



Modellierung der Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten

Masterthesis

Zur Erlangung des Mastergrades
„Master of Science“, abgekürzt „MSc“

im Rahmen des
Universitätslehrganges „Geographical Information Science & Systems“
(UNIGIS MSc) am Fachbereich Geoinformatik (Z_GIS),
Fakultät für Digitale und Analytische Wissenschaften,
Universität Salzburg

eingereicht von

Philipp Blumenkamp

Gutachter:
Assoz. Prof. Dr. Dirk Tiede

Krefeld, Juni 2025

Danksagung

An dieser Stelle möchte ich meinen Dank an die Personen aussprechen, die mich bei der Erstellung dieser Masterarbeit unterstützt haben.

Mein besonderer Dank gilt Herrn Felix Kröber für die fachliche Betreuung, die Unterstützung bei der Erarbeitung des Themas und die stets sehr konstruktiven Gespräche bei Fragen und Problemen. Ebenso danke ich Herrn Prof. Dr. Dirk Tiede für die Betreuung sowohl bei fachlichen als auch administrativen Fragen.

Ein weiter Dank gilt dem gesamten UNIGIS-Team. Besonders hervorheben möchte ich hier Frau Prof. Dr. Gudrun Wallentin für die sehr strukturierte Vermittlung wissenschaftlicher Arbeitsweisen und die hervorragend aufbereiteten Materialien, die eine ideale Vorbereitung zur Bearbeitung dieser Masterarbeit boten.

Zusammenfassung

Die Verbesserung der Energieeffizienz von Gebäuden ist von zentraler Bedeutung im Kontext der Energiewende und des Klimawandels. Energieausweise informieren über die Energieeffizienz einzelner Gebäude und ermöglichen durch statistische Auswertungen breite Anwendungsmöglichkeiten. Die statistische Auswertung der Daten ist derzeit beschränkt, da die Datenqualität oft nicht gewährleistet ist, die Daten häufig in Datensilos vorliegen und die Erstellung von Energieausweisen nur unter bestimmten Bedingungen verpflichtend ist.

Diese Arbeit untersucht die Modellierung der Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten mit einer Bodenauflösung von 3,5 Metern, die umfangreich aufbereitet und in Landoberflächentemperaturen umgerechnet wurden. Die Thesen basiert auf der Hypothese, dass die Abwärme von Gebäuden in den Wärmebildern detektierbar ist. Für die Modellierung wurden Daten von ca. 43.000 Energieausweisen verwendet. Die Modellierung mit Regressionsmethoden zeigte jedoch keinen ausreichenden Zusammenhang zwischen der Landoberflächentemperatur und der Energieeffizienz von Gebäuden. Eine Hauptursache dafür ist der Aufnahmezeitpunkt der Thermaldaten, bei dem die Lufttemperaturen zu hoch waren, um die Abwärme der Gebäude zuverlässig zu messen. Zusätzlich fehlen wichtige Informationen zum Dachmaterial der Gebäude und den Oberflächentemperaturen der Gebäudefassaden. Die vorherrschende Bebauungsstruktur im Untersuchungsgebiet – geprägt durch kleine Reihenhäuser – stellt eine weitere Herausforderung im Hinblick auf die Bodenauflösung der Thermaldaten dar.

Dennoch implizieren alle berechneten Regressionen einen negativen Trend, sodass ein energieineffizientes Gebäude durchschnittlich höhere Landoberflächentemperaturen aufweist. Diese Ergebnisse bestätigen die zugrundeliegende Annahme, dass die Detektion der Abwärme von Gebäuden bei der vorliegenden Bodenauflösung möglich ist. Die Aggregation der Daten auf größere räumliche Ebenen konnte die Modellgüte verbessern, wobei auch ein multiskaliger Ansatz keinen hinreichend starken statistischen Zusammenhang erkennen lässt. Die in den Residuen erkennbaren räumlichen Muster konnten durch die eingesetzten räumlichen Regressionsansätze auf Basis der verfügbaren Daten nicht zuverlässig abgebildet werden.

Abstract

Improving the energy efficiency of buildings is a key priority in the context of the energy transition and climate change. Energy performance certificates provide information on the energy efficiency of buildings and enable statistical analyses that support energy planning and strategic decision-making. However, the statistical use of the data is currently limited, as the data quality is not guaranteed, fragmented data storage and the fact that certificate creation is only mandatory under certain conditions.

This thesis investigates the modelling of the energy efficiency of buildings using thermal remote sensing data with a ground resolution of 3.5 m, which has been extensively processed and converted into land surface temperatures. The thesis is based on the hypothesis that the waste heat of buildings can be detected in thermal imagery. Data from approximately 43,000 energy performance certificates were used for modelling. The modelling with regression methods could not prove a sufficient strong correlation between the land surface temperature and the energy efficiency of buildings. One of the main reasons for this is the time at which the thermal data was recorded, when the air temperatures were too high to reliably measure the waste heat from the buildings. In addition, important information such as the roof material of the buildings and the surface temperatures of the building facades is missing. The predominant building structure in the study area - characterised by small terraced houses - poses a further challenge with regard to the ground resolution of the thermal data.

Nevertheless, all calculated regressions indicate a negative trend, meaning that an energy-inefficient building has higher average land surface temperatures. These results confirm the underlying assumption that the detection of waste heat from buildings is possible at the available ground resolution. Aggregating the data to larger spatial levels improved the quality of the model, although even a multiscale approach does not reveal a sufficiently strong statistical correlation. The spatial patterns recognisable in the residuals could not be reliably mapped by the spatial regression approaches used on the basis of the available data.

Inhaltsverzeichnis

Danksagung	I
Zusammenfassung	II
Abstract	III
Inhaltsverzeichnis	IV
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
Abkürzungsverzeichnis	VIII
1 Einleitung	1
1.1 Problemstellung	1
1.2 Stand der Forschung	1
1.3 Forschungsfrage	2
1.4 Aufbau der Arbeit	2
2 Daten und Methoden	3
2.1 Daten	3
2.1.1 Untersuchungsgebiet	3
2.1.2 Thermaldaten	4
2.1.3 Klimatische Bedingungen	4
2.1.4 Energy Performance Certificates	5
2.1.5 Gebäudepolygone	6
2.2 Vorprozessierung der Daten	7
2.2.1 Zeitlicher Trend	7
2.2.2 Erstellung der Mosaik	11
2.2.3 Berechnung der mittleren Brightness Temperature je Gebäude	13
2.2.4 Berechnung der Land Surface Temperatures	18
2.3 Deskriptive Statistiken zu den Daten	24
2.3.1 Deskriptive Statistik zu den Energy Performance Certificates	24
2.3.2 Deskriptive Statistik zu den Land Surface Temperatures der Gebäudedächer	27
2.4 Methoden zur Modellierung der Energieeffizienz	30
2.4.1 Modellierung mittels globaler Regressionsansätze	30
2.4.2 Betrachtung verschiedener räumlicher Ebenen	32
2.4.3 Explizite Berücksichtigung räumlicher Effekte	34
3 Ergebnisse	36
3.1 Modellierung mittels globaler Regressionsansätze	36
3.2 Betrachtung verschiedener räumlicher Ebenen	40
3.3 Explizite Berücksichtigung räumlicher Effekte	42
4 Diskussion	43

5	Schlussfolgerungen und Ausblick.....	47
	Literaturverzeichnis	49
	Anhang.....	53
	Anhang A: Python-Skript zur Berechnung des zeitlichen Trends und zur Normalisierung der Daten.....	53
	Anhang B: Python-Skript zur Berechnung der gewichteten Mittelwerte für die Brightness Temperatures	55
	Anhang C: Python-Skript zur Erstellung der Diagramme für die deskriptiven Statistiken ...	56
	Anhang D: Python-Skript zur Modellierung der Energieeffizienz mittels globaler Regressionsansätze und als multiskaliger Ansatz.....	58
	Anhang E: Python-Skript zur Modellierung der Energieeffizienz mittels räumlich expliziter Regressionsansätze	60

Abbildungsverzeichnis

Abbildung 1: Untersuchungsgebiet im County Merseyside	3
Abbildung 2: Visualisierung des zeitlichen Trends am 12.11.2022 auf einer Karte	8
Abbildung 3: Gewichtete Regressionsanalyse zwischen Brightness Temperature und Time am 12.11.2022	9
Abbildung 4: Gewichtete Regressionsanalyse zwischen Brightness Temperature und Time am 13.11.2022	9
Abbildung 5: Beispiele für Artefakte.....	11
Abbildung 6: Mosaik der Thermalbilder vom 12.11.2022	12
Abbildung 7: Mosaik der Thermalbilder vom 13.11.2022	13
Abbildung 8: Beispiel zur Selektion der Rasterzellen für die BTs anhand der Mittelpunkte ...	15
Abbildung 9: Beispiel zur Selektion der Rasterzellen für die BTs anhand der Flächenanteile	16
Abbildung 10: Schematische Darstellung der Geoprozessierung zur Berechnung der mittleren BTs für die Gebäude.....	17
Abbildung 11: Beispielansicht zur LST der Gebäude gemeinsam mit der jeweiligen Energieeffizienzklasse.....	23
Abbildung 12: Anzahl der Objekte je Energieeffizienzklasse der Gebäude	24
Abbildung 13: Histogramm für die Energieeffizienz der Gebäude	25
Abbildung 14: Anzahl der Objekte je Energieeffizienzklasse der Gebäudedächer	26
Abbildung 15: Verteilung der Gebäudeflächen in Kategorien.....	27
Abbildung 16: Histogramm für die LSTs am 12.11.2022.....	29
Abbildung 17: Histogramm für die LSTs am 13.11.2022.....	29
Abbildung 18: Histogramm für die Mittelwerte der LSTs beider thermischer Infrarotmessungen	30
Abbildung 19: Beispielansicht zu den verschiedenen räumlichen Ebenen.....	33
Abbildung 20: Vergleich mehrerer linearer Regressionen zwischen verschiedenen LST-Werten und den Energieeffizienzwerten der Gebäude bzw. Dächer.....	37
Abbildung 21: Vergleich mehrerer linearer Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Filterung der Gebäude nach deren Flächen	39
Abbildung 22: Vergleich mehrerer linearer Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Aggregation der Werte auf verschiedene räumliche Ebenen	41

Tabellenverzeichnis

Tabelle 1: Wetterdaten der Beobachtungszeiträume	5
Tabelle 2: Regressionsparameter für die Berechnung des zeitlichen Trends.....	10
Tabelle 3: Emissionsgrade für Dachmaterialien	22
Tabelle 4: Deskriptive statistische Maße der LST-Werte	28
Tabelle 5: Parameter der linearen Regressionen zwischen verschiedenen LST-Werten und den Energieeffizienzwerten der Gebäude bzw. Dächer	36
Tabelle 6: Parameter der linearen Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Filterung der Gebäude nach deren Flächen	38
Tabelle 7: Parameter der linearen Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Aggregation der Werte auf verschiedene räumliche Ebenen	40
Tabelle 8: Parameter der räumlichen Regressionsmodelle Spatial Lag und Spatial Error zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer ..	42

Abkürzungsverzeichnis

API	Application Programming Interface
BT	Brightness Temperature
EPC	Energy Performance Certificate
LST	Land Surface Temperature
NDVI	Normalized Difference Vegetation Index
TES	Temperature and Emissivity Separation
TISI	Temperature Independent Spectral Indices
UPRN	Unique Property Reference Number

1 Einleitung

1.1 Problemstellung

Die Verbesserung der Energieeffizienz von Gebäuden ist von zentraler Bedeutung im Kontext der Energiewende und des Klimawandels. Nach Schätzungen der Europäischen Kommission sind Gebäude für 40 % des Energieverbrauchs sowie für 36 % der energiebedingten Treibhausgasemissionen in der Europäischen Union verantwortlich (European Commission, 2020). Energieausweise informieren über die Energieeffizienz einzelner Gebäude und tragen zur Erreichung nationaler Energieeffizienzziele bei, indem sie Potenziale zur Energieeinsparung aufzeigen (European Parliament, 2010). Innerhalb der Europäischen Union wurden ca. 75 % des Gebäudebestands mit einer schlechten Energieeffizienz klassifiziert (European Commission, 2020).

Neben der Information über die Energieleistung von Gebäuden eröffnen Energieausweise durch statistische Analysen breitere Anwendungsmöglichkeiten als ursprünglich beabsichtigt, insbesondere in der Energieplanung und strategischen Entscheidungsfindung (López-González et al., 2016). Allerdings ist die statistische Auswertung von Energieausweisen beschränkt, da die Datenqualität oft nicht gewährleistet ist und die Daten häufig in Datensilos vorliegen (Pasichnyi et al., 2019). Zudem decken die Daten nicht den gesamten Gebäudebestand ab, da die Ausstellung von Energieausweisen nur in bestimmten Situationen verpflichtend ist, wie beispielsweise beim Bau, Verkauf oder der Vermietung von Gebäuden oder Gebäudeteilen (European Parliament, 2010, Art. 12).

1.2 Stand der Forschung

Die Qualität von Energieausweisdaten kann mit zusätzlichen Daten gesichert werden (Pasichnyi et al., 2019). Die Energieleistung von Gebäuden wurde bereits mit datengetriebenen Methoden modelliert. Hierzu gehören die Anwendung von Deep-Learning Methoden unter Verwendung einer Vielzahl von Datensätzen (Ali et al., 2020) sowie Schätzungen basierend auf den Informationen, die von Immobilieneigentümern gesammelt wurden (Wederhake et al., 2022).

Die Infrarot-Thermografie ermöglicht die Untersuchung der gebauten Umwelt auf verschiedenen räumlichen Ebenen (Martin et al., 2022). Thermische Fernerkundungsdaten wurden in zahlreichen Publikationen zur Detektion von städtischen Hitzeinseln sowie deren Indikatoren verwendet (Parlow, 2021; Zhao et al., 2018; Zhu et al., 2019). Wärmebilder von Drohnen wurden effektiv eingesetzt, um die Energieeffizienz einzelner Bauwerke zu bewerten und Wärmeverluste in der Gebäudehülle zu inspizieren (Z. Mayer et al., 2023). Zwischen der aus thermischen Satellitendaten abgeleiteten Bodenoberflächentemperatur und dem Energieverbrauch je Fläche wurde eine starke Korrelation auf Blockebene festgestellt (Jazizadeh & Taleghani, 2016). Die Modellierung der Energieeffizienz von Gebäuden mit

luftgestützten Wärmebildern und wenigen Energieausweisdaten innerhalb eines Baublocks mit heterogenem Gebäudebestand erzielte gute Ergebnisse (Anselmo et al., 2023).

1.3 Forschungsfrage

Das Ziel dieser Masterarbeit ist die Modellierung der Energieeffizienz von Gebäuden mit hochauflösenden thermischen Fernerkundungsdaten. Aus diesem Ziel ergeben sich folgende zu beantwortende Teilfragen:

1. Mit welcher Genauigkeit kann die Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten mit einer Bodenauflösung von 3,5 m statistisch modelliert werden?
2. Wie wirkt sich die Betrachtung verschiedener räumlicher Ebenen auf das Modell aus?
3. Wie wirkt sich die explizite Berücksichtigung räumlicher Effekte im Modell aus?

1.4 Aufbau der Arbeit

Die Arbeit ist wie folgt aufgebaut: Zunächst wird die verwendete Datengrundlage strukturiert dargelegt. Die thermischen Fernerkundungsdaten wurden umfangreich vorprozessiert, indem zeitliche Trends eliminiert, Artefakte in den Mosaiken minimiert und die gemessenen Werte auf Einzelgebäude bezogen sowie in Landoberflächentemperaturen umgerechnet wurden. Mittels deskriptiver statistischer Methoden und explorativer Datenanalyse werden die Verteilungen der aufbereiteten Daten veranschaulicht. Anschließend werden die methodischen Ansätze zur Bearbeitung der drei Teilfragen beschrieben. Kapitel 3 präsentiert die Ergebnisse der unterschiedlichen Modellierungen der Energieeffizienz, die im darauffolgenden Kapitel 4 kritisch diskutiert sowie mit den Ergebnissen anderer Publikationen verglichen werden und damit Antworten auf die zuvor beschriebene Forschungsfrage geben. Das abschließende Kapitel 5 fasst die Ergebnisse der Arbeit zusammen und zeigt offene Fragestellungen, die in weiterführenden Arbeiten betrachtet werden können.

2 Daten und Methoden

2.1 Daten

2.1.1 Untersuchungsgebiet

Das in Abbildung 1 dargestellte Untersuchungsgebiet befindet sich im County Merseyside im Vereinigten Königreich und erstreckt sich über beide Seiten des Flusses Mersey.



Abbildung 1: Untersuchungsgebiet im County Merseyside

Durch seine Lage und Struktur handelt es sich um ein überwiegend urban geprägtes Gebiet mit dichter Bebauung. Der westliche Teil liegt im Metropolitan Borough of Wirral. Dieser Bereich ist von Wohngebieten dominiert und umfasst daneben auch größere Grün- und Gewerbeflächen. Entlang des Flusses Mersey befindet sich eine große Hafenanlage mit einem angrenzenden Gewerbegebiet. Der östliche Abschnitt des Untersuchungsgebietes umfasst

große Teile der Stadt Liverpool. Dieser Bereich ist stark urbanisiert mit einer dichten Wohnbebauung sowie Geschäftsgebieten. Zusätzlich befinden sich in diesem Bereich mehrere Parkanlagen sowie Gewerbegebiete. Die Gesamtfläche des Untersuchungsgebietes beträgt etwa 83,9 km².

2.1.2 Thermaldaten

Die Thermaldaten wurden von dem Satellitenunternehmen SatelliteVU mittels thermischer Infrarotmessung aus einem Flugzeug aufgenommen. Die Daten wurden mit einer Bodenauflösung von 3,5 Metern aufbereitet, entsprechend der Bildqualität der geplanten HotSat-Satelliten, die hochauflösende thermische Infrarotbilder liefern sollen. Der verwendete Thermalscanner arbeitet im mittleren Infrarotbereich mit einer Wellenlänge von 3,7 µm bis 5 µm. Die Aufnahmen erfolgen aus einer nahezu senkrechten Blickrichtung mit einem Off-Nadir Winkel von 3,8°. Dieser Winkel ermöglicht lediglich die Beobachtung der Dächer von Gebäuden und nicht deren Fassaden. Eine geometrische Korrektur der Rasterdaten wurde durch den Datenanbieter bereits ausgeführt. Die Pixelwerte in den Rasterdaten stellen die *Brightness Temperature* (BT) dar. Die weitere Vorprozessierung der Daten hin zur *Land Surface Temperature* (LST) erfolgte im Rahmen dieser Forschungsarbeit.

Die thermische Infrarotmessung wurde an zwei aufeinanderfolgenden Tagen, dem 12.11.2022 in Ost-West Richtung und dem 13.11.2022 in West-Ost Richtung, zur Abendzeit zwischen 19:52 und 21:26 durchgeführt. Für das gesamte Untersuchungsgebiet liegen die Thermaldaten an beiden Befliegungstagen fast vollständig vor. Die Wärmebilder bestehen aus 21 GeoTIFF-Dateien vom 12.11.2022 und 22 GeoTIFF-Dateien vom 13.11.2022.

