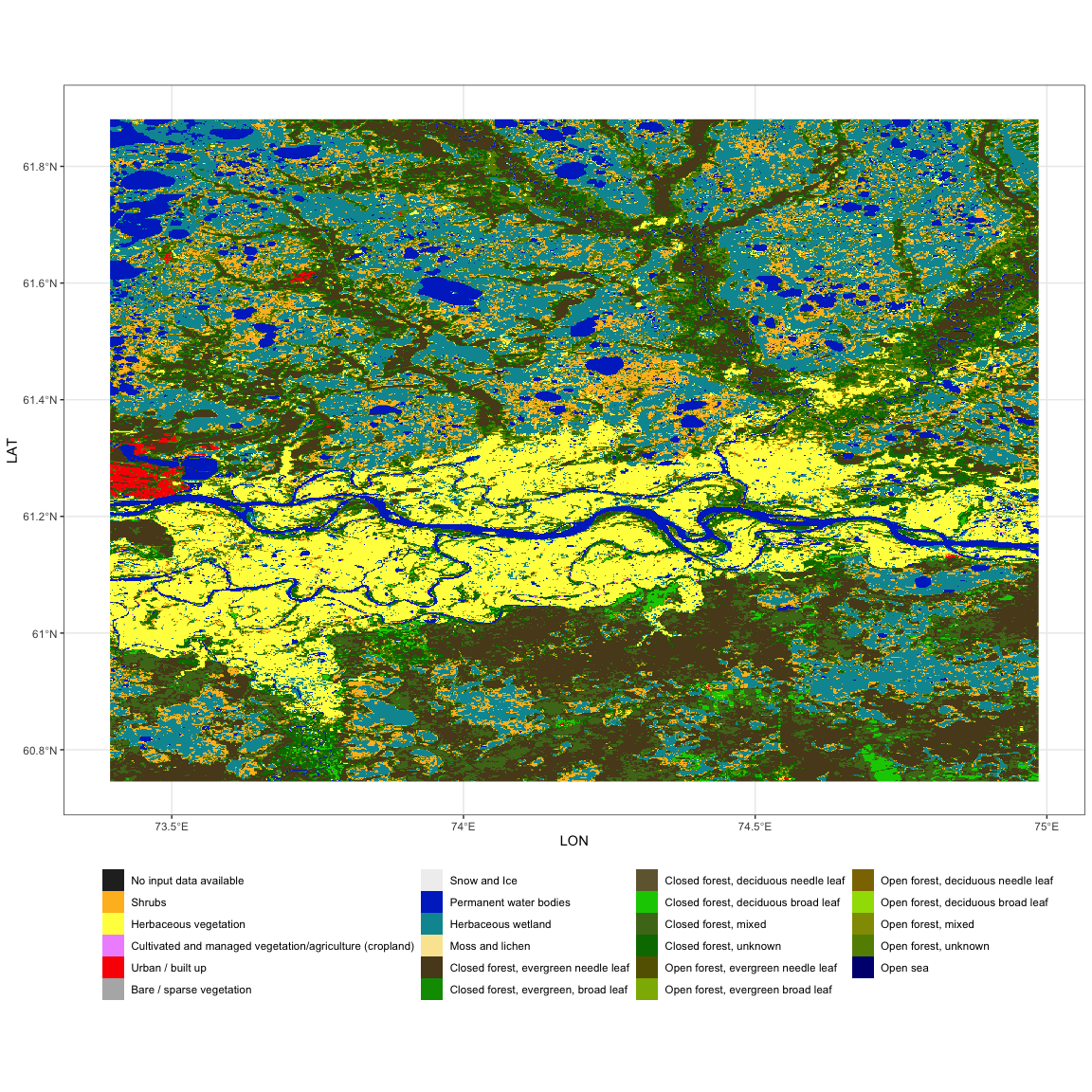


Рис. 10: Типы земельного покрова по данным Copernicus Global Land Cover (участок Сургут)