Üblicherweise beginnt die Heizperiode etwa Anfang Oktober und endet etwa Ende April, sodass die Aufnahmen während der Heizperiode gemessen wurden. Dies ist wichtig, da die zugrunde liegende wissenschaftliche Fragestellung auf der These basiert, dass die Abwärme von Gebäuden aus thermischen Infrarotmessungen hervorgeht. Aufgrund der höheren Innentemperatur gegenüber der Außentemperatur entsteht ein Wärmefluss von innen nach außen, der auf Wärmebildern sichtbar wird. Bei energieeffizienten Gebäuden sollte dieser Wärmefluss möglichst gering ausfallen (Boomsma et al., 2016). Die Oberflächentemperaturen sind durch die zeitliche Aufnahme in der Abendzeit, und damit einige Stunden nach Sonnenuntergang, nur noch in geringem Maße von der Aufwärmung durch Sonneneinstrahlung beeinflusst (Bitelli et al., 2015, S. 2157).

2.1.3 Klimatische Bedingungen

Da die beobachtete Oberflächentemperatur unter anderem von lokalen klimatischen Schwankungen beeinflusst wird (Hemachandran et al., 2018, S. 241–242; Voogt & Oke, 2003, S. 371), wurden die klimatischen Bedingungen von der nächstgelegenen Wetterstation

herangezogen. Der Flughafen Liverpool hat eine Wetterstation, welche halbstündlich Wetterdaten speichert. Tabelle 1 zeigt die relevanten Daten dieser Wetterstation für die beiden Beobachtungszeiträume. Die Daten wurden über die Webseite <https://www.timeanddate.com> bezogen. Da auch die Feuchtigkeit der Oberflächen die gemessenen Temperaturwerte beeinflusst, ist zu erwähnen, dass es am 13.11.2022 in der Zeit von 11:50 bis 13:20 zu Niederschlägen kam.

Tabelle 1: Wetterdaten der Beobachtungszeiträume

Datum	Uhrzeit	Temperatur	Wetterlage	Windrichtung
12.11.2022	19:50	12° C	Vorbeiziehende Wolken	Ost nach West
12.11.2022	20:20	13° C	Vorbeiziehende Wolken	Südost nach Nordwest
12.11.2022	20:50	13° C	Vorbeiziehende Wolken	Südsüdost nach Nordnordwest
12.11.2022	21:20	12° C	Vorbeiziehende Wolken	Südsüdost nach Nordnordwest
13.11.2022	19:50	11° C	Vorbeiziehende Wolken	Ostsüdost nach Westnordwest
13.11.2022	20:20	11° C	Vorbeiziehende Wolken	Ostsüdost nach Westnordwest
13.11.2022	20:50	11° C	Vorbeiziehende Wolken	Ostsüdost nach Westnordwest
13.11.2022	21:20	10° C	Vorbeiziehende Wolken	Ostsüdost nach Westnordwest

2.1.4 Energy Performance Certificates

Die *Energy Performance Certificates* (EPCs) enthalten Informationen zur Energieeffizienz von Gebäuden. Die wichtigsten Informationen in den EPCs sind die Energieeffizienzbewertung (Skala von 1 bis 100, wobei 100 die höchste Effizienz darstellt), die Energieeffizienzklasse (Skala von A bis G, wobei A die höchste Effizienz darstellt), der jährliche Energieverbrauch und die damit verbundenen CO₂-Emissionen pro Jahr, der Gebäudetyp, die Heiztechnik, die Dämmeigenschaften verschiedener Bauelemente wie Dach, Wände, Fenster und Bodenplatte sowie empfohlene Maßnahmen zur Verbesserung der Energieeffizienz.

Im Vereinigten Königreich werden die EPC-Daten vom Ministry of Housing, Communities and Local Government zur Verfügung gestellt. Die britische Regierung führt ein öffentliches Register über die EPCs und stellt die Daten mittels Bulk-Download oder über *Application Programming Interfaces* (APIs) der Forschung bereit. Diese freie Datenverfügbarkeit war der Hauptgrund dafür, einen Thermaldatensatz aus dem Vereinigten Königreich für die vorliegende Arbeit zu verwenden.

Im Vereinigten Königreich wurde ab 2007 schrittweise eine Verpflichtung zur Erstellung von Energieausweisen für Gebäude beim Bau, dem Verkauf oder der Vermietung eingeführt. Die in dieser Arbeit genutzten Energieausweise wurden im Februar 2024 aus dem offiziellen Register heruntergeladen. Für das Untersuchungsgebiet liegen insgesamt 179.932 EPCs vor.

Da das Register bereits seit 2007 geführt wird, existieren für viele Gebäude mehrere EPCs. Für die weitere Analyse sind nur die aktuellsten EPCs für jedes Gebäude von Bedeutung. Eine Filterung nach den jeweils jüngsten EPCs je Objekt verringert die Anzahl der Datensätze auf 101.434 EPCs.

Obwohl die EPCs Informationen zur Lage enthalten, verfügen sie nicht über direkte Koordinaten. Zur Georeferenzierung wurden die *Unique Property Reference Numbers* (UPRNs) herangezogen. Die UPRN ist ein eindeutiger Identifikator (Geocode) für Gebäude und andere adressbasierte Orte im Vereinigten Königreich. Die Ordnance Survey als Behörde für Landesvermessung und Kartographie im Vereinigten Königreich stellt Daten zur Georeferenzierung in einem offenen Datensatz bereit, der zu jeder UPRN eindeutige Lagekoordinaten enthält.

Die verwendeten Thermaldaten können für die Gebäude lediglich Informationen zu der LST der Dachoberflächen liefern. Aufgrund dessen sind Informationen über die Dächer der Gebäude von großer Bedeutung. Die EPCs enthalten neben der Energieeffizienz des Gebäudes auch Angaben zur Energieeffizienz des Daches, in Form einer fünfstufigen Bewertungsskala (*Very Poor, Poor, Average, Good* und *Very Good*). Zudem enthalten die EPCs ein Textattribut *ROOF_DESCRIPTION*, welches die Dämmung beziehungsweise den Wärmedurchgangskoeffizienten des Daches beschreibt. Die Werte des Feldes setzen sich wie folgt zusammen (der Platzhalter * wird für variable Begriffe verwendet):

- *Flat* * (Beispiele für den Wert des Platzhalters sind *insulated* oder *no insulation*),
- *Pitched* * (Beispiele für den Wert des Platzhalters sind *300+ mm loft insulation* oder *no insulation*),
- *Roof room(s)* * (Beispiele für den Wert des Platzhalters sind *insulated* oder *no insulation*),
- *Thatched* oder *Thatched, with additional insulation*,
- *Other premises above*,
- *Another dwelling above*.

Einige EPCs haben keinen Wert für die Energieeffizienz des Daches. Diese EPCs beschreiben Wohnungen in Mehrfamilienhäusern, welche sich nicht über das Dachgeschoss erstrecken und im Attribut *ROOF_DESCRIPTION* Werte wie *Other premises above* oder *Another dwelling above* enthalten. Diese EPCs wurden für die weitere Analyse ausgeschlossen aufgrund der fehlenden Informationen zur Dachenergieeffizienz. Nach Ausschluss dieser Daten verbleiben 98.974 EPCs für die Modellierung der Energieeffizienz der betrachteten Gebäude.

2.1.5 Gebäudepolygone

Damit die Thermaldaten mit der Energieeffizienz von Gebäuden verknüpft werden können, müssen die Rasterdaten auf Gebäudepolygone bezogen werden. Gebäudepolygone werden

von verschiedensten Datenanbietern zur Verfügung gestellt (z. B. OpenStreetMap oder die Ordnance Survey). Die Gebäude Features aus der Ordnance Survey National Geographic Database bieten die höchste Genauigkeit sowie Vollständigkeit und wurden daher für diese Untersuchung verwendet.

Die Ordnance Survey stellt die Daten über eine API bereit. Neben der Geometrie enthalten die Gebäudefeatures Angaben zur Art der Nutzung, die räumliche Beziehung zu benachbarten Gebäuden (klassifiziert als *Standalone*, *Semi-Connected*, *End-Connected* oder *Multi-Connected*) und der jeweiligen Grundfläche.

Im Untersuchungsgebiet befinden sich 204.079 Gebäude. Die Daten wurden im Februar 2024 heruntergeladen.

2.2 Vorprozessierung der Daten

2.2.1 Zeitlicher Trend

Die Visualisierung der Thermaldaten in einem Geoinformationssystem zeigt am 12.11.2022 deutliche Unterschiede zwischen den Werten im östlichen und westlichen Teilbereich. Der östliche Bereich des Untersuchungsgebiets, der große Teile der Stadt Liverpool umfasst, weist signifikant höhere Temperaturen auf als der westlich des Flusses Mersey gelegene Teilbereich im Metropolitan Borough of Wirral.

Abbildung 2 zeigt das Mosaik der Thermalbilder vom 12.11.2022, welches ohne Korrektur der Emissionsgrade sowie ohne Berücksichtigung der Artefakte zwischen den einzelnen Wärmebildern erstellt wurde. Zusätzlich sind die Zeitstempel für die Beobachtungszeitpunkte der einzelnen Aufnahmen dargestellt. Die Darstellung lässt einen zeitlichen Trend der BTs vermuten: Der westliche Teil des Untersuchungsgebiets, der zeitlich später erfasst wurde, weist niedrigere BTs auf. Ein derartiger zeitlicher Trend lässt sich für die Thermaldaten vom 13.11.2022 nicht erkennen.

Die Schwankungen bei der Erfassung von Oberflächentemperaturen können insbesondere bei hochaufgelösten Daten bereits innerhalb weniger Stunden auftreten. Sowohl die Lage als auch die Intensität thermischer Messergebnisse werden durch Beschattung, unterschiedliche Feuchtigkeitsbedingungen und meteorologische Dynamiken beeinflusst (Coutts et al., 2016, S. 649).

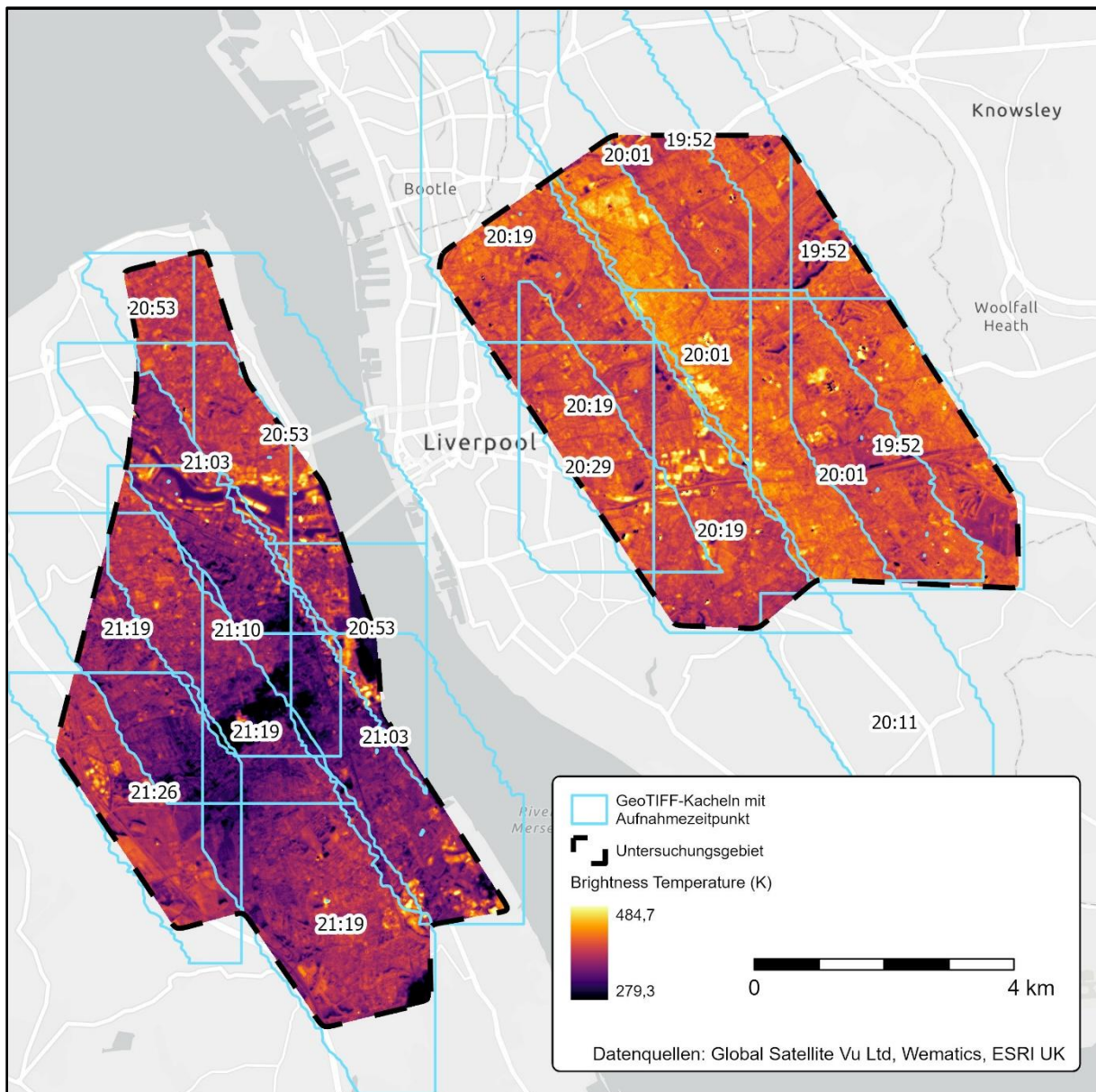


Abbildung 2: Visualisierung des zeitlichen Trends am 12.11.2022 auf einer Karte

Die Hypothese eines vorliegenden zeitlichen Trends in den Thermaldaten wurde mit einer gewichteten Regressionsanalyse getestet. Für beide Befliegungstage wurde jeweils eine separate Regressionsanalyse mit einem Python-Skript berechnet (siehe Anhang A). Die Regression soll den Zusammenhang zwischen der abhängigen Variable BT und der unabhängigen Variable Time (Aufnahmezeitpunkt in Sekunden) erklären.

Für die Regressionsanalyse wurden die Medianwerte der BTs je Wärmebildkachel verwendet. Die Wahl des Medians begründet sich durch dessen Robustheit gegenüber Ausreißern. Aufgrund der Variabilität des Emissionsgrades und dessen starken Wirkung auf die Oberflächentemperaturen kann es zu extremen Einzelwerten kommen, deren Einfluss durch die Verwendung des Medians deutlich reduziert wird.

Um die unterschiedlichen Flächengrößen der einzelnen Wärmebildkacheln in der Regression zu berücksichtigen, wurde die Anzahl der enthaltenen Pixel je Kachel als Gewichtungsfaktor in der Regression verwendet. Damit wird sichergestellt, dass Wärmebilder mit einer höheren Datenmenge einen größeren Einfluss auf das Modell haben.

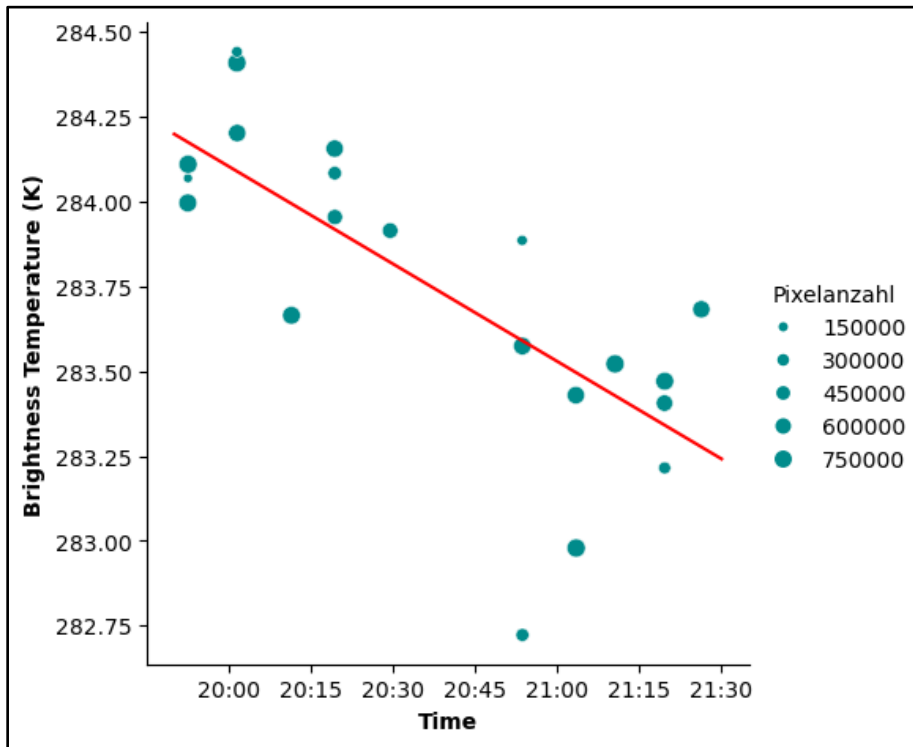


Abbildung 3: Gewichtete Regressionsanalyse zwischen Brightness Temperature und Time am 12.11.2022

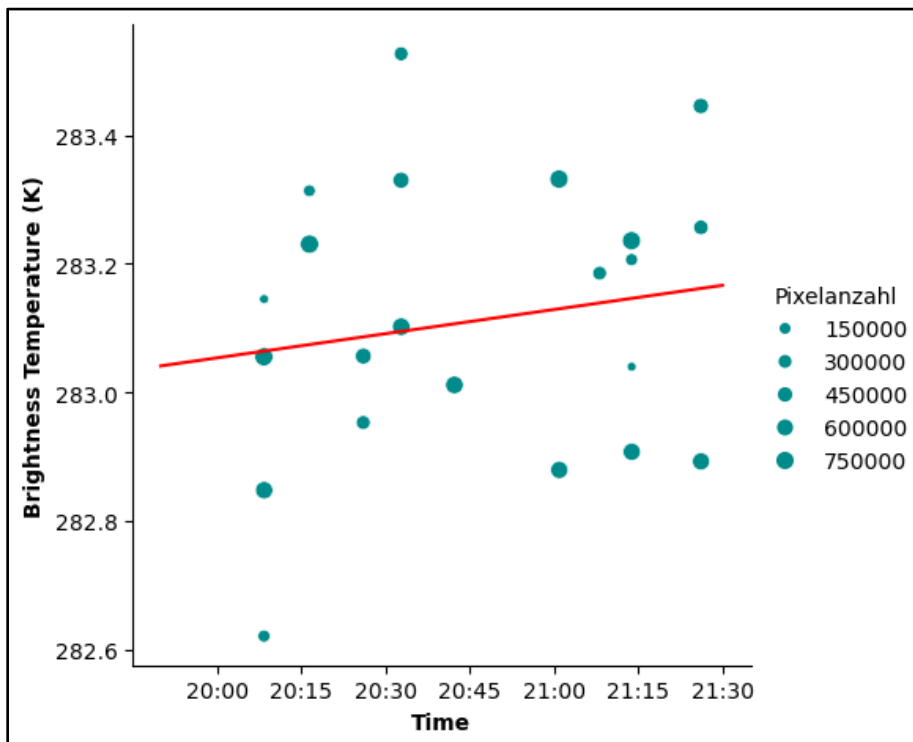


Abbildung 4: Gewichtete Regressionsanalyse zwischen Brightness Temperature und Time am 13.11.2022

Abbildung 3 und Abbildung 4 zeigen die Ergebnisse der gewichteten Regressionsanalysen für die thermischen Infrarotaufnahmen vom 12.11.2022 beziehungsweise 13.11.2022. Die abhängige Variable BT ist auf der Y-Achse dargestellt und die unabhängige Variable Time auf der X-Achse.