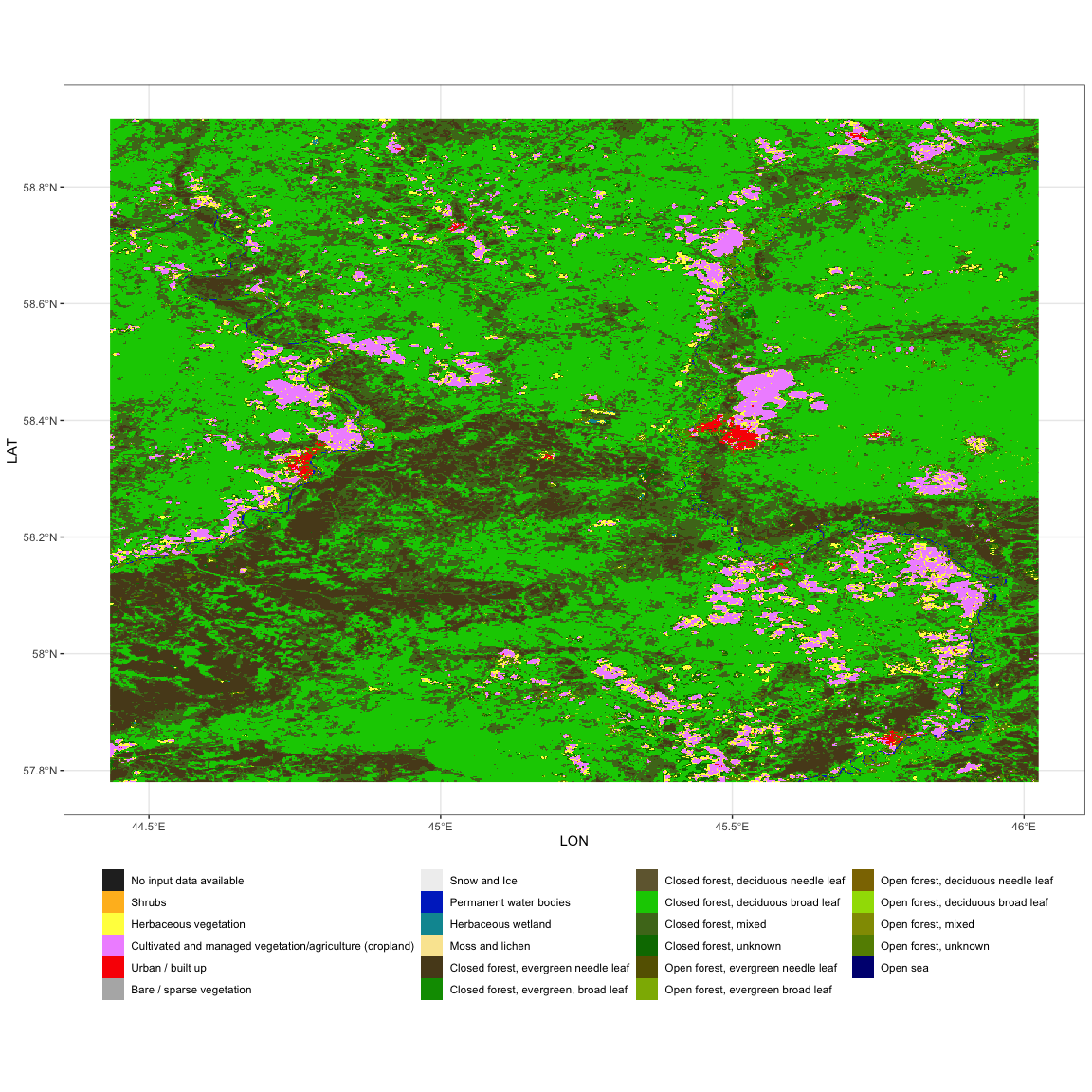


Рис. 11: Типы земельного покрова по данным Copernicus Global Land Cover (участок Шарья)

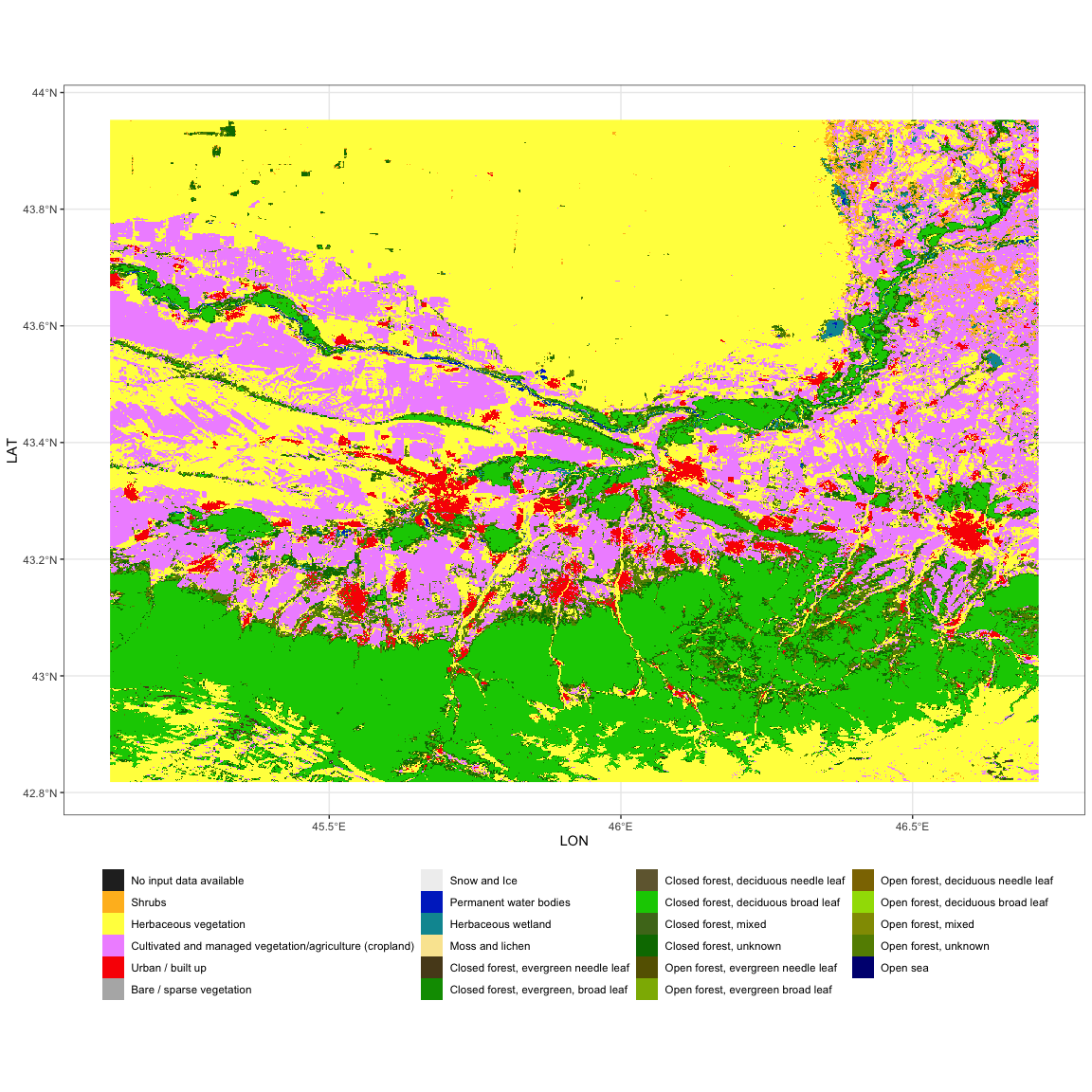


Рис. 12: Типы земельного покрова по данным Copernicus Global Land Cover (участок Грозный)

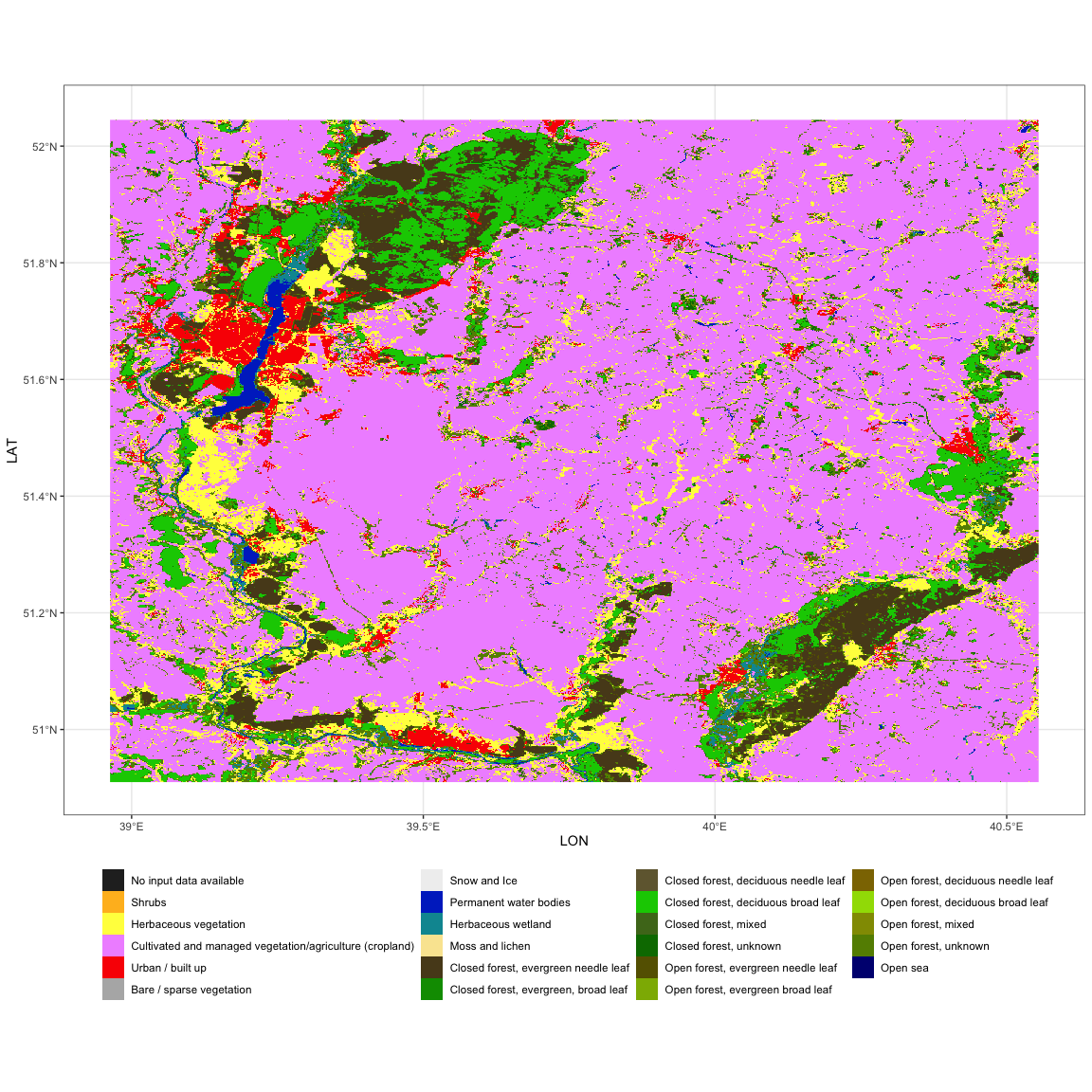


Рис. 13: Типы земельного покрова по данным Copernicus Global Land Cover (участок Воронеж)

Нами был проведен корреляционный и регрессионный анализ зависимости плотностных характеристик детализации социально-экономических элементов топографических карт от уровня хозяйственной освоенности территории. Диаграммы рассеяния между плотностью точек/объектов/пересечений и долей урбанизированных территорий/сельхозугодий представлены на Рис. 14. Дополнительно на них нанесена линия регрессии с доверительными интервалами вероятности 0.95.