Die berechneten Regressionsparameter sind in Tabelle 2 zusammengefasst.

Tabelle 2: Regressionsparameter für die Berechnung des zeitlichen Trends

Datum	Achsenabschnitt	Steigung	Pearson r	r ²	p-Wert
12.11.2022	295,599	-0,000160	-0,75	0,57	6,71*10 ⁻⁹
13.11.2022	281,553	0,000021	0,20	0,03	0,45

Die Regression für die Daten vom 12.11.2022 zeigt mit einem Pearson r Wert von -0,75 einen starken negativen Zusammenhang zwischen der Brightness Temperature und der unabhängigen Variable Time. Eine spätere Beobachtung korreliert mit einer niedrigeren Temperatur, was die Hypothese eines zeitlichen Trends in den Daten bestätigt. Das Bestimmtheitsmaß r² mit einem Wert von 0,57 kann bei Betrachtung der Komplexität der beobachteten thermischen Daten als gut betrachtet werden, da zahlreiche potenzielle Unsicherheitsfaktoren – wie Emissivität, Wolkenbedeckung, Feuchtigkeit oder Windrichtung – die Messergebnisse beeinflussen. Der p-Wert nahe Null weist auf einen hochsignifikanten Zusammenhang hin.

Die Regressionsanalyse für die Daten vom 13.11.2022 bestätigt, dass an diesem Tag kein zeitlicher Trend die BT beeinflusst. Die Varianz der BTs wird nur zu etwa 3 % durch das Modell erklärt. Der schwache positive Zusammenhang (Pearson r Wert von 0,20) kann als zufällig betrachtet werden, da der p-Wert von 0,45 keine statistische Signifikanz darlegt.

Mögliche Erklärungsversuche für den am 12.11.2022 beobachteten zeitlichen Trend liegen in variierenden Feuchtigkeitsbedingungen und meteorologischen Veränderungen. Insbesondere die natürliche Abkühlung im Tagesverlauf infolge des Sonnenuntergangs kann zu Temperaturveränderungen während der Aufnahme führen, da eine spätere Beobachtung eine fortgeschrittenere Abkühlungsphase durchlaufen hat. Am 13.11.2022 könnte dieser Effekt bereits zu Beginn der Aufnahme weitgehend abgeschlossen gewesen sein. Durch den Regen um die Mittagszeit waren die Oberflächen weniger Sonneneinstrahlung ausgesetzt und konnten durch die erhöhte Feuchtigkeit weniger Wärme aufnehmen. Auch lokal wechselnde Bewölkung sowie Änderungen der Windrichtung während der Befliegung am 12.11.2022 können die Veränderungen der Temperaturen hervorgerufen haben.

In einer Forschungsarbeit zur Schätzung der LST aus thermischen Infrarotaufnahmen wird vorgeschlagen, die LSTs auf die lokale Sonnenzeit zu normalisieren, um unterschiedliche

lokale Sonnenzeiten einzelner Pixel auszugleichen (Li et al., 2013, S. 32). In Anlehnung an diese Methode wurden die BTs vom 12.11.2022 auf den Beobachtungszeitpunkt des ersten Wärmebilds normalisiert (siehe letzter Teil des Python-Skripts in Anhang A). Die normalisierten Temperaturwerte werden in den weiteren Analysen verwendet.

2.2.2 Erstellung der Mosaik

Bei der Erstellung der Mosaik treten Artefakte beziehungsweise sichtbare Übergänge zwischen den einzelnen Thermalaufnahmen auf. Abbildung 5 veranschaulicht vier Beispiele solcher Artefakte in den resultierenden Mosaiken.

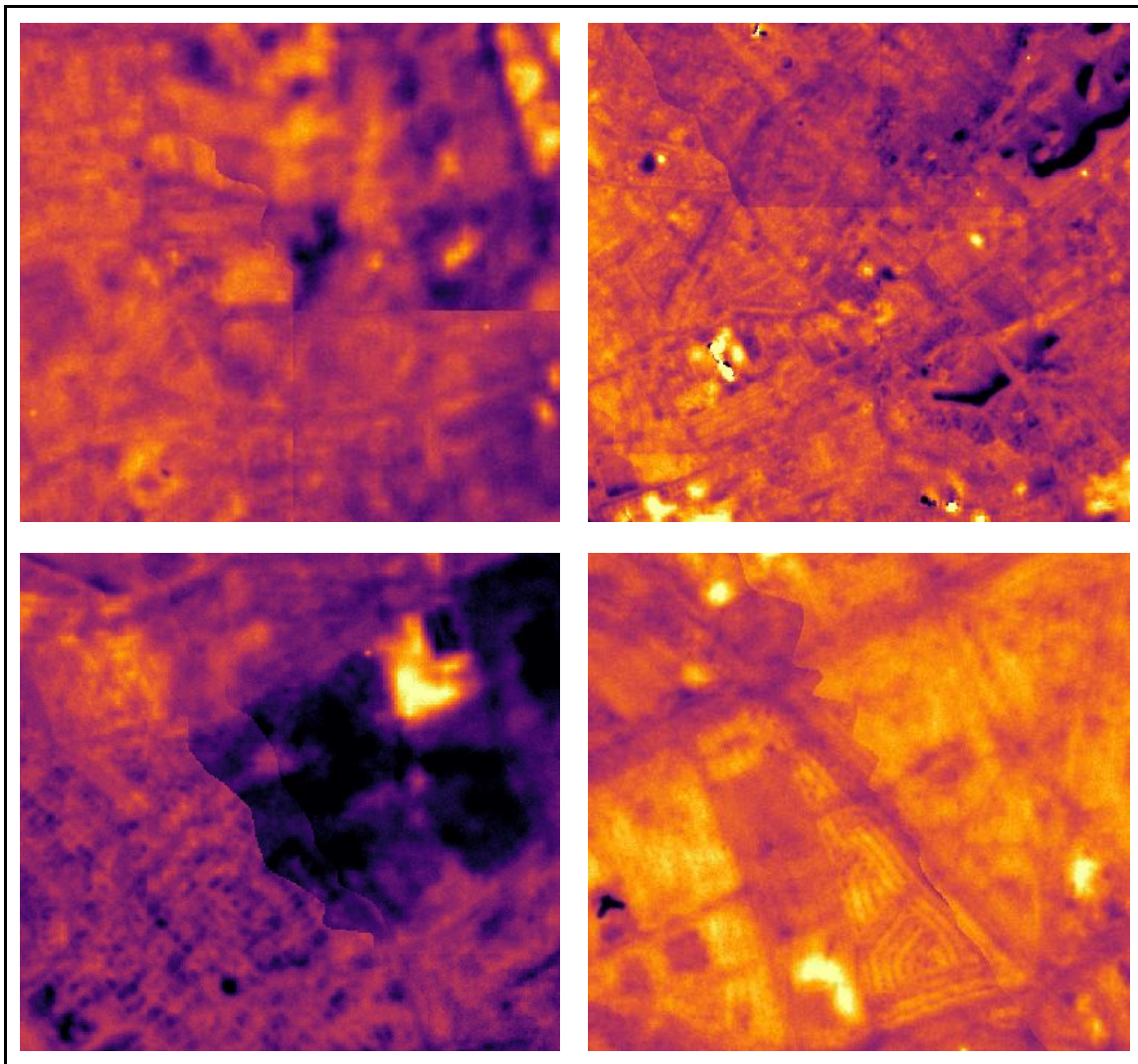


Abbildung 5: Beispiele für Artefakte

Die Thermalbilder besitzen sowohl Überlappungsbereiche als auch Kanten, in denen die Bilder direkt aneinanderstoßen. Durch Ausrichtungsfehler oder photometrische Unterschiede können in diesen Bereichen sichtbare Artefakte entstehen. Die Übergänge können mithilfe von *Blending*-Algorithmen geglättet werden (Pandey & Pati, 2019, S. 240).

ArcGIS Pro nutzt die *Transition smoothing method* als *Blending*-Algorithmus. Dieser Algorithmus reduziert die sichtbaren Unterschiede in den Übergangsregionen, indem die

Pixelwerte durch gewichtete Mittelwerte der überlappenden Raster ersetzt werden (Pandey & Pati, 2019, S. 247). Die Übergänge zwischen Bildern wurden durch sogenannte *Seamlines* gekennzeichnet, die auf Grundlage der Footprints der einzelnen GeoTIFF-Dateien erstellt wurden.

Eine vollständige Glättung aller Übergänge im Mosaik vom 12.11.2022 erwies sich als schwierig (siehe Abbildung 6). Zwar trägt die Regression des zeitlichen Trends zu einer verbesserten Anpassung bei, sie kann jedoch den zeitlichen Trend nicht in jeder Einzelaufnahme vollständig eliminieren. Auch die zuvor beschriebenen Methoden führten nicht zu einem artefaktfreien Mosaik, ohne eine umfangreiche Anpassung der Daten vorzunehmen.

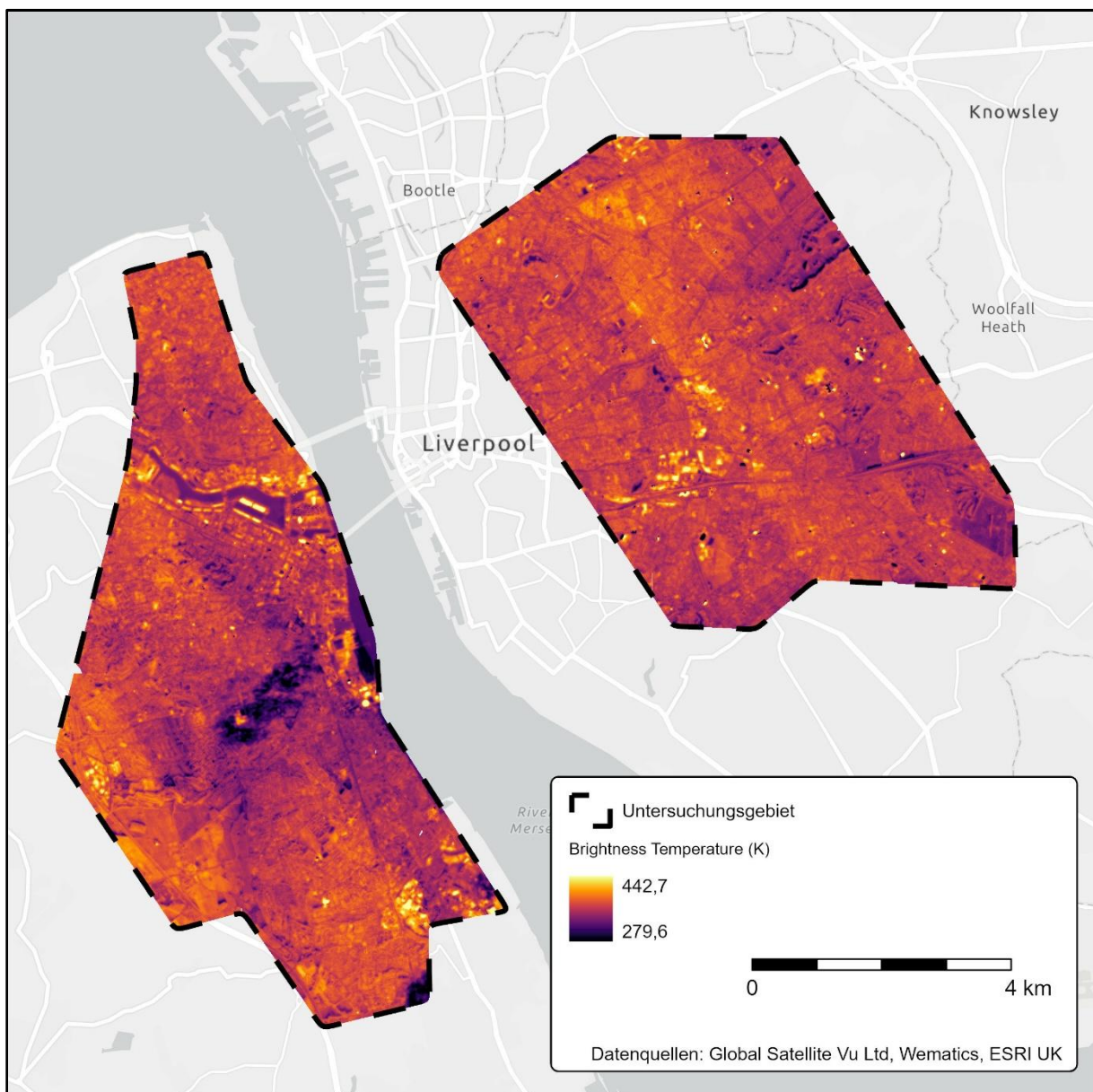


Abbildung 6: Mosaik der Thermalbilder vom 12.11.2022

Das Mosaik vom 13.11.2022 lässt weichere Übergänge erkennen. Die Verwendung von *Blending-Algorithm*en und *Seamlines* erzeugt kontinuierliche Übergänge zwischen den Bildern.

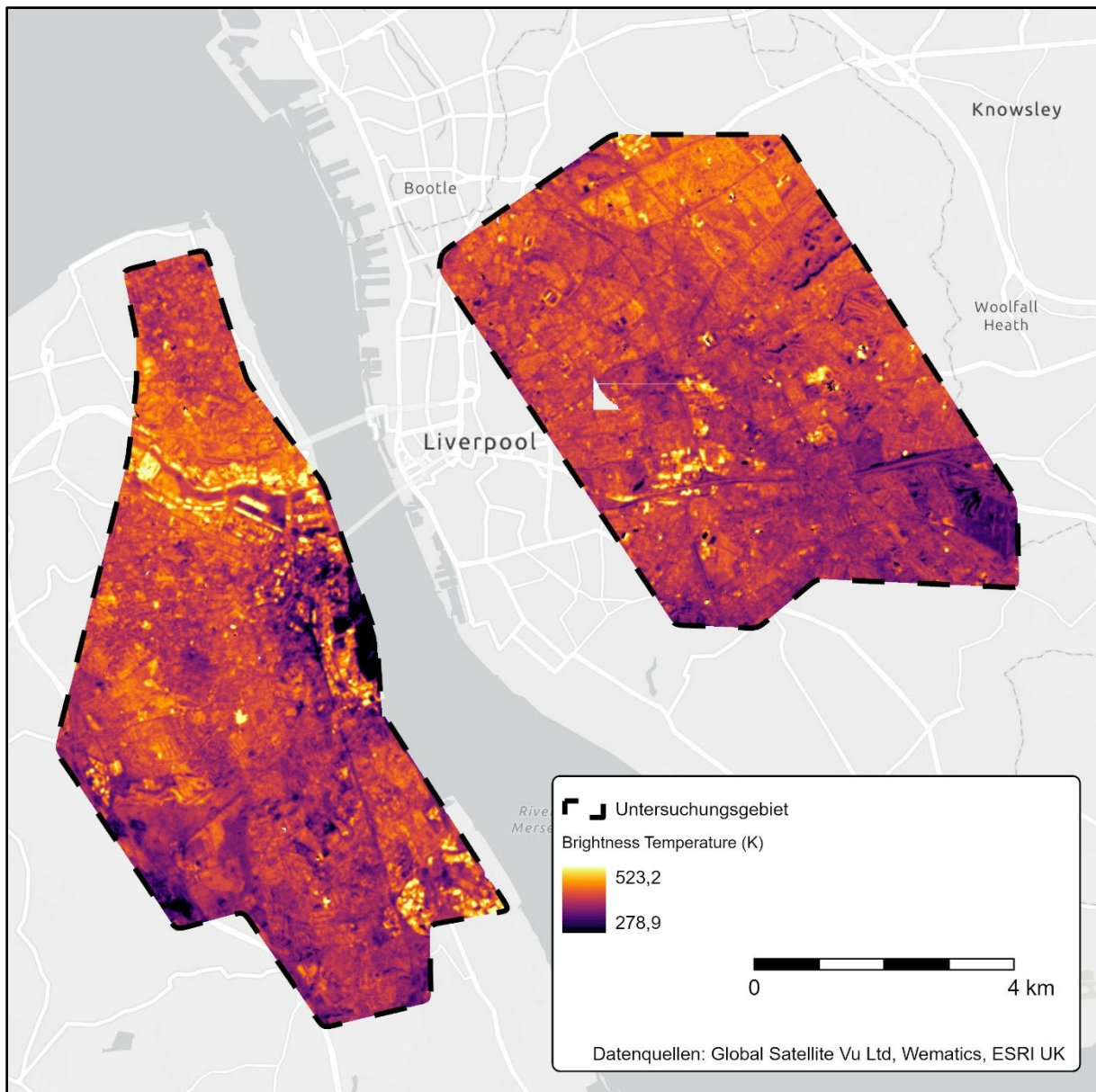


Abbildung 7: Mosaik der Thermalbilder vom 13.11.2022

2.2.3 Berechnung der mittleren Brightness Temperature je Gebäude

Üblicherweise würde nach der Mosaikbildung die Berechnung der LSTs erfolgen. Für die Beantwortung der Forschungsfrage sind nur die LSTs der Gebäude im Untersuchungsgebiet von Bedeutung. Da zur Berechnung der LST aus der BT insbesondere die Emissivität der Landoberflächen berücksichtigt werden muss, welche je nach Oberfläche beziehungsweise Oberflächenmaterial unterschiedliche Emissionsgrade aufweist, wurden zunächst die mittleren BTs für die Gebäude berechnet. Anschließend kann die Emissivität gezielt für die

Dachmaterialien der Gebäude in die Berechnung einbezogen werden. Diese Vorgehensweise vereinfacht den Arbeitsablauf der Vorprozessierung der Thermaldaten.

Gebäudedächer umfassen häufig neben der eigentlichen Dachhaut diverse Objekte wie Dachfenster, Schornsteine oder Lüftungsöffnungen. Da die Bodenauflösung der Thermaldaten mit 3,5 m eine Unterscheidung zwischen der Dachhaut und den verschiedenen Kleinstelementen nicht zulässt, wurden die Dächer als homogene Einheiten betrachtet. Diese Annahme lag ebenfalls der Studie von Hemachandran et al. (2018, S. 242) zugrunde, die Thermaldaten mit einer Auflösung von 0,5 m verwendet.