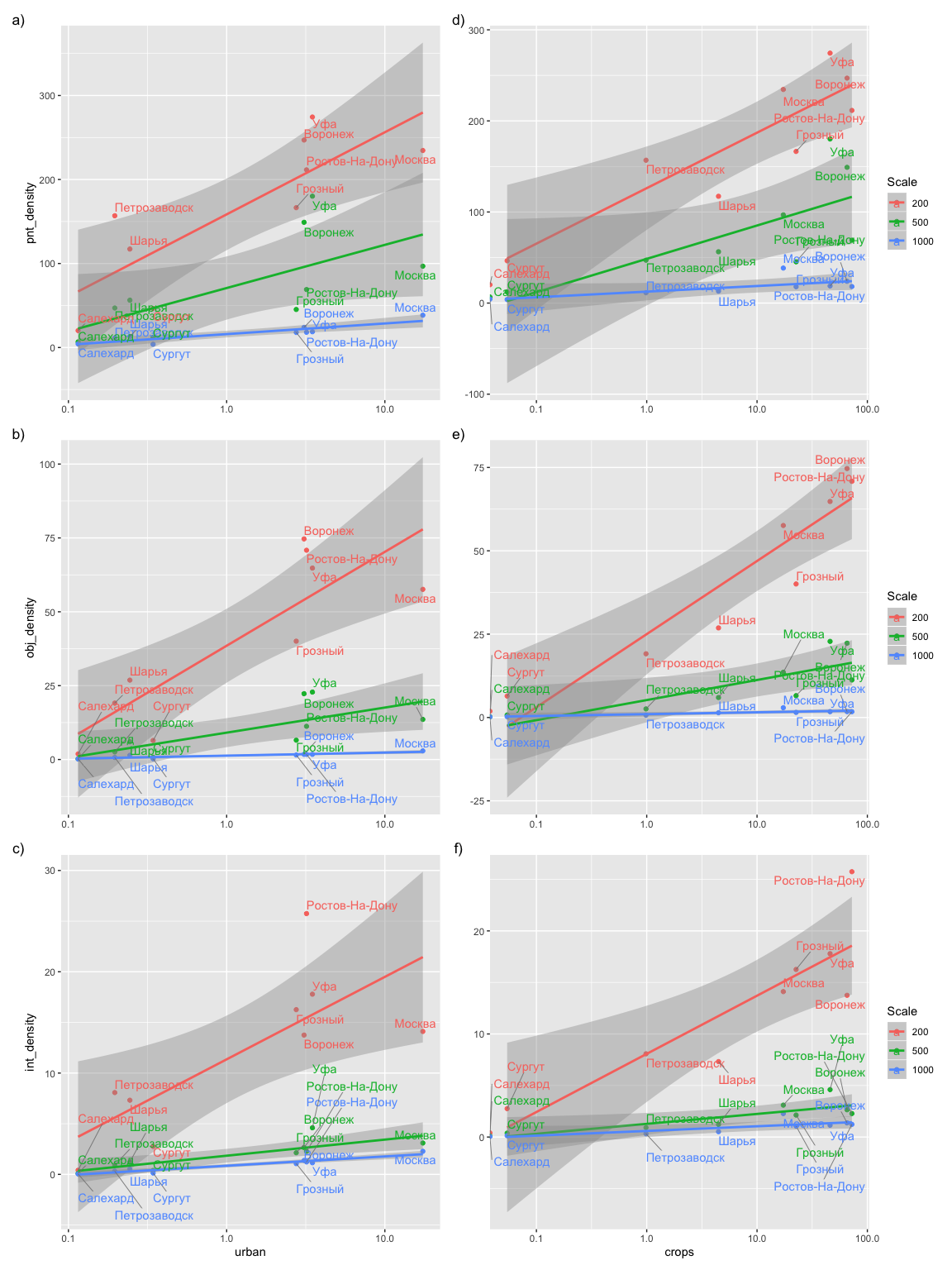


Рис. 14: Зависимость плотностных показателей детализации социально-экономических элементов карт от уровня компонент хозяйственной освоенности территории: a-c) — урбанизированные территории, d-f) — сельхозугодья

По результатам анализа была обнаружена статистически значимая (p < 0.05) зависимость между **логарифмом** процентной доли хозяйственно освоенных земель и плотностью пространственных объектов на карте. Результаты представлены в виде корреляционных матриц, сгруппированных по масштабам на Рис. 15.

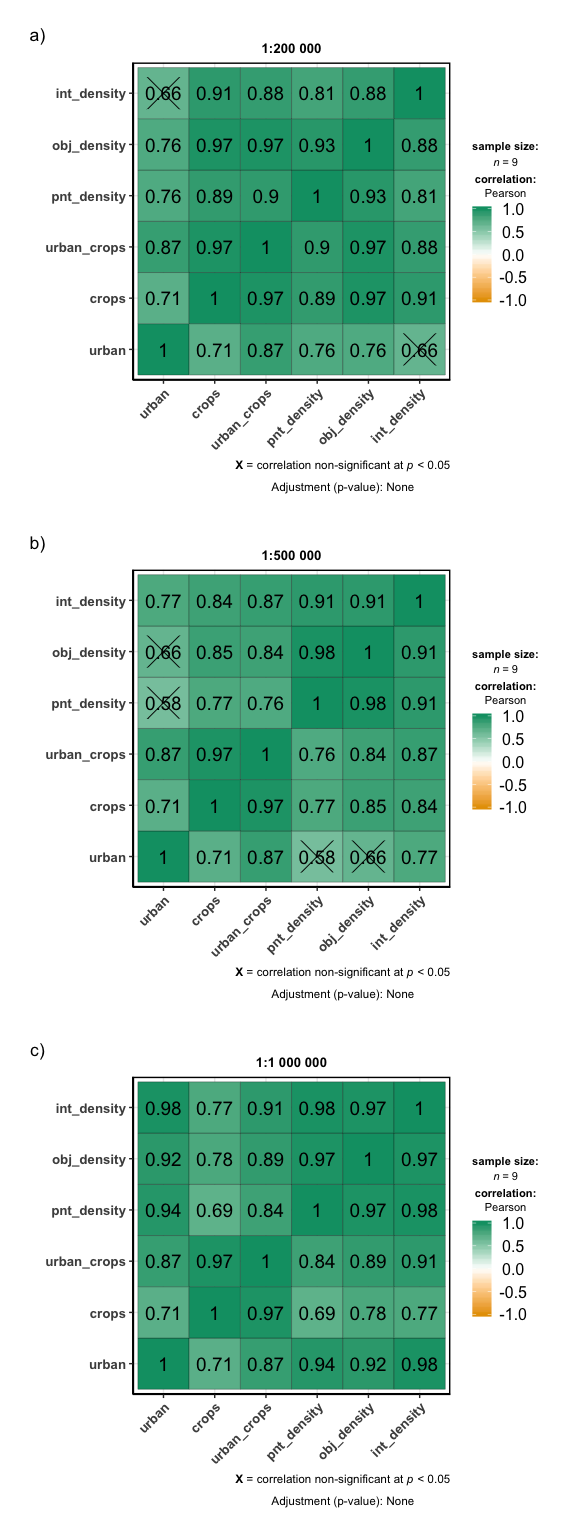


Рис. 15: Корреляционные матрицы между долей хозяйственно освоенных земель и относительными (плотностными) параметрами детализации карты

В масштабе 1:200 000 оказалась статистически не значима зависимость между плотностью пересечений и урбанизированностью, а в масштабе 1:500 000 — зависимости между плотностью точек, плотностью объектов и также урбанизированностью. Также можно обратить внимание на то, что в более крупных масштабах доля сельскохозяйственных территорий начинает играть более значимую роль в повышении плотности изображения, что может быть связано с подробным показом подъездных проселочных дорог. При этом стабильно высоким оказывается синтетический показатель urban\_crops, получаемый путем суммирования логарифма доли урбанизированных территорий и логарифма доли сельхозугодий (к обоим долям перед логарифмированием прибавляется единица, чтобы избежать отрицательных и бесконечных величин там где доля составляет менее 1%).

Построенные графики показывают что характер отношения между плотностными характеристиками и уровнем хозяйственного освоения зависит от масштаба карты. Характер кривой определяется средним значением нагрузки. Чем она выше (масштаб карты крупнее), тем сильнее уклон кривой. Поскольку нашей целью является нахождение закономерности, не зависящей от масштаба карты, для дальнейшего анализа мы перешли к значениям показателей, нормированных на среднее по всем регионам для соответствующего масштаба. Результирующий график с линией регрессии, построенной уже для данных по всем масштабам, представлен на Рис. 16, а соответствующая корреляционная матрица — на Рис. 17. Как видно, для общей выборки комплексный показатель, сочетающий долю урбанизированных территорий и долю с/х земель, обладает более высоким коэффициентом корреляции, нежели каждый тип земель по-отдельности.