Um die mittleren BTs für jedes Gebäudedach zu berechnen, mussten zunächst die relevanten Rasterzellen je Gebäude ausgewählt werden. Diese Auswahl ist nicht trivial, da die Rasterzellen einerseits möglichst wenig von äußeren Faktoren beeinflusst werden sollen und andererseits die Größe der Stichprobe nicht erheblich verkleinert werden soll. Damit die resultierenden mittleren Temperaturwerte möglichst unverzerrt bleiben, sollten Mischpixel vermieden werden. Mischpixel sind Zellen, welche mehrere Oberflächen beschreiben wie z. B. das Gebäudedach zusammen mit Vegetation oder Straßen. Zudem weisen die Außenwände der Gebäude andere Emissionsgrade als die Dächer der Gebäude auf. Gulbe et al. (2017, S. 49) verkleinern die Gebäudepolygone um eine Pufferfläche von 0,5 m, um die Auswirkung der Emissionsgrade der Wände sowie anderer Oberflächen zu minimieren. Hemachandran et al. (2018, S. 244) gehen in ihrer Studie zur Ermittlung von Schwachstellen an der Dachisolierung einen Schritt weiter und unterteilen die Dächer in zwei Zonen, wobei eine Zone den Randbereich eines Gebäude mit einer Breite von einem Meter umfasst.

In dieser Forschungsarbeit wurde die Ausdehnung der Gebäudepolygone um eine Pufferfläche von einem Meter reduziert. Die Rasterzellen wurden anschließend anhand der skalierten Gebäudepolygone selektiert.

Die Verschneidung der Rasterzellen für die BTs mit den Vektorobjekten für die Gebäude kann mit verschiedenen Vergleichsoperatoren erfolgen. Die meisten Geoverarbeitungen in Geoinformationssystemen nutzen zur statistischen Auswertung von Rastern die Mittelpunkte der einzelnen Zellen. ArcGIS Pro verwendet diese Methode beispielsweise beim Geoprozessierungswerkzeug *Zonal Statistics as Table*. Eine räumliche Verschneidung auf Grundlage der Mittelpunkte der Rasterzellen birgt jedoch ein großes Risiko für Mischpixel. Abbildung 8 zeigt eine Reihenbebauung zusammen mit den BTs, den ursprünglichen Gebäudepolygonen, deren skalierte Ausdehnung nach Verringerung um die Pufferfläche sowie die Grenzen der Rasterzellen und deren Mittelpunkte. Eine Auswahl der Rasterzellen anhand der Mittelpunkte würde in dem dargestellten Beispiel auch solche Zellen zur Berechnung der Mittelwerte einbeziehen, die große Flächenanteile außerhalb der skalierten Gebäudepolygone umfassen.

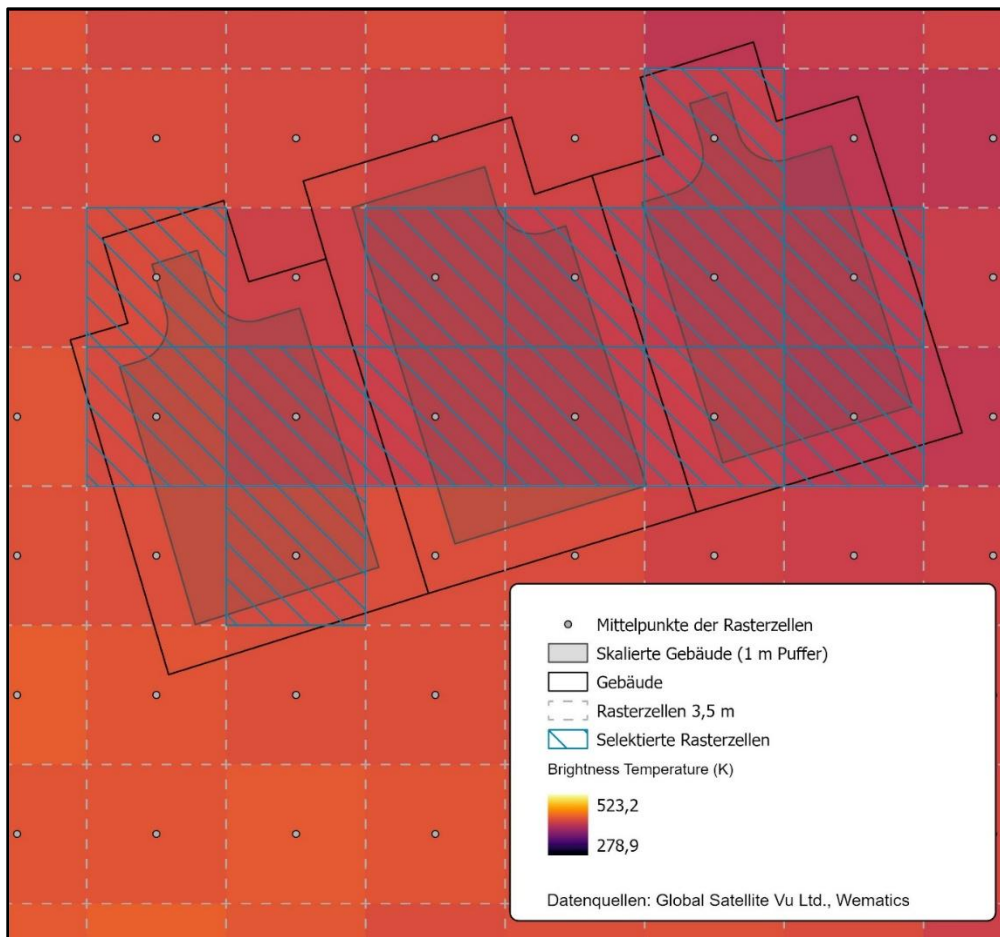


Abbildung 8: Beispiel zur Selektion der Rasterzellen für die BTs anhand der Mittelpunkte

Die Selektion der Rasterzellen zur Berechnung der mittleren BTs für jedes Gebäude wurde infolgedessen auf die überlappenden Flächenanteile der Zellen mit den skalierten Gebäuden gestützt. Verschiedene Grenzwerte für den Flächenanteil wurden als Basis zur Selektion getestet.

Eine Selektion, die ausschließlich vollständig überlappende Rasterzellen für die Thermaldaten vom 13.11.2022 mit den skalierten Gebäuden auswählt, würde den Großteil der Gebäude aus der weiteren Auswertung ausschließen. Nur bei 8.433 Gebäuden überlagern die verkleinerten Gebäudepolygone mindestens eine Rasterzelle vollständig. Ein Grenzwert des Flächenanteils von 50 % würde 90.007 Gebäude einbeziehen. Der Einfluss äußerer Faktoren auf die BT bei einem Grenzwert von 50 % für den Flächenanteil kann bis zu 50 % betragen, was die Zuverlässigkeit der mittleren BTs erheblich verringern würde. Ein Grenzwert von 75 % stellt einen guten Kompromiss zwischen der Größe der verbleibenden Stichprobe und der Reduzierung des Einflusses von Mischpixeln dar. Bei den Thermaldaten vom 13.11.2022 überdecken bei diesem Grenzwert die skalierten Geometrien von 56.464 Gebäuden mindestens eine Rasterzelle zu 75 % oder mehr.

Abbildung 9 zeigt die Selektion der Rasterzellen für die BTs anhand der überlappenden Flächenanteile von mindestens 75 % am vorherigen Beispiel einer Reihenbebauung.

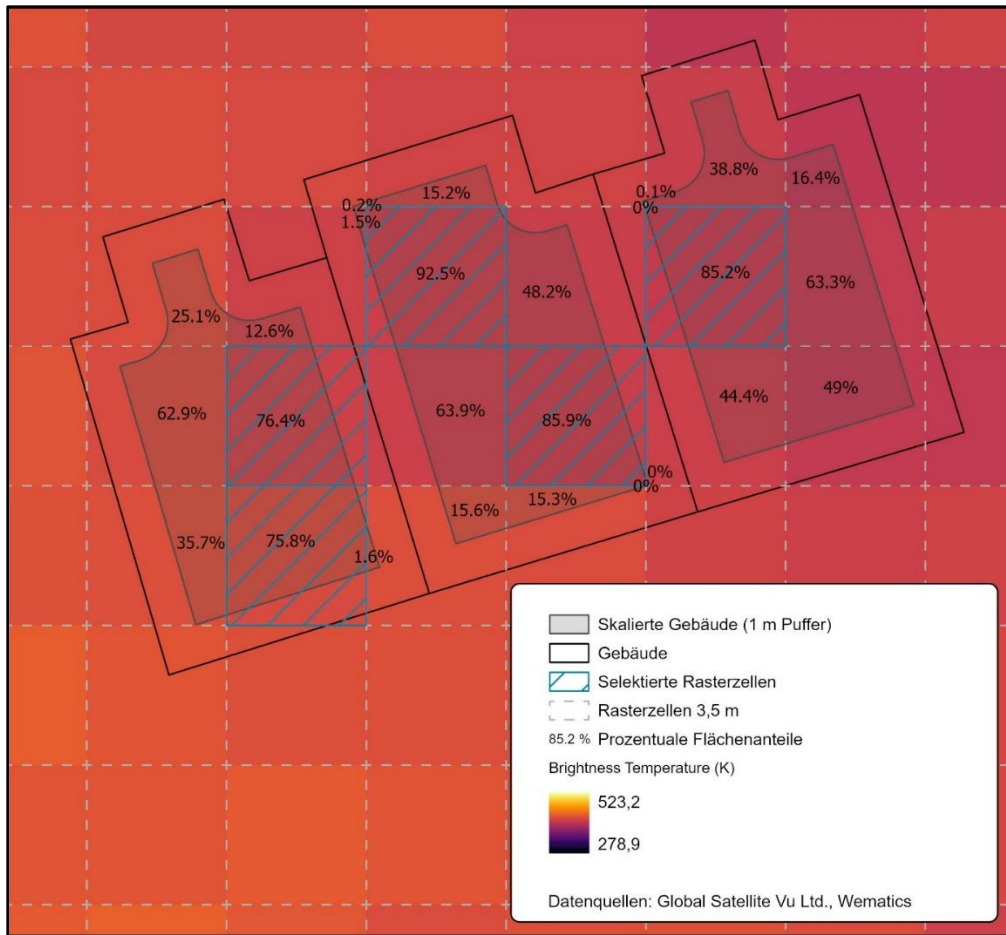


Abbildung 9: Beispiel zur Selektion der Rasterzellen für die BTs anhand der Flächenanteile

Für die thermischen Aufnahmen vom 12.11.2022 überlagern die skalierten Gebäudepolygone von 56.491 Objekten mindestens eine Rasterzelle zu 75 % oder mehr. Damit die Daten der beiden Aufnahmetage in der Analyse miteinander verglichen werden können, wurden nur Gebäude berücksichtigt, bei denen mindestens eine Rasterzelle der Thermaldaten vom 12.11.2022 sowie eine Rasterzelle der Thermaldaten vom 13.11.2022 mit den skalierten Gebäudepolygonen zu mindestens 75 % überlappt. Die Größe der Stichprobe für Gebäude mit EPCs verringert sich dadurch auf 44.922.

Die mittleren BTs je Gebäude wurden mittels der Temperaturwerte der Pixel aus den vorherigen Selektionen berechnet. Der Flächenanteil der Rasterzellen wurde als Gewichtungsfaktor bei der Berechnung der Mittelwerte verwendet, sodass Rasterzellen, welche einen größeren Anteil der skalierten Gebäude abdecken, einen höheren Einfluss auf die resultierenden mittleren BTs haben.

Die Berechnung der gewichteten Mittelwerte erfolgte mit einem Python Code mit ArcGIS Pro Notebook (siehe Anhang B). Die berechneten mittleren BTs konnten anschließend über das Schlüsselattribut Gebäude-ID mit den Gebäudeobjekten verknüpft werden.

Abbildung 10 zeigt den schematischen Ablauf der Geoprozessierung zur Berechnung der mittleren BTs für die Gebäude unter Verwendung eines ArcGIS Pro Modells.

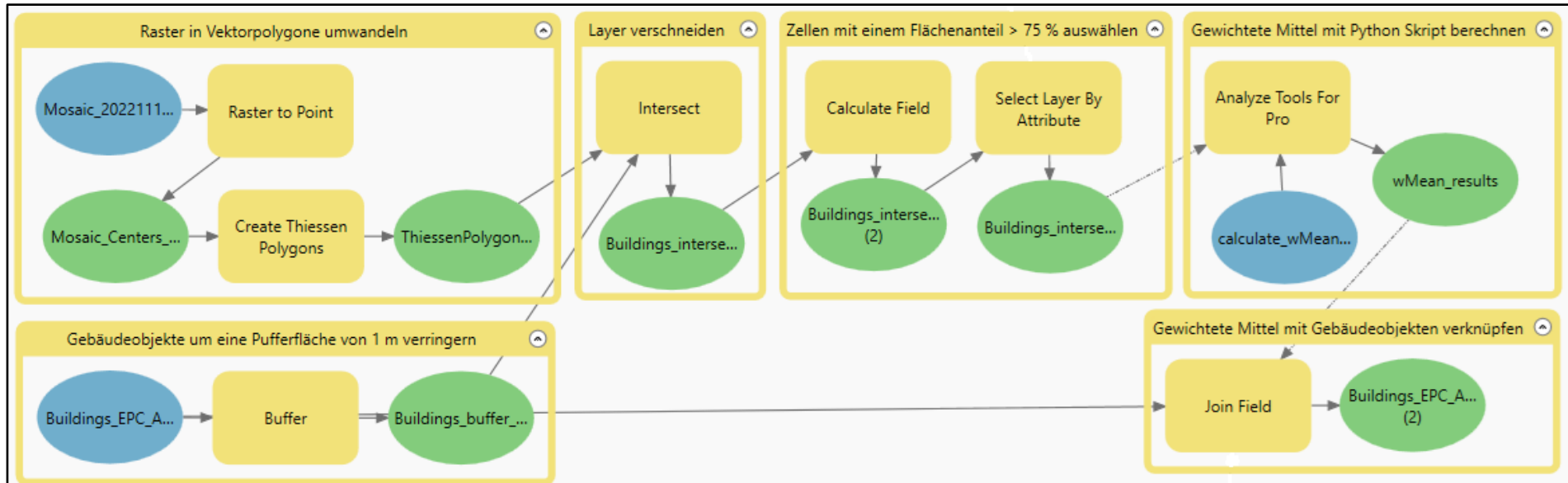


Abbildung 10: Schematische Darstellung der Geoprozessierung zur Berechnung der mittleren BTs für die Gebäude (erstellt mit dem ModelBuilder aus ArcGIS Pro)

2.2.4 Berechnung der Land Surface Temperatures

Zur Berechnung der LST aus der BT müssen atmosphärische Effekte korrigiert und die Emissivität der Landoberflächen berücksichtigt werden. Zu den atmosphärischen Korrekturen gehören die atmosphärische Transmission, die aufwärts und abwärts gerichtete Strahlung der Atmosphäre sowie die reflektierte Umgebungsstrahlung. Für eine Korrektur dieser Effekte ist eine genaue Kenntnis der vertikalen Profile des atmosphärischen Wasserdampfs, der Temperatur und des Luftdrucks notwendig, die räumlich und zeitlich stark variieren und mit Radiosonden bestimmt werden (Li et al., 2013, S. 17; Rüdiger et al., 2024, S. 3).

Solche Parameter lagen für die durch SatelliteVU durchgeführte Befliegung nicht vor. Da die Flughöhe während der Aufnahmen konstant blieb und das Untersuchungsgebiet räumlich begrenzt ist, kann angenommen werden, dass die atmosphärischen Effekte innerhalb des Gebietes weitgehend homogen sind. Aus diesem Grund wurde auf eine atmosphärische Korrektur verzichtet.

Für die Modellierung der Energieeffizienz der Gebäude sind die relativen Unterschiede der Oberflächentemperaturen zwischen den Gebäuden von Relevanz. Die atmosphärischen Effekte würden diese Temperaturdifferenzen nicht wesentlich beeinflussen. Bei Vergleichen zwischen den unterschiedlichen Aufnahmetagen muss dieser Unsicherheitsfaktor jedoch berücksichtigt werden.

Die Emissivität ist der Parameter mit dem größten Einfluss auf die LST (Rüdiger et al., 2024, S. 37). Sie beschreibt den Emissionsgrad einer Oberfläche, beziehungsweise die Fähigkeit einer Oberfläche Wärme zu abstrahieren. Der Emissionsgrad ist definiert als das Verhältnis zwischen der emittierten Strahlung einer Oberfläche zur Strahlung eines idealen schwarzen Körpers und kann Werte zwischen 0 (keine Emission) und 1 (vollständige Emission) annehmen.

Der Emissionsgrad kann mit verschiedenen Methoden ermittelt werden:

- 1) **Globaler Emissionsgrad:** Bei dieser Methode wird ein konstanter Wert für die Emissivität über das gesamte Untersuchungsgebiet hinweg angenommen. Der Wert orientiert sich an den Emissionsgraden typischer Flächen im Untersuchungsgebiet und wird als Durchschnittswert von Literaturangaben festgelegt.
- 2) **Landnutzungsklassen:** Eine weit verbreitete Methode zur Festlegung des Emissionsgrades ist die Anpassung an die jeweiligen Landnutzungsklassen. Den Pixelwerten werden je nach zugeordneter Landnutzungsklasse spezifische Emissionsgrade zugewiesen, die durch Feldmessungen ermittelt oder aus der Fachliteratur entnommen werden. Die Landnutzungsklassen können aus verschiedenen Quellen abgeleitet werden. Häufig sind lokale Kartierungen der unterschiedlichen Landnutzungen bei den jeweiligen Vermessungs- und Katasterbehörden vorhanden. Für

großflächige Analysen mit niedriger Auflösung können Fernerkundungsprodukte wie die *CORINE Land Cover* verwendet werden. Diese Methode findet in zahlreichen Studien Anwendung. Hierzu gehören die Studie von Coutts et al. (2016) zur Analyse von Hotspots mit hohen städtischen Lufttemperaturen in unterschiedlichen räumlichen Ebenen sowie die Studie von Kamaraj et al. (2021) zur Untersuchung der wärmeverursachenden städtischen Oberflächenmaterialien.

- 3) **NDVI:** Eine weitere übliche Methode berücksichtigt die Emissivität unter Zuhilfenahme des *Normalized Difference Vegetation Index* (NDVI). Der NDVI ist ein Indikator zur Bewertung der Vegetationsdichte und -gesundheit, der aus dem sichtbaren roten Licht sowie nahinfraroten Spektralbereich berechnet wird. Jedem Pixel wird basierend auf dem zugehörigen NDVI ein individueller Emissionsgrad zugewiesen. Dieses Verfahren wird häufig bei der Analyse städtischer Hitzeinseln anhand thermischer Infrarotaufnahmen eingesetzt, z. B. in den Studien von Liu et al. (2015) und Zhao et al. (2018) zur Untersuchung der Effekte sowie den wärmeverursachenden Faktoren städtischer Hitzeinseln.
- 4) **TES:** Der *Temperature and Emissivity Separation* (TES) Algorithmus ermöglicht die gleichzeitige Schätzung der Temperatur und der Emissivität auf Basis von Thermaldaten mit mehr als vier Wärmebändern. Die Methodik besteht aus drei Modulen und nutzt sowohl die emittierte Strahlung der Landoberfläche als auch die abwärts gerichtete atmosphärische Strahlung über alle Wellenlängenspektren hinweg. Das erste Modul liefert vorläufige Schätzwerte für die LST und die Emissivität. Im zweiten Modul wird das normalisierte Emissionsspektrum geschätzt. Das dritte Modul wandelt die *Maximum Minimum Difference* und das Emissionsspektrum in die absolute Emissivität um und berechnet schließlich die LST (Oltra-Carrió et al., 2012, S. 299).
- 5) **TISI:** Der *Temperature Independent Spectral Indices* (TISI) Algorithmus schätzt die Emissivität temperaturunabhängig auf Grundlage des Verhältnisses thermischer Infrarotstrahlung bei unterschiedlichen Wellenlängen. Voraussetzung ist die Verfügbarkeit eines Tages- und Nachtbildes. Der Emissionsgrad wird zunächst im mittleren Infrarotbereich bestimmt, der nur gering von der Temperatur beeinflusst wird. Abschließend wird aus diesen Daten die Emissivität der thermalen Infrarotkanäle abgeleitet (Oltra-Carrió et al., 2012, S. 299).