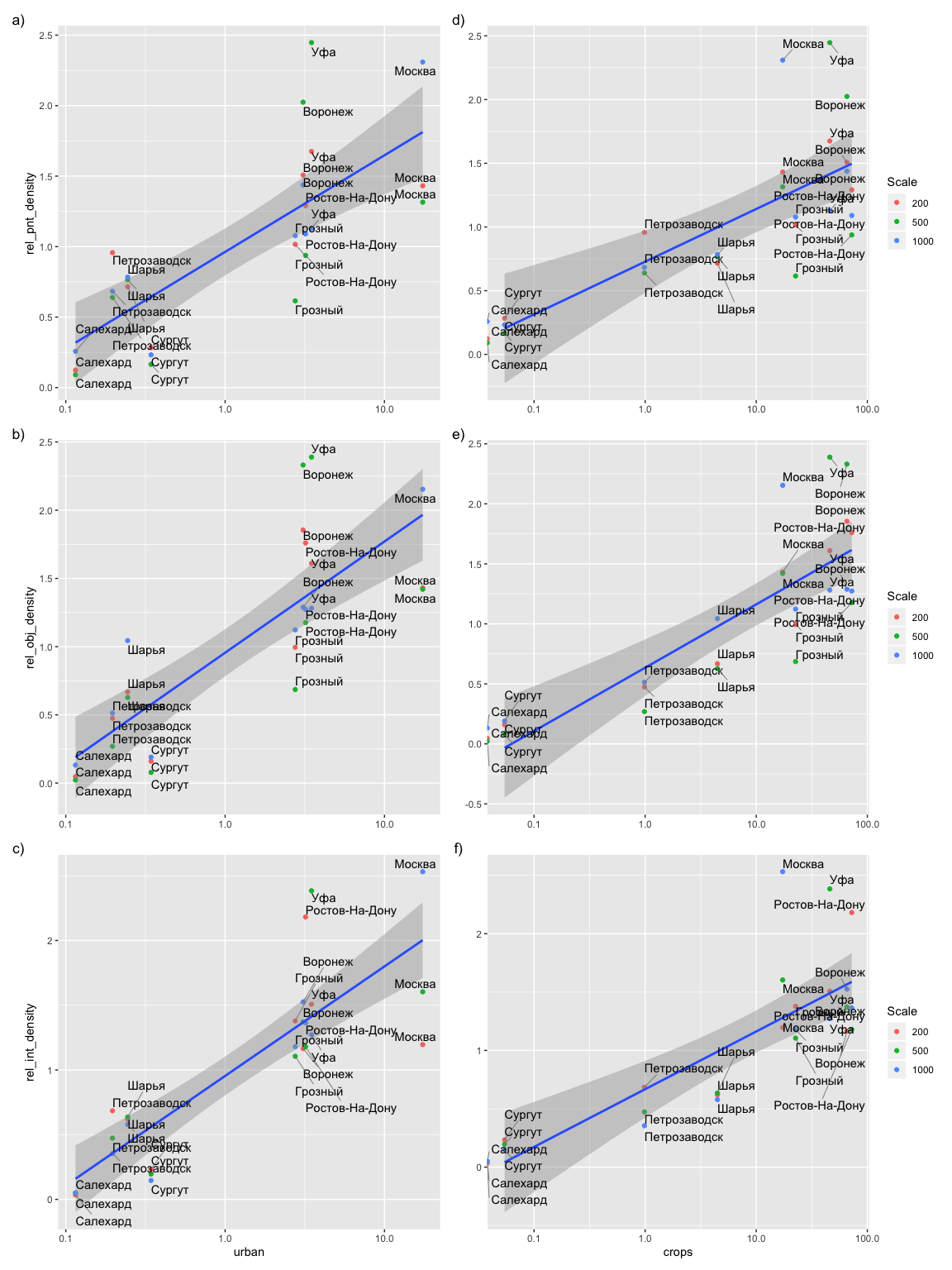


Рис. 16: Зависимость нормированных на среднее плотностных показателей детализации социально-экономических элементов карт всех масштабов от уровня компонент хозяйственной освоенности территории: a-c) — урбанизированные территории, d-f) — сельхозугодья

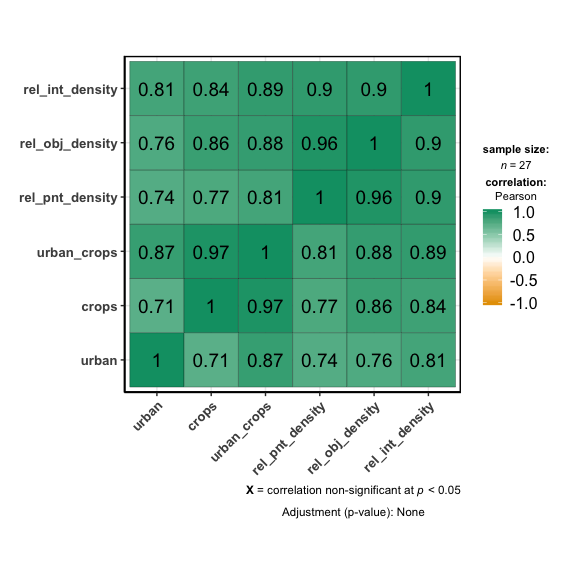


Рис. 17: Корреляционная матрица между долей хозяйственно освоенных земель и относительными (плотностными) параметрами детализации карты, нормированными на среднее (по всем масштабам)

Далее мы нашли коэффициенты линейной регрессии следующего вида:

где — величина -й нормированной на среднее по масштабу метрики детализации, — доля урбанизированных территорий, — доля сельскохозяйственных угодий. Результаты регрессионного анализа приведены в Таблице 3. Коэффициенты, полученные по данным со всех масштабов, помечены как NA.

Таблица 3: Коэффициенты линейной регрессии между логарифмом процентной доли урбанизированных территорий, логарифмом процентной доли с/х угодий и плотностными параметрами детализации карты

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| var | scale | alpha | beta | gamma |
| rel\_pnt\_density | 200 | 0.3378216 | 0.1430135 | 0.2215381 |
| rel\_pnt\_density | 500 | 0.1729355 | 0.0580641 | 0.3309348 |
| rel\_pnt\_density | 1000 | 0.3163132 | 0.6199278 | 0.0162904 |
| rel\_pnt\_density | NA | 0.2756901 | 0.2736685 | 0.1895877 |
| rel\_obj\_density | 200 | 0.0849434 | 0.0996705 | 0.3502060 |
| rel\_obj\_density | 500 | -0.0360285 | 0.1020455 | 0.4013627 |
| rel\_obj\_density | 1000 | 0.2624474 | 0.5020154 | 0.0925862 |
| rel\_obj\_density | NA | 0.1037874 | 0.2345771 | 0.2813850 |
| rel\_int\_density | 200 | 0.1839314 | 0.0071351 | 0.3490974 |
| rel\_int\_density | 500 | 0.1277016 | 0.2719571 | 0.2542461 |
| rel\_int\_density | 1000 | 0.0608824 | 0.7423555 | 0.0714907 |
| rel\_int\_density | NA | 0.1241718 | 0.3404825 | 0.2249447 |