Ein Vergleich der Methoden 3) bis 5) zur Ermittlung des Emissionsgrades von Oberflächen städtischer Gebiete zeigt für die Methoden TES und TISI gute Ergebnisse. Der TES-Algorithmus konnte die LSTs für künstliche urbane Oberflächen am genauesten wiedergeben. Die Berücksichtigung der Emissivität mit dem NDVI erwies sich hingegen als ungeeignet, um zwischen verschiedenen künstlichen Oberflächen zu unterscheiden (Oltra-Carrió et al., 2012).

Da bei der Modellierung der Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten ausschließlich künstliche Oberflächen (Gebäudedächer) betrachtet werden, ist eine Nutzung des NDVI nicht zielführend. Die Algorithmen TES und TISI erfordern mehrere Wärmebänder oder multispektrale Daten mit hoher Auflösung. Solche Daten liegen dem Autor nicht vor und wurden während der thermischen Infrarotaufnahmen im Untersuchungsgebiet nicht erfasst. Aus diesem Grund können die Methoden 3) bis 5) für die Berechnung der LSTs ausgeschlossen werden.

Nach dem Ausschluss dieser drei Methoden verbleiben die Möglichkeiten, die Emissivität der Dachoberflächen mit einem globalen Emissionsgrad oder durch unterschiedliche Emissionsgrade je nach Landnutzungsklasse zu berücksichtigen. Bei hohen Bodenauflösungen vernachlässigt jedoch die Nutzung von Landnutzungsclassen das unterschiedliche Dachmaterial der Gebäude (Hemachandran et al., 2018, S. 242). In dem spezifischen Anwendungskontext kann das Verfahren so angepasst werden, dass unterschiedlichen Dachmaterialien verschiedene Emissionsgrade zugewiesen werden. Bitelli et al. (2015) nutzen Emissionsgrade aus der Literatur für verschiedene Dachmaterialien. Damit den Dachoberflächen Emissionsgrade auf Basis des Materials der Dacheindeckung zugeordnet werden können, müssen die Gebäudedächer zunächst nach deren Oberflächenmaterial klassifiziert werden. Allerdings enthalten die vorliegenden EPCs lediglich Informationen zu den Dämmeigenschaften der Gebäudedächer, nicht aber zu den Materialien der Dacheindeckung. Eine Unterscheidung zwischen Materialien wie Ziegel, Beton oder Metall erfordert hochauflösende Multispektralbilder (Coutts et al., 2016, S. 641), die im Rahmen dieser Masterarbeit nicht zur Verfügung stehen.

Die Emissionsgrade der meisten städtischen Oberflächen – mit Ausnahme metallischer Oberflächen – liegen in einem engen Bereich, wodurch potenzielle Fehler bei der Berechnung der LSTs gering bleiben (Rüdiger et al., 2024, S. 37). Gängige Materialien für geneigte Dächer sind Ton- oder Betonziegel und bei älteren Gebäuden Schiefer. Flachdächer älterer Gebäude bestehen häufig aus Bitumen, während bei neueren Gebäuden meist EPDM, ein synthetischer Kautschuk, verwendet wird. Metaldächer kommen vor allem bei Gewerbe- und Industriegebäuden zum Einsatz. Diese Annahmen stützen sich insbesondere auf die Zusammenfassung des *UK Roofing Market Report 2024-2028* der AMA Research, einem Anbieter von Marktforschungs- und Analyseberichten (<https://www.amaresearch.co.uk/report/uk-roofing-market-report-2024-2028>, Abruf am 04.03.2025). Offizielle Statistiken zu den Anteilen der Dachmaterialien im Vereinigten Königreich existieren nicht. Bei Wohnhäusern wird Metall als Dachmaterial selten verwendet, unter anderem wegen den hohen Anschaffungskosten, der traditionellen Ästhetik von Ziegel- oder Schieferdächern sowie einer hohen Geräusentwicklung bei Regen oder Hagel.

Damit der große Unterschied der Emissionsgrade zwischen Metalldächern und herkömmlichen Dachmaterialien die Modellierung nicht beeinflusst, wurde die Analyse auf Wohngebäude begrenzt. Hierzu wurde das Attribut *buildinguse* der Gebäudefeatures verwendet. Dadurch reduziert sich die Anzahl der Gebäude mit EPC-Daten von 44.922 auf 43.000. Viele Nichtwohngebäude – insbesondere Industrie oder landwirtschaftliche Bauten – unterliegen nicht der Verpflichtung zur Erstellung eines Energieausweises (UK Government, 2012, Reg. 5(1)(d)). Dadurch bleibt die Anzahl der Objekte, die durch die Filterung nach Wohnnutzung ausgeschlossen wurden, überschaubar. In Anbetracht des erheblichen Einflusses der Emissivität auf die LST ist die Verringerung des Stichprobenumfangs um ca. vier Prozent als vernachlässigbar zu bewerten.

Basierend auf den zuvor aufgestellten Hypothesen zu den häufigsten Dachmaterialien würde eine Unterteilung der Gebäude sowohl hinsichtlich der Dachform (geneigt oder flach) als auch nach Baujahr (alt oder neu) eine bestmögliche Zuweisung der Emissionsgrade ermöglichen. Die EPCs umfassen Angaben zu den Baualtersklassen der Gebäude (z. B. 1930 bis 1949). Da keine belastbaren Statistiken zu dem Anteil der eingesetzten Dachmaterialien je Baujahr beziehungsweise Baualtersklasse vorliegen, gestaltet sich eine Kategorisierung in neu und alt schwierig. Darüber hinaus ist fraglich, ob das Baujahr als Information ausreicht, da das ursprüngliche Dachmaterial durch Sanierungen verändert worden sein könnte. Unterschiedliche Emissionsgrade wurden infolgedessen nur für verschiedene Dachformen berücksichtigt. Zur Differenzierung der Gebäude nach deren Dachform wurde das Attribut *ROOF_DESCRIPTION* aus den EPC-Daten herangezogen. Als Flachdächer wurden alle Dächer mit einem Merkmal beginnend mit dem Begriff *Flat* festgelegt. Für alle übrigen Gebäude wurde ein geneigtes Dach angenommen.

Tabelle 3 zeigt Emissivitätswerte verschiedener Dachmaterialien, welche aus der Fachliteratur entnommen wurden. Als Quelle dienen die von Kotthaus et al. (2014) ermittelten Emissionsgrade, die Materialproben aus London verwendet. Die räumliche Nähe der Studie zum Untersuchungsgebiet sowie die detaillierte Differenzierung der Dachmaterialien waren ausschlaggebend für die Auswahl der Werte dieser Studie. Da in der Literatur keine Emissionsgrade für Reetdächer verfügbar sind, wurden Reetdächer unter Anwendung einer Filterung des Attributs *ROOF_DESCRIPTION* aus dem Datensatz exkludiert (Merkmalsausprägungen: *Thatched*, *Thatched, with additional insulation* oder *Roof room(s) thatched*), um weitere Unsicherheiten zu minimieren. Hiervon sind nur fünf Gebäude betroffen, sodass für die weitere Analyse 42.995 Gebäude zur Verfügung stehen.

Tabelle 3: Emissionsgrade für Dachmaterialien (entnommen aus Kotthaus et al., 2014)

Dachmaterial	Emissionsgrad ε
Tonziegel	0,92-0,95
Betonziegel	0,91-0,96
Schiefer	0,90-0,93
Bitumen	0,93
EPDM	0,93-0,94
Metall	0,16-0,97

Aufgrund der hohen Kosten für die Anschaffung und Verlegung wird Schiefer heutzutage nur noch selten als Dachmaterial eingesetzt. Bei älteren Bestandsgebäuden ist ein Schieferdach insbesondere bei denkmalgeschützten Gebäuden zu vermuten, welche von der Pflicht zur Erstellung eines EPC ausgenommen sind (UK Government, 2012, Reg. 5(1)(a)). Der Emissionsgrad für Gebäude mit einem geneigten Dach wird infolgedessen ohne Einbeziehung eines Wertes für Schiefer als Mittelwert der Emissionsgrade von Ton- und Betonziegeln zu 0,94 festgesetzt. Für Gebäude mit Flachdächern wurde der Emissionsgrad für die üblichen Dachmaterialien Bitumen und EPDM zu 0,93 gemittelt. Diese Emissionsgrade stellen unter den getroffenen Annahmen und im Rahmen der verfügbaren Daten die bestmögliche Annäherung dar.

Neben den genannten Unsicherheitsfaktoren in den zugrundeliegenden Prämissen, ist zu berücksichtigen, dass die Emissivität von Oberflächen durch fortschreitende Verwitterung beeinflusst werden kann (Kotthaus et al., 2014, S. 206).

Zur Berechnung der LST aus der BT bei bekanntem Emissionsgrad und nur einem verfügbaren thermalen Infrarotkanal wird folgende Formel angewendet (Liu et al., 2015, S. 4812):

$$LST = \frac{BT}{1 + \left(\frac{\lambda \times BT}{\rho}\right) \ln \varepsilon}$$

$$\rho = \frac{h \times c}{\sigma} \approx 14380 \mu m$$

Dabei ist BT die *Brightness Temperature* in Kelvin, λ die Wellenlänge der emittierten Strahlung (4,35 μm), ε der Emissionsgrad der Landoberfläche, h die Planck Konstante ($6,626 \times 10^{-34}$ Js), c die Lichtgeschwindigkeit ($2,998 \times 10^8$ m/s) und σ die Boltzmann-Konstante ($1,38 \times 10^{-23}$ J/K). Anschließend kann die Temperatur durch Subtraktion von 273,15 von Kelvin in Grad Celsius umgerechnet werden.

Dieses Verfahren wurde auf alle mittleren Gebäudetemperaturwerte angewendet. Abbildung 11 zeigt einen Ausschnitt des Untersuchungsgebietes, in dem die mittleren LSTs der Gebäude gemeinsam mit der jeweiligen Energieeffizienzklasse dargestellt sind.



Abbildung 11: Beispielansicht zur LST der Gebäude gemeinsam mit der jeweiligen Energieeffizienzklasse

2.3 Deskriptive Statistiken zu den Daten

2.3.1 Deskriptive Statistik zu den Energy Performance Certificates

Zur Modellierung der Energieeffizienz der Gebäude mittels thermischer Infrarotaufnahmen unter Verwendung der Daten aus den EPCs ist ein genaues Verständnis der Daten essenziell. Die folgenden Seiten beschreiben die Verteilung der relevanten Merkmale aus den EPCs. Die Beschreibungen beziehen sich auf die gefilterte Stichprobe mit 42.995 Gebäuden. Zur Erstellung der Diagramme wurde ein Python-Skript verwendet (siehe Anhang C).

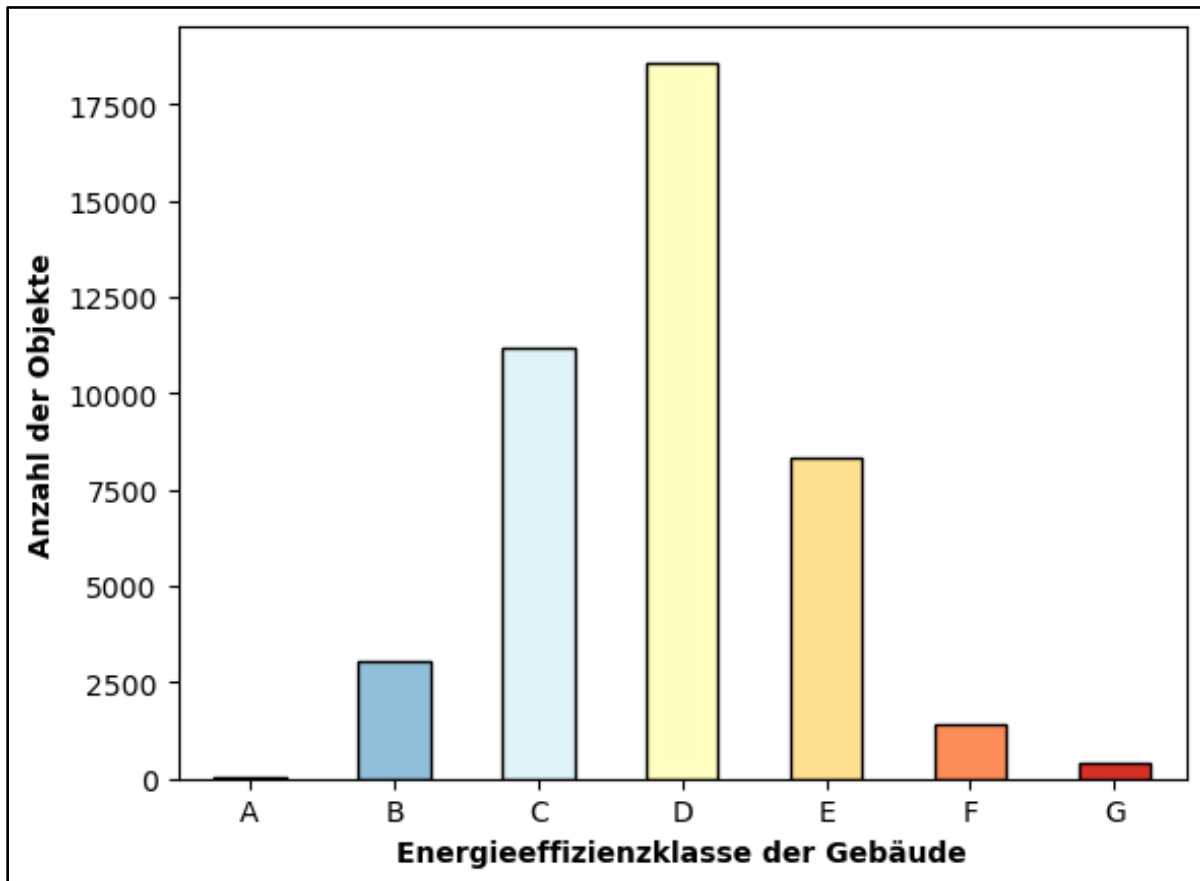


Abbildung 12: Anzahl der Objekte je Energieeffizienzklasse der Gebäude

Abbildung 12 stellt die Anzahl der Gebäude je Energieeffizienzklasse in Form eines Säulendiagramms dar. Die Verteilung zeigt deutlich, dass ein Großteil der Gebäude eine mittlere Energieeffizienz aufweist – etwa 89 % der Gebäude entfallen auf die Energieeffizienzklassen C bis E. Sehr energieeffiziente und sehr energieineffiziente Gebäude (Energieeffizienzklassen A und G) sind im Datensatz kaum vorhanden. Auch die Energieeffizienzklassen B (energieeffizient) und F (energieineffizient) sind selten vertreten.

Das Histogramm in Abbildung 13 veranschaulicht die Verteilung der numerischen Werte der Energieeffizienz der Gebäude auf einer Skala von 1 bis 100, wobei höhere Werte eine bessere Energieeffizienz repräsentieren.

Die Verteilung der Daten weist eine leichte Linksschiefe auf, da die meisten Gebäude im Wertebereich zwischen 60 und 72 liegen und der Median geringfügig über dem Mittelwert liegt. In dem Wertebereich zwischen 80 und 90 ergibt sich eine weitere, weniger ausgeprägte Häufung. Die Randbereiche der Skala mit Werten zwischen 1 und 20 sowie 90 und 100 enthalten nur wenige Objekte, wobei der Wert 1 – als Indikator für eine sehr schlechte Energieeffizienz – überproportional häufig auftritt. Zusammenfassend lässt sich sagen, dass die Mehrheit der Gebäude eine mittlere bis gute Energieeffizienz besitzen.

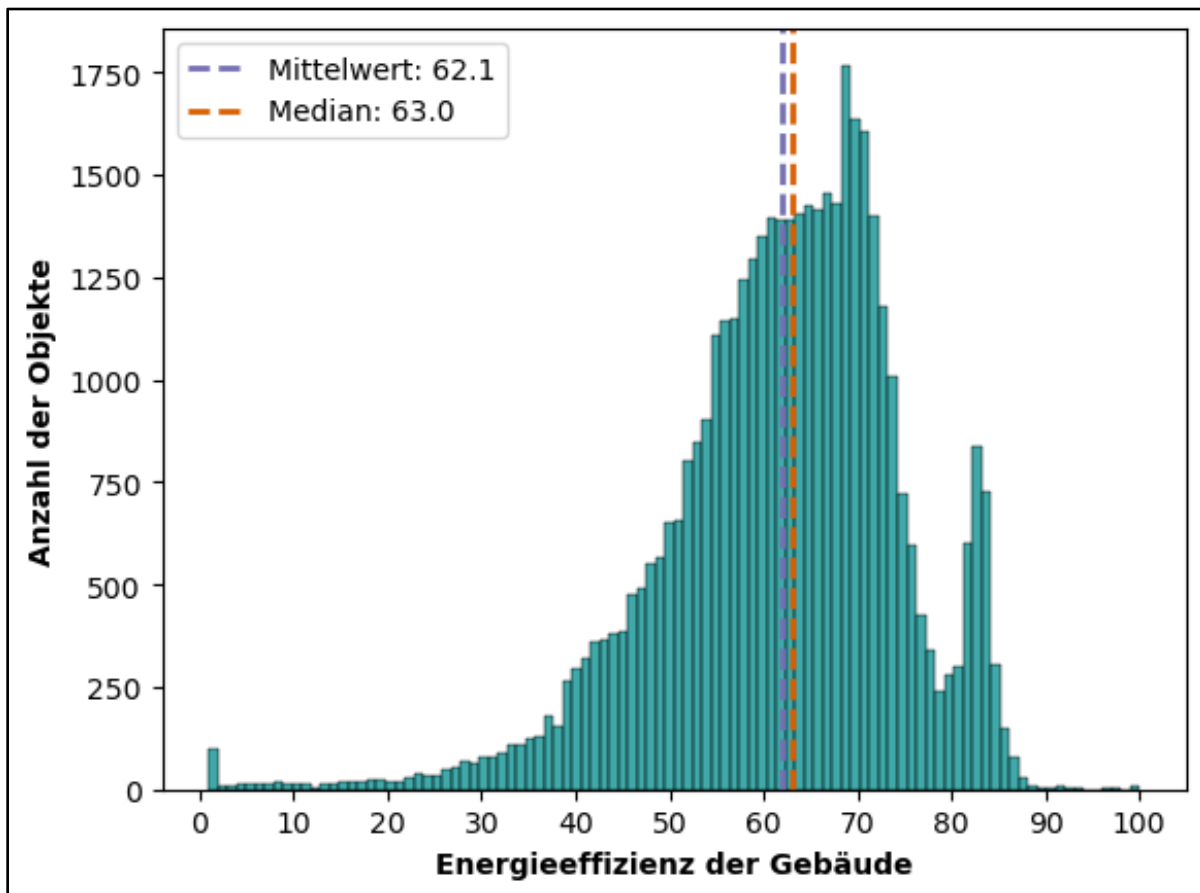


Abbildung 13: Histogramm für die Energieeffizienz der Gebäude

Da die thermischen Infrarotaufnahmen aufgrund der nahezu senkrechten Blickrichtung ausschließlich zur Beurteilung der Gebäudedächer geeignet sind, kommt der Energieeffizienz der Dächer eine besondere Bedeutung zu. Abbildung 14 zeigt die Häufigkeit der Objekte in den fünf Kategorien der Energieeffizienz gemäß den EPCs (*Very Poor*, *Poor*, *Average*, *Good* und *Very Good*).

Mit etwa 47 % entfällt der größte Anteil auf die Kategorie *Good*, die auf eine gute Energieeffizienz der Dachflächen hinweist. Die Kategorien *Very Poor*, *Average* und *Very Good* sind annähernd gleich häufig vertreten, während *Poor* mit nur etwa 5 % den geringsten Anteil bildet.

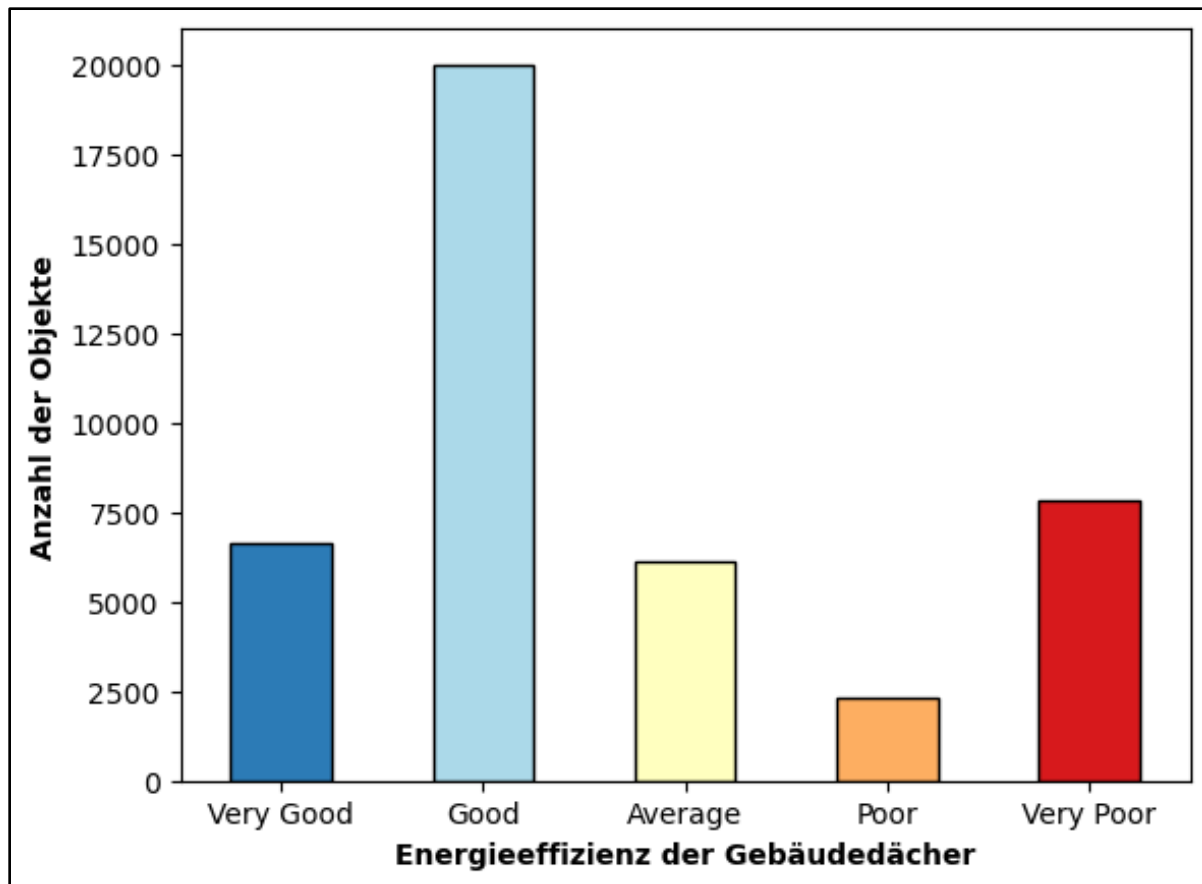


Abbildung 14: Anzahl der Objekte je Energieeffizienzklasse der Gebäudedächer

Sowohl die Gesamtenergieeffizienz der Gebäude als auch die spezifische Bewertung der Dächer ordnen die meisten Objekte mittleren bis guten Energieeffizienzklassen zu. Dennoch weichen die Verteilungen deutlich voneinander ab. Besonders auffällig sind die signifikanten Unterschiede in der Häufigkeit der Objekte in den Randbereichen. Während extrem gute oder schlechte Gesamtwerte (Klassen A oder G) im Datensatz kaum vertreten sind, zeigt die Bewertung der Dächer eine vergleichsweise gleichmäßigere Verteilung auch in den Kategorien *Very Poor* und *Very Good*.

Die wissenschaftlichen Fragestellungen beziehen sich darauf, mit welcher Genauigkeit die Energieeffizienz von Gebäuden mit den zur Verfügung stehenden Thermaldaten modelliert werden kann. Entscheidend dafür ist, ob die Gebäude bei einer Bodenauflösung von 3,5 Metern ausreichend detektierbar und differenzierbar sind. Hierbei ist auch die Gebäudefläche der jeweiligen Objekte maßgebend. Die Heterogenität der Gebäudeflächen stellt eine besondere Herausforderung dar.

Abbildung 15 zeigt die Verteilung der Gebäudeflächen mit einem Säulendiagramm, wobei die Flächen in zehn Wertebereiche unterteilt wurden. Die geometrische Fläche der Gebäudepolygone deckt eine Spanne von 32,6 m² bis 3126,8 m² ab. Der Mittelwert beträgt 75,3 m² und liegt damit oberhalb des Medians von 65,9 m², was auf die Verzerrung durch einige Gebäude mit sehr großen Grundflächen zurückzuführen ist.

Der Großteil der Gebäude weist eine Fläche zwischen 40 m² und 80 m² auf. Mit zunehmender Flächengröße nimmt die Anzahl der Gebäude in den Kategorien erheblich ab.

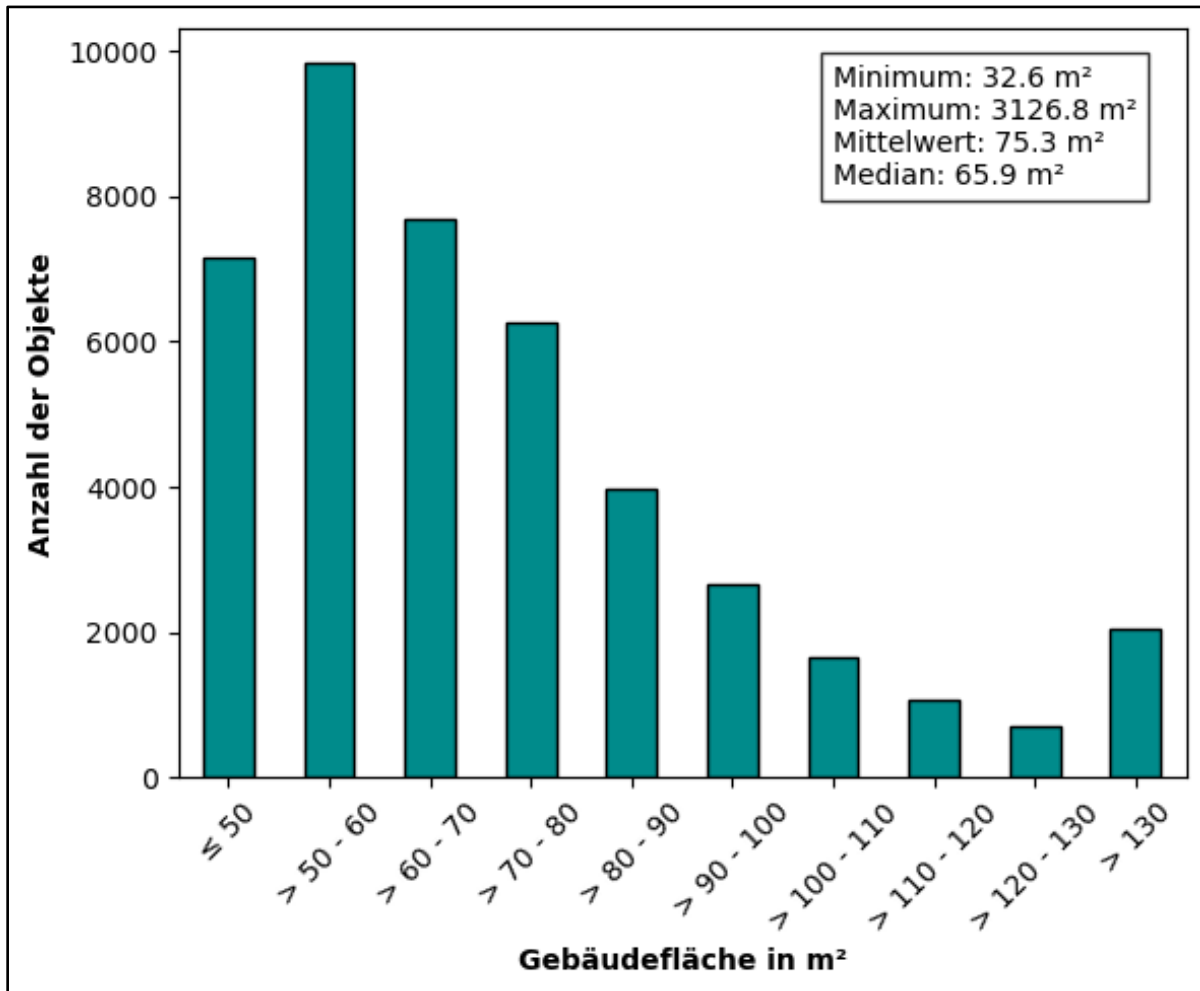


Abbildung 15: Verteilung der Gebäudeflächen in Kategorien

2.3.2 Deskriptive Statistik zu den Land Surface Temperatures der Gebäudedächer

Zur Modellierung der Energieeffizienz der Gebäude werden die in Absatz 2.2 ermittelten mittleren LSTs der Gebäude verwendet. Für jedes Gebäude liegen die LSTs aus beiden thermischen Infrarotaufnahmen vor. Zusätzlich wurde für jedes Gebäude der Mittelwert aus beiden Messungen berechnet.

Zur Beschreibung der Daten wurden deskriptive statistische Kennwerte berechnet und die Verteilungen in Histogrammen visualisiert (siehe Python-Code in Anhang C).

Tabelle 4 bietet eine Übersicht über die deskriptiven statistischen Maße der LSTs. Die Mittelwerte und Mediane der einzelnen Datensätze liegen sehr nahe beieinander, was auf eine nahezu symmetrische Verteilung der Daten hindeutet. Die geringen Spannweiten lassen darauf schließen, dass die LSTs stark konzentriert sind und wenige Ausreißer auftreten. Die Minima, Maxima und Spannweiten der LSTs vom 13.11.2022 sowie der gemittelten LSTs sind

fast identisch. Im Vergleich dazu zeigt sich für die Messung vom 12.11.2022 eine um etwa 1,1 °C größere Spannweite, was vor allem auf ein höheres Maximum zurückzuführen ist.

Die Standardabweichung der mittleren LSTs zeigt mit einem Wert von 0,288 °C die geringste Streuung. Die LSTs sind demnach über beide Messungen sehr stabil.

Tabelle 4: Deskriptive statistische Maße der LST-Werte

Datensatz	LSTs vom 12.11.2022	LSTs vom 13.11.2022	LSTs Mittelwerte beider Messungen
Minimum	9,440	9,650	9,778
Maximum	13,885	12,975	13,096
Spannweite	4,445	3,325	3,318
Mittelwert	12,321	11,239	11,780
Q1 – Unteres Quartil	12,142	11,036	11,610
Median / Q2 – Mittleres Quartil	12,354	11,243	11,789
Q3 – Oberes Quartil	12,546	11,443	11,970
Standardabweichung	0,364	0,305	0,288
Skewness	-1,235	-0,110	-0,427
Kurtosis	7,786	3,195	4,318

Abbildung 16 zeigt das Histogramm der LSTs vom 12.11.2022. Die Verteilung ist linksschief, aber stark um den Mittelwert konzentriert. Die Skewness von -1,24 deutet darauf hin, dass einige wenige Gebäude mit niedrigen LSTs die Verteilung leicht nach links verzerren. Die Kurtosis von 7,79 bestätigt eine ausgeprägte Konzentration der LSTs um den Mittelwert, während gleichzeitig einige Extremwerte auftreten.

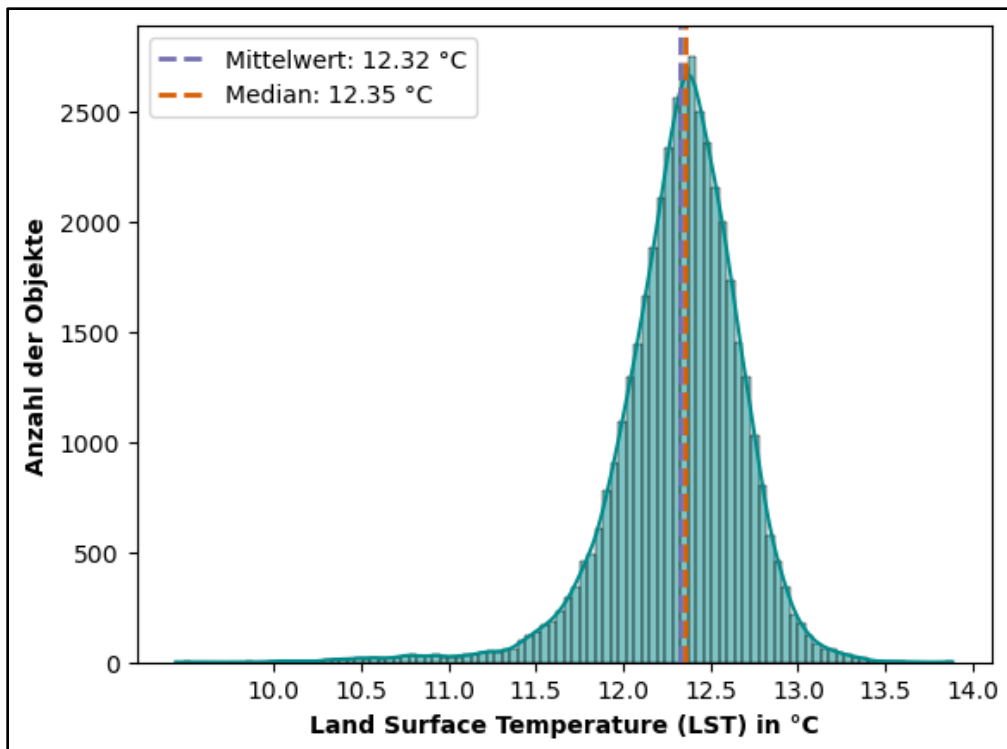


Abbildung 16: Histogramm für die LSTs am 12.11.2022

Das Histogramm der LSTs vom 13.11.2022 (siehe Abbildung 17) deutet auf eine annähernd normalverteilte Datenstruktur hin. Diese Annahme wird durch die Skewness nahe Null sowie die Kurtosis mit einem Wert von 3,20 gestützt.

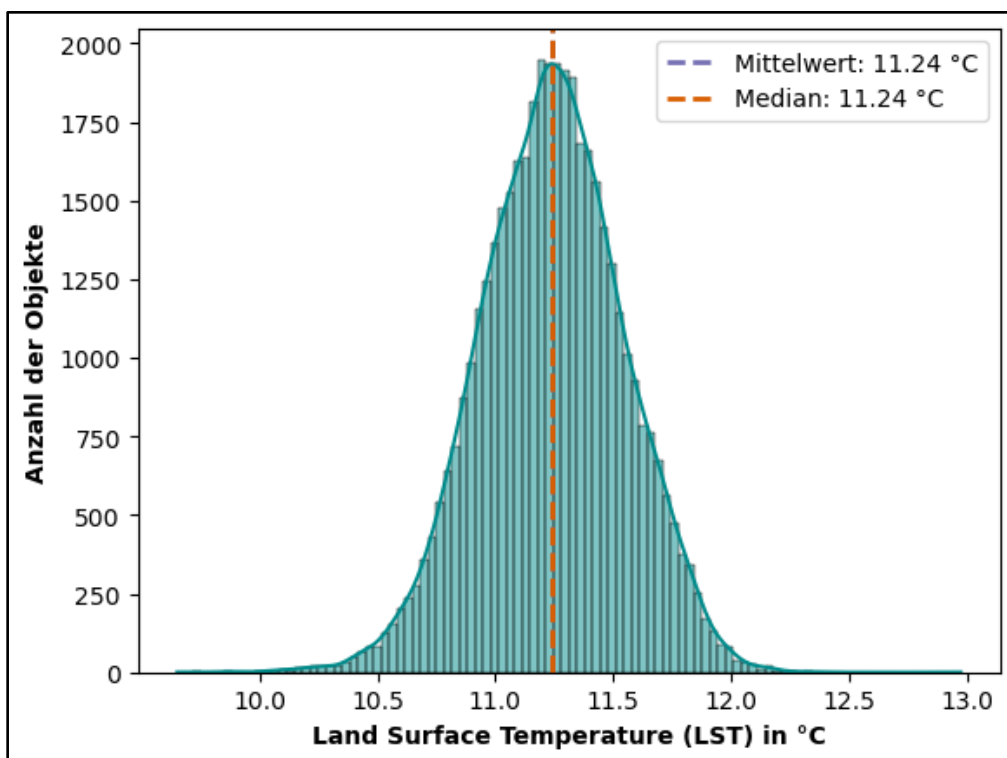


Abbildung 17: Histogramm für die LSTs am 13.11.2022

Abbildung 18 stellt das Histogramm der Mittelwerte der LSTs beider Messungen dar. Das Histogramm zeigt eine gleichmäßige Verteilung mit einer leichten Linksschiefe. Der Wert der Skewness von -0,43 entspricht einer nahezu symmetrischen Verteilung mit einer leichten Verzerrung nach links. Die Kurtosis von 4,32 deutet darauf hin, dass die meisten LSTs um den Mittelwert konzentriert sind, während Ausreißer weniger häufig auftreten.

Die Mittelwerte eignen sich demnach als repräsentative LSTs, da sie eine geringe Verzerrung aufweisen und die Verteilung nahe an einer Normalverteilung liegt.