На заключительном этапе анализа была проанализирована возможность приводить плотностные характеристики детализации к средней для масштаба величине с использованием найденных регрессий. Зная для каждой территории долю урбанизированных и сельскохозяйственных земель, получаем ожидаемое превышение плотности объектов по отношению к средней для масштаба плотности. Поделив фактическую плотность на ожидаемое превышение, получаем приведенную величину плотности. Результирующие данные представлены на Рис. 18. Построенные графики свидетельствуют о том, что выполненное приведение способствует существенному нивелированию территориальных контрастов в плотности картографических объектов, обусловленных неравномерным хозяйственным освоением территории. На представленных диаграммах уже нет случаев, когда плотность объектов отличается на 1-2 порядка, хотя отличия в несколько раз по-прежнему имеются.

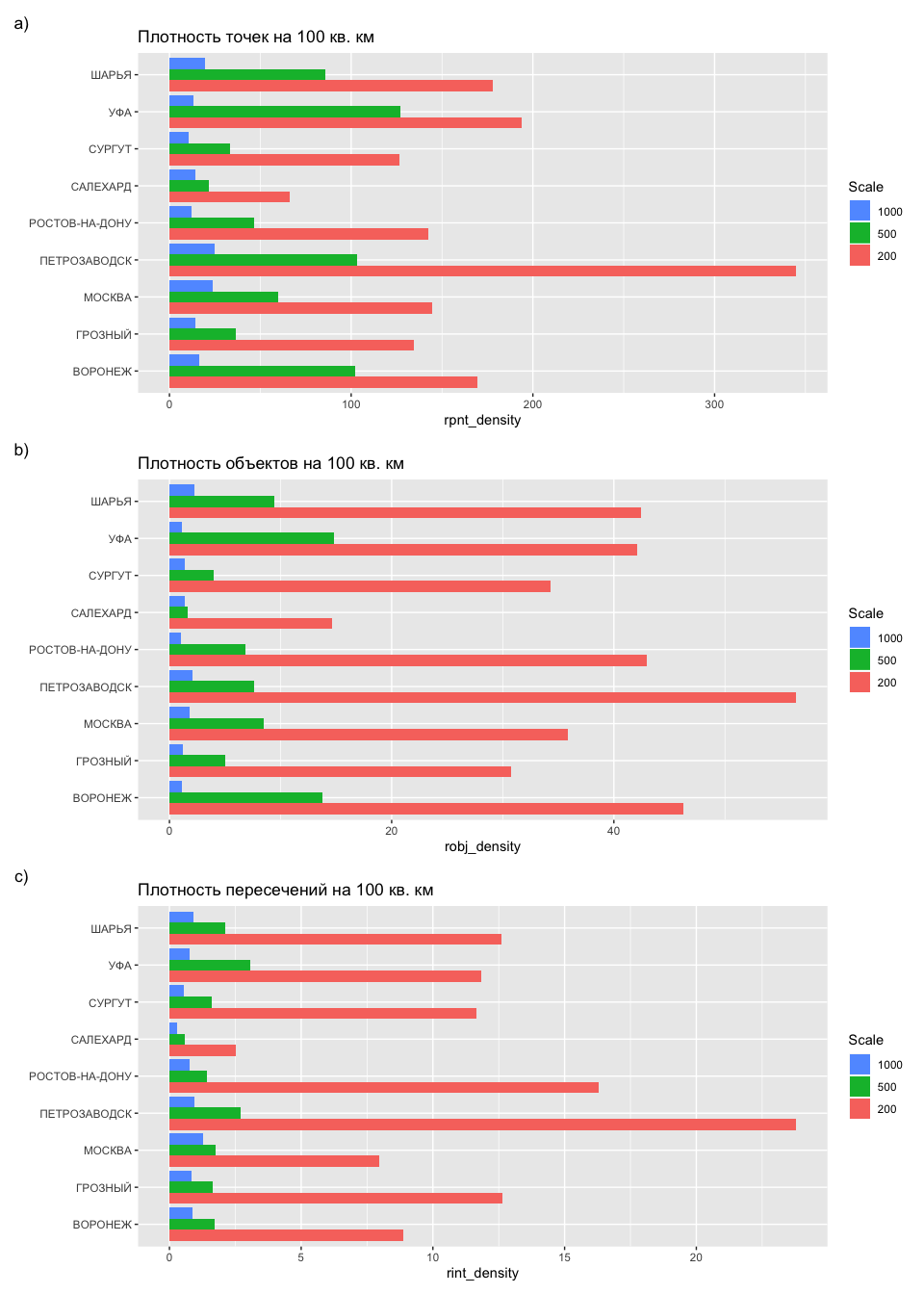


Рис. 18: Относительные характеристики геометрической детализации карты (плотность)

Диаграммы размаха (Рис. 19) демонстрируют существенно более компактный интерквартильный разброс, при этом медианные значения смещаются незначительно, что позволяет говорить о более качественной разделимости уровней детализации после выполненного преобразования данных. Это статистически подтверждается результатами теста Стьюдента на равенство средних, приведенными в Таблице 4: p-значения для всех масштабных переходов имеют величину ниже 0.005, что существенно лучше, чем аналогичые результаты, полученные для исходных, не преобразованных данных.