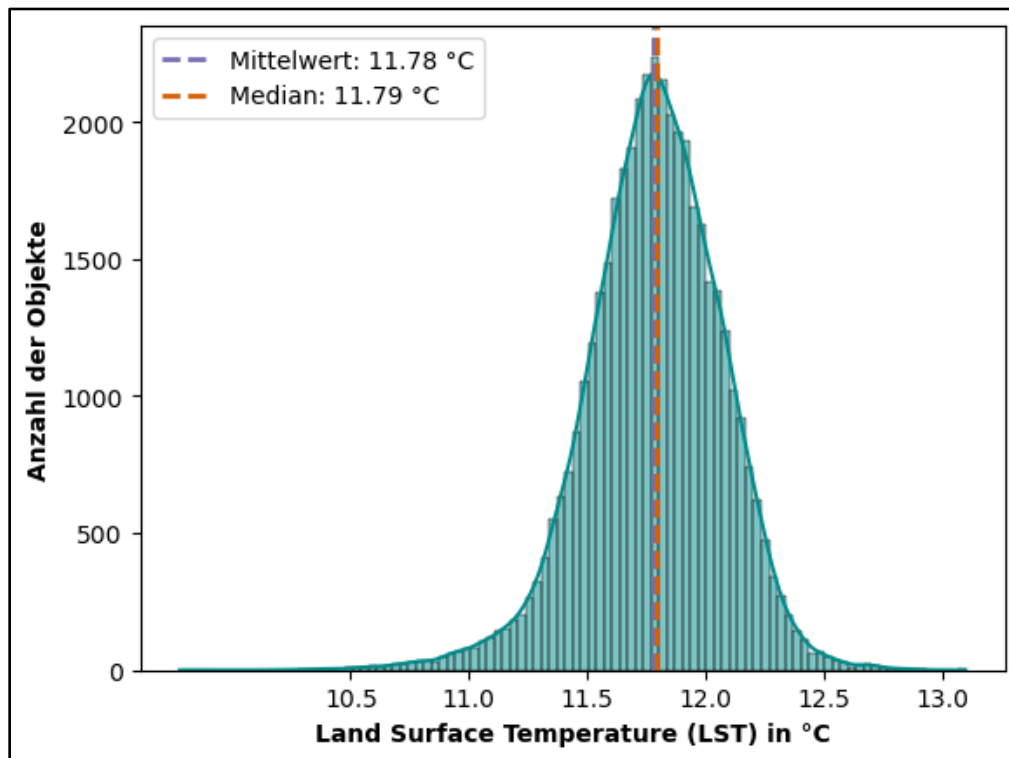


Abbildung 18: Histogramm für die Mittelwerte der LSTs beider thermischer Infrarotmessungen

2.4 Methoden zur Modellierung der Energieeffizienz

2.4.1 Modellierung mittels globaler Regressionsansätze

Die statistische Modellierung der Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten wurde mittels linearer Regression durchgeführt. Die Regression soll die abhängige Variable Energieeffizienz durch die unabhängige Variable LST erklären. Für die unabhängige Variable wurden alle drei LST-Datensätze herangezogen – die Messwerte vom 12.11.2022, vom 13.11.2022 sowie die Mittelwerte beider Aufnahmen.

Die Modellierung erfolgte sowohl für die Gesamtenergieeffizienz der Gebäude als auch separat für die Energieeffizienz der Gebäudedächer. Während für die Energieeffizienz der Gebäude kontinuierliche numerische Werte im Bereich von 1 bis 100 vorliegen, handelt es sich bei der Energieeffizienz der Dächer um ein ordinalskaliertes Merkmal mit fünf Kategorien. Zur Berechnung der Regressionsanalyse wurden die fünf Merkmalsausprägungen in

numerische Werte überführt, indem schlecht isolierte Dächer einen Wert von 1 erhielten und gut gedämmte Dächer einen Wert von 5.

Obwohl die Eingangsdaten zur Dachenergieeffizienz diskrete Werte darstellen, wurde der Zusammenhang mit Hilfe linearer Regressionsmethoden modelliert, da eine feinere Abstufung in Form von Fließkommazahlen als plausibel angesehen werden kann. Die Anwendung eines linearen Regressionsansatzes setzt voraus, dass die Abstände zwischen den Kategorien äquidistant sind. Diese Annahme wird als zutreffend erachtet, da die Energieeffizienzkategorien der Dächer auf Wärmedurchgangskoeffizienten der verwendeten Dämmmaterialien basieren.

Der Zusammenhang zwischen der Energieeffizienz der Gebäude bzw. deren Dächer und den LSTs wurde mittels linearer Regressionsmethoden modelliert, da ein logarithmischer Ansatz keinen signifikanten Vorteil hinsichtlich der Modellgüte zeigte. In einer Studie wurden Zusammenhänge zwischen den jährlichen Energieverbräuchen von Gebäuden pro Flächen in verschiedenen Baublöcken und den LSTs sowohl mit linearen als auch logarithmischen Regressionsmethoden untersucht. Ein logarithmischer Ansatz zeigte in dieser Studie nur geringfügig bessere Modellanpassungen (Jazizadeh & Taleghani, 2016).

Die Berechnung der verschiedenen Regressionen wurde mit Python durchgeführt (siehe Anhang D).

Die linearen Regressionen zeigen die beste Modellanpassung für die Dachenergieeffizienz bei Verwendung der Mittelwerte der LSTs. Auf Grund dieser Ergebnisse wurden für alle folgenden Analysen die Energieeffizienz der Dächer sowie die gemittelten LSTs herangezogen.

Daneben wurden weitere Regressionen mit einem gefilterten Datensatz berechnet, bei dem die Gebäude nach ihrer Flächengröße anhand unterschiedlicher Grenzwerte ausgewählt wurden. Wie in Abschnitt 2.3.1 erläutert, liegt ein Großteil der Gebäudeflächen unterhalb von 60 m². Die Regressionsanalysen sollen Aussagen darüber ermöglichen, ob und inwieweit die Genauigkeit der Modellierung der Energieeffizienz durch thermische Fernerkundungsdaten von der Grundfläche der Gebäude abhängig ist. Die Bodenauflösung von 3,5 Metern könnte insbesondere bei kleinen Gebäuden die Zuverlässigkeit der thermischen Detektion beeinträchtigen. Bei diesen Gebäuden wurde die LST – wie im Verfahren zur Selektion der Rasterzellen je Gebäude in Abschnitt 2.2.3 beschrieben – anhand gewichteter Mittelwerte aus einer sehr geringen Anzahl von Zellen berechnet. In der Mehrheit der Fälle basieren die berechneten LSTs auf nur ein bis zwei BTs. Für die linearen Regressionen wurden Mindestgrößen für die Gebäudeflächen von 50 m², 60 m², 70 m², 80 m², 100 m² und 120 m² definiert. Die Filterung der Gebäudeflächen und die Berechnung der linearen Regressionen mit den ausgewählten Daten wurde mit Python durchgeführt (siehe Anhang D).

2.4.2 Betrachtung verschiedener räumlicher Ebenen

Die Betrachtung verschiedener räumlicher Ebenen soll zeigen, wie sich das Modell bei der Aggregation der Daten auf unterschiedliche räumliche Einheiten verändert. Als räumliche Einheiten wurden sowohl administrative Grenzen als auch modellierte räumliche Einheiten verwendet.

Als administrative Einheiten wurden die *Wards* genutzt, die in England die kleinste verwaltungstechnische Einheit darstellen. Im Untersuchungsgebiet befinden sich lediglich 53 *Wards*, die eine ausreichende Anzahl von Gebäuden umfassen. Eine Aggregation auf größere administrative Einheiten würde zu einer zu geringen Stichprobe führen, sodass keine weitere administrative Einheit zur Aggregation herangezogen wurde.

Als nicht-administrative Einheiten kamen Hexagonraster unterschiedlicher Größe zum Einsatz. Die kleinste gewählte Einheit sind Hexagone mit einer Seitenlänge von 100 m, was einer räumlichen Ebene benachbarter Gebäude entspricht. Die größte verwendete Rastergröße orientiert sich an der durchschnittlichen Fläche eines *Wards* und entspricht einem Hexagon mit einer Seitenlänge von 1026 m. Zusätzlich wurden Zwischenstufen mit Seitenlängen von 200 m, 300 m und 500 m verwendet.

Diese Abstufungen der räumlichen Ebenen sollen ermöglichen, Rückschlüsse darüber zu ziehen, ob und inwieweit die Aggregation der Werte zu einer verbesserten Anpassung des Modells führt.

Abbildung 19 stellt exemplarisch die verschiedenen räumlichen Bezugseinheiten dar, die zur Aggregation der Daten verwendet wurden.

Die Geoobjekte der verschiedenen räumlichen Bezugseinheiten wurden mit den Gebäudepolygonen räumlich verschnitten. Als Kriterium für die Verschneidung wurde festgelegt, dass ein Gebäude vollständig in einer Bezugseinheit enthalten sein muss. Diese Option verhindert redundante Gebäudeobjekte bei der anschließenden Aggregation der Daten. Nach der Geoprozessierung besitzt jedes Gebäudepolygon Informationen darüber, in welcher Bezugseinheit der jeweiligen räumlichen Ebene es verortet ist.

Die linearen Regressionen für die aggregierten Daten wurden mit Python berechnet (siehe Anhang D). Dazu wurde insbesondere die Methode *groupby()* der Pandas-Bibliothek eingesetzt. Für jede räumliche Bezugseinheit wurden Mittelwerte der LSTs sowie der Energieeffizienzwerte der Gebäudedächer ermittelt und als Regressionsvariablen verwendet.

Daten einer räumlichen Bezugseinheit wurden von der Regression ausgeschlossen, wenn die enthaltenen Gebäudepolygone die Bezugseinheit nicht hinreichend repräsentieren. Dies ist der Fall, wenn in einer Bezugseinheit nur sehr wenige Gebäude liegen – häufig an den Rändern des Untersuchungsgebietes. Als maßgebender Parameter für die Einbeziehung in

die Analyse diente die Anzahl der Gebäude innerhalb der jeweiligen Bezugseinheit. Für die Hexagonraster mit einer Seitenlänge von 100 m wurde ein Mindestwert von vier Gebäuden festgelegt. Bei größeren räumlichen Einheiten wurde dieser Grenzwert proportional erhöht.

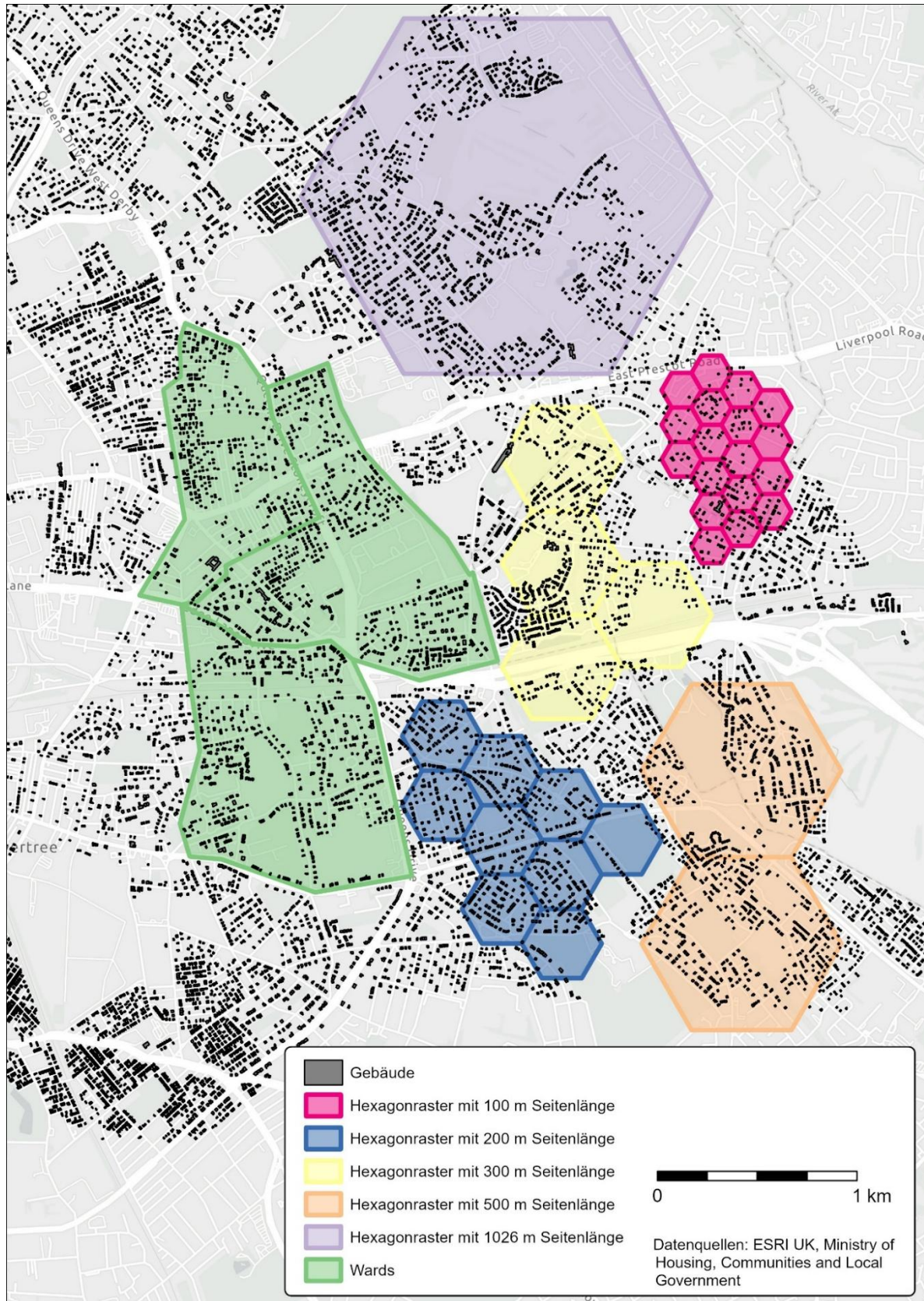


Abbildung 19: Beispielansicht zu den verschiedenen räumlichen Ebenen

2.4.3 Explizite Berücksichtigung räumlicher Effekte

Um zu prüfen, ob die Modelle räumlichen Effekten unterliegen, wurde zunächst eine räumliche Autokorrelationsanalyse durchgeführt. Die Analyse basiert auf den Zentroiden der Gebäudepolygone. Zunächst wurde mittels globaler räumlicher Autokorrelation untersucht, ob in den Residuen der linearen Regression räumlich strukturierte Abhängigkeiten vorliegen. Als Ausgangsmodell wurde die lineare Regression zwischen der Dachenergieeffizienz und der mittleren LST verwendet.

Für die Definition der räumlichen Beziehungen wurde eine Gewichtsmatrix mittels inverser Entfernungen berechnet, sodass nähere Gebäude ein höheres Gewicht erhalten. Als Grenzwert für die Entfernung wurden 100 m definiert. Dieser Wert entspricht annähernd der Ausdehnung der Hexagonraster mit 100 m Seitenlänge und repräsentiert somit das unmittelbare nachbarschaftliche Umfeld eines Gebäudes. Vor Durchführung der Analyse mussten 19 Objekte aus dem Datensatz entfernt werden, da diese Objekte innerhalb des Entfernungsschwellenwertes von 100 m über keine Nachbarn verfügen. Ein weiteres Problem stellt die Aufsplitterung der Gebäude in räumlich getrennte Gruppen dar. Durch die heterogene räumliche Verteilung der Gebäude sowie natürliche Barrieren – insbesondere des Flusses Mersey – ergeben sich Gruppen von Gebäuden, die untereinander nicht verbunden sind. Hierdurch kann keine zusammenhängende räumliche Struktur für das gesamte Untersuchungsgebiet modelliert werden. Um die Stabilität der folgenden räumlichen Modelle zu gewährleisten, wurde die Analyse auf die größte zusammenhängende Gruppe beschränkt. Diese Gruppe stellt den überwiegenden Teil der Gebäude östlich des Flusses Mersey dar und umfasst 23.770 Objekte.

Die Ergebnisse der globalen Autokorrelationsanalysen zeigen eine leicht positive Autokorrelation mit einem Moran's I von 0,166. Der hohe Z-Score von 68,43 sowie ein p-Wert nahe Null deuten auf ein hoch signifikantes räumliches Muster in den Residuen hin. In dem Modell verbleiben räumliche Abhängigkeiten, welche durch die Berechnung eines räumlichen Regressionsmodells berücksichtigt werden sollten.

Die räumliche Abhängigkeit wurde mithilfe eines Spatial Lag Modells abgebildet. Diese Modellierung soll zeigen, ob die Energieeffizienz eines Gebäudes durch die Energieeffizienz benachbarter Gebäude beeinflusst wird. Zusätzlich wurde mit derselben Datengrundlage ein Spatial Error Modell berechnet. Mit diesem Modell kann nachgewiesen werden, ob die räumliche Abhängigkeit ausschließlich im Fehlerterm und nicht in der abhängigen Variable selbst liegt. Als Gewichtsmatrix wurde die gleiche Matrix verwendet, welche auch zur Autokorrelationsanalyse diente. Die räumlichen Regressionen wurden mit den Python-Modulen *libpysal.weights* und *spreg* berechnet (siehe Anhang E).

Im Anschluss an die räumlichen Regressionsmodelle wurde jeweils eine globale räumliche Autokorrelationsanalyse der Residuen berechnet. Ein Vergleich zwischen den Moran's I Werten der globalen linearen Regression und den räumlichen Regressionen gibt Aufschluss darüber, inwieweit die räumliche Struktur durch das jeweilige Modell erklärt wurde.

Darüber hinaus wurde überprüft, ob das berechnete Modell signifikante Cluster oder Ausreißer in den räumlichen Mustern der Residuen zeigt. Hierzu wurde eine lokale räumliche Autokorrelationsanalyse mit den Residuen der linearen Regression zwischen der Dachenergieeffizienz und der mittleren LST für die Hexagonraster mit einer Seitenlänge von 100 m durchgeführt. Die lokale räumliche Autokorrelationsanalyse konnte keine signifikanten Ausreißer oder Cluster feststellen. Insofern wurde auf die Berechnung eines Modells für räumlich variierende Zusammenhänge, z. B. eine geografisch gewichtete Regression, verzichtet

3 Ergebnisse

3.1 Modellierung mittels globaler Regressionsansätze

Die berechneten Regressionsparameter für die verschiedenen linearen Regressionen sind in Tabelle 5 zusammengefasst. Alle Modelle zeigen einen sehr leichten bis leicht negativen Zusammenhang zwischen der Energieeffizienz der Gebäude beziehungsweise der Dächer und den LSTs.

Die Korrelationskoeffizienten und Bestimmtheitsmaße verbessern sich sowohl bei der Modellierung der Dachenergieeffizienz im Vergleich zur gesamten Gebäudeenergieeffizienz als auch bei der Verwendung der gemittelten LSTs im Vergleich zu den LSTs der Einzelmessungen. Während die Modelle für die Energieeffizienz der gesamten Gebäude nur 0,2 % bis 0,5 % der Varianz durch die LSTs abbilden, können die Modelle für die Dachenergieeffizienz zwischen 1,8 % und 2,5 % der Varianz erklären. Die p-Werte nahe Null belegen eine hohe statistische Signifikanz der Zusammenhänge.

Tabelle 5: Parameter der linearen Regressionen zwischen verschiedenen LST-Werten und den Energieeffizienzwerten der Gebäude bzw. Dächer

LST-Datensatz	12.11.2022	13.11.2022	Mittelwerte beider Messungen	12.11.2022	13.11.2022	Mittelwerte beider Messungen
	Gebäude (1 bis 100)	Gebäude (1 bis 100)	Gebäude (1 bis 100)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)
Achsenabschnitt	94,029	83,165	98,847	9,407	9,995	11,870
Steigung	-2,594	-1,877	-3,122	-0,491	-0,591	-0,723
Pearson r	-0,073	-0,044	-0,070	-0,135	-0,136	-0,158
r²	0,005	0,002	0,005	0,018	0,019	0,025
p-Wert	6,19*10 ⁻⁵²	4,61*10 ⁻²⁰	3,05*10 ⁻⁴⁷	4,82*10 ⁻¹⁷⁵	2,67*10 ⁻¹⁷⁷	2,63*10 ⁻²³⁷

Abbildung 20 zeigt die Scatterplots der Modelle zusammen mit den jeweiligen Regressionsgeraden. Auf der Y-Achse ist die abhängige Variable Energieeffizienz und auf der X-Achse die unabhängige Variable LST dargestellt.

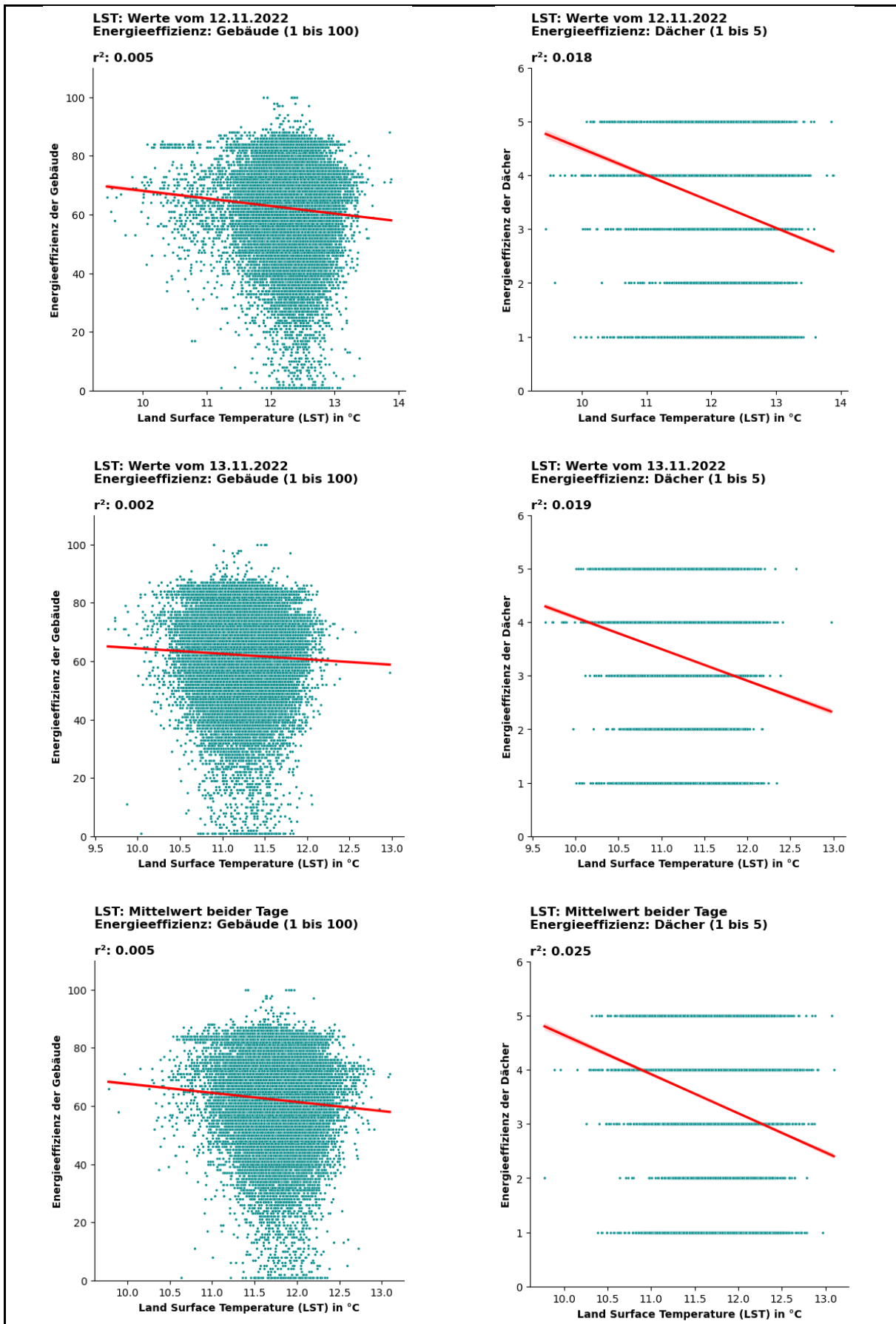


Abbildung 20: Vergleich mehrerer linearer Regresstionen zwischen verschiedenen LST-Werten und den Energieeffizienzwerten der Gebäude bzw. Dächer

Die berechneten Regressionsparameter für die verschiedenen linearen Regressionen mit einer Filterung der Gebäude nach deren Grundflächen sind in Tabelle 6 zusammengefasst. Alle Regressionen zeigen einen leicht negativen Zusammenhang zwischen der Energieeffizienz der Gebäudedächer und den mittleren LSTs. Die Modelle erklären lediglich 3,3 % bis 5,1 % der Varianz der Energieeffizienz durch die LSTs. Die p-Werte nahe Null deuten auf eine hohe Signifikanz der Ergebnisse hin.

Die Filterung nach der Gebäudefläche führt bei höheren Grenzwerten zu einer verbesserten Modellanpassung – mit Ausnahme der Analyse von Gebäuden mit einer Fläche über 120 m². Das Bestimmtheitsmaß steigt kontinuierlich bei Grenzwerten von 50 m² über 60 m² hin zu 70 m² an. Aber einer Flächengröße von 70 m² bleibt die erklärte Varianz konstant, während sie bei einem Grenzwert von 120 m² wieder abnimmt. Mit zunehmendem Grenzwert reduziert sich gleichzeitig auch die Größe der Stichprobe.

Tabelle 6: Parameter der linearen Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Filterung der Gebäude nach deren Flächen

LST-Datensatz	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen
Energieeffizienz-Datensatz	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)
Bedingung Gebäudeflächen	>50 m ²	>60 m ²	>70 m ²	>80 m ²	>100 m ²	>120 m ²
Objektanzahl	35849	26028	18337	12069	5447	2744
Achsenabschnitt	13,578	15,089	16,473	16,612	16,424	14,649
Steigung	-0,872	-1,006	-1,127	-1,143	-1,131	-0,980
Pearson r	-0,182	-0,205	-0,226	-0,226	-0,225	-0,197
r²	0,033	0,042	0,051	0,051	0,051	0,039
p-Wert	2,19*10 ⁻²⁶⁵	8,80*10 ⁻²⁴⁴	9,21*10 ⁻²¹¹	5,63*10 ⁻¹⁴⁰	1,31*10 ⁻⁶³	2,18*10 ⁻²⁵

Abbildung 21 zeigt die Scatterplots der Modelle mit nach Gebäudeflächen gefilterten Eingangsdaten, jeweils gemeinsam mit den zugehörigen Regressionsgeraden. Auf der Y-Achse ist die abhängige Variable Energieeffizienz und auf der X-Achse die unabhängige Variable LST dargestellt.

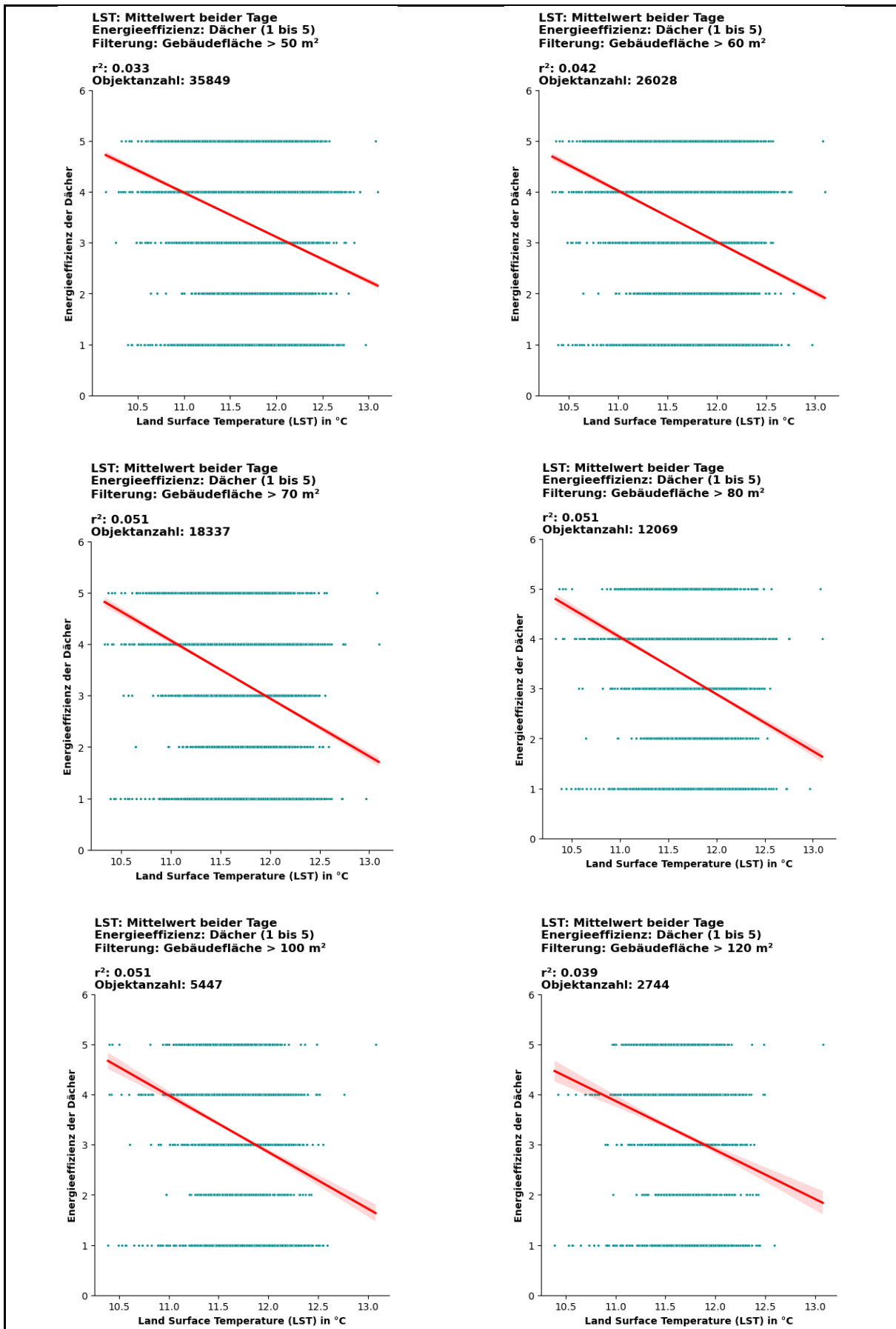


Abbildung 21: Vergleich mehrerer linearer Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Filterung der Gebäude nach deren Flächen

3.2 Betrachtung verschiedener räumlicher Ebenen

Tabelle 7 stellt die berechneten Regressionsparameter für die unterschiedlichen linearen Regressionen bei Betrachtung verschiedener räumlicher Ebenen in einer Übersicht dar. Alle Regressionen zeigen einen leichten negativen Zusammenhang zwischen der Energieeffizienz der Gebäudedächer und den mittleren LSTs. Die Modelle für die aggregierten Energieeffizienzwerte erklären nur 2,7 % bis 6,1 % der Varianz durch die LSTs.

Die Ergebnisse für Hexagonraster mit Seitenlängen von 100 m bis 500 m sind statistisch signifikant. Für Hexagonraster mit einer Seitenlänge von 1026 m sowie für die *Wards* sind die p-Werte über 0,05, wodurch diese Modelle nicht statistisch signifikant sind. Der Anteil der erklärten Varianz steigt mit zunehmender Größe der Hexagonraster bis zu einer Seitenlänge von 500 m. Die lineare Regression mit den aggregierten Werten für die *Wards* zeigt hingegen keine signifikante Verbesserung im Vergleich zur Modellierung auf Gebäudeebene. Mit zunehmender Größe der räumlichen Bezugseinheiten reduziert sich gleichzeitig auch die Größe der Stichprobe.

Tabelle 7: Parameter der linearen Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Aggregation der Werte auf verschiedene räumliche Ebenen

LST-Datensatz	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen	Mittelwerte beider Messungen
Energieeffizienz-Datensatz	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)	Dach (1 bis 5)
Räumliche Ebene	Hexagon (100 m)	Hexagon (200 m)	Hexagon (300 m)	Hexagon (500 m)	Hexagon (1026 m)	Wards
Objektanzahl	2179	685	332	143	46	53
Achsenabschnitt	7,650	7,001	7,320	8,324	7,181	7,007
Steigung	-0,365	-0,309	-0,333	-0,417	-0,317	-0,307
Pearson r	-0,164	-0,175	-0,196	-0,247	-0,236	-0,179
r ²	0,027	0,031	0,038	0,061	0,056	0,032
p-Wert	1,26*10 ⁻¹⁴	4*10 ⁻⁶	3,25*10 ⁻⁴	0,003	0,115	0,199

Abbildung 22 zeigt die Scatterplots der Modelle für die verschiedenen räumlichen Ebenen zusammen mit den zugehörigen Regressionsgeraden. Auf der Y-Achse ist die abhängige Variable Energieeffizienz und auf der X-Achse die unabhängige Variable LST dargestellt.

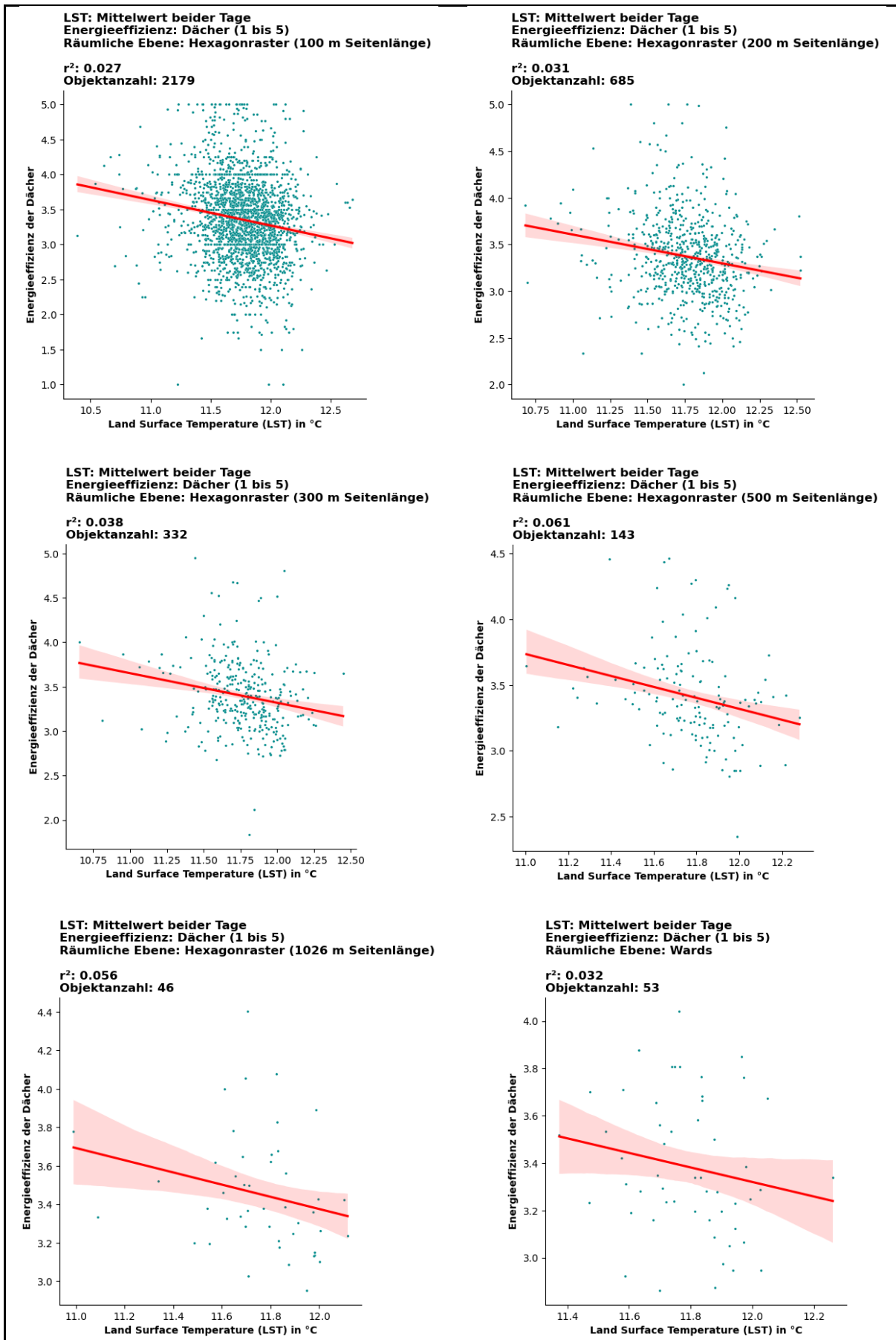


Abbildung 22: Vergleich mehrerer linearer Regressionen zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer bei Aggregation der Werte auf verschiedene räumliche Ebenen

3.3 Explizite Berücksichtigung räumlicher Effekte

Tabelle 8 fasst die berechneten Regressionsparameter der räumlichen Modelle Spatial Lag und Spatial Error in einer Übersicht zusammen. Beide Modelle zeigen einen negativen Zusammenhang zwischen der LST und der Energieeffizienz der Dächer. Der Anteil der erklärten Varianz des Spatial Lag Modells zeigt gegenüber der globalen linearen Regression eine deutliche Verbesserung. Die p-Werte nahe Null deuten auf eine hohe Signifikanz der Ergebnisse hin. Der räumliche Parameter ρ hat einen Wert von -2,536 und liegt damit außerhalb des zulässigen Bereichs (-1 bis 1), was auf ein Modellproblem hinweist. Infolgedessen verschlechtert sich die räumliche Struktur, was sich in einem deutlich erhöhten Moran's I Wert der Residuen zeigt. Das Spatial Error Modell erreicht mit einem Pseudo r^2 von 0,025 lediglich die Modellgüte der globalen linearen Regression, ohne diese zu verbessern. Der hohe Wert für λ deutet darauf hin, dass im Modell räumlich strukturierte Einflussgrößen fehlen. Obwohl λ signifikant ist, kann das räumliche Modell keine relevante Verbesserung erzielen.

Tabelle 8: Parameter der räumlichen Regressionsmodelle Spatial Lag und Spatial Error zwischen den mittleren LST-Werten und den Energieeffizienzwerten der Gebäudedächer

Modell	Spatial Lag Modell	Spatial Error Modell
LST-Datensatz	Mittelwerte beider Messungen	Mittelwerte beider Messungen
Energieeffizienz-Datensatz	Dach (1 bis 5)	Dach (1 bis