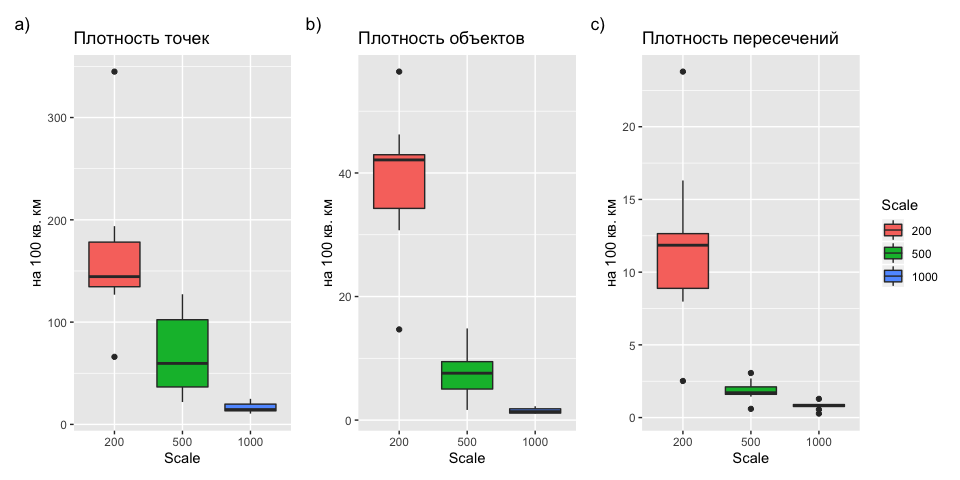


Рис. 19: Разброс относительных характеристик детализации карты (плотность)

Таблица 4: p-значения теста Стьюдента на равенство средних для нормированных относительных характеристик детализации

|  |  |  |  |
| --- | --- | --- | --- |
| param | scale1 | scale2 | p-value |
| pnt\_density | 200 | 500 | 0.0047953 |
| pnt\_density | 500 | 1000 | 0.0030152 |
| obj\_density | 200 | 500 | 0.0000220 |
| obj\_density | 500 | 1000 | 0.0019757 |
| int\_density | 200 | 500 | 0.0007702 |
| int\_density | 500 | 1000 | 0.0020417 |

Таким образом, можно утверждать, что разработанный нами подход позволяет учесть неравномерность хозяйственного освоения территории при оценке детализации картографических данных. Тем не менее, остающиеся неравномерности в детализации требуют дальнейших поисков предикторов, в том числе в части сложности природной компоненты ландшафта.

# Выводы

На основе проведенных исследований были сделаны следующие выводы:

1. Характеристики геометрической детализации карты целесообразно разделить на относительные, характеризующие плотность размещения объектов, и абсолютные, характеризующие их размеры. Эти два вида характеристик являются объектами цензов и норм отбора объектов на картах, но также имеют географическую обусловленность.
2. Относительные характеристики детализации менее надежно характеризуют уровень детализации данных, поскольку зависят от степени хозяйственной освоенности территории и сложности природного ландшафта. Чем более освоена территория и сложнее ландшафт, тем выше будет плотность элементов карты. Первый фактор влияет на плотность представления социально-экономических объектов, второй фактор — на плотность представления природных объектов на картах.
3. Надежность абсолютных характеристик (размеров объектов) зависит от сложности картографируемой территории: при невысокой сложности плотность объектов низка и выборка не позволяет получить состоятельные оценки средних и минимальных размеров объектов. При этом, однако, абсолютные характеристики не имеют явной зависимости от сложности картографируемой территории.
4. Разделение карт по детализации на основе относительных (плотностных) характеристик не может быть надежно осуществлено без привлечения независимой (внешней) информации о сложности территории в социально-экономическом и природном отношении. Для этих целей могут быть использованы геоинформационные продукты, полученные на основе данных дистанционного зондирования, такие как растры типов земельного покрова.
5. Для компенсации влияния сложности территории на относительные характеристики детализации может быть использован переход к приведенным значениям плотности размещения объектов путем их нормирования на среднюю для масштаба плотность. Приведенная плотность показывает статистически значимую корреляционную связь с логарифмированными значениями доли урбанизированных и сельскохозяйственных земель, и может быть смоделирована на их основе путем линейной регрессии. Таким образом, подобрав единожды модель регрессии между этими величинами и зная долю урбанизированных и сельскохозяйственных земель для участка, покрытого данными, можно предсказать, насколько покрывающие территорию данные должны отличаться по плотности размещения объектов от среднестатистических для данного масштаба. Применение данного подхода позволяет качественно улучшить разделимость уровней детализации на основе плотностных характеристик: дисперсии величин и p-значения теста Стьюдента на отличия в их средних снижаются в несколько раз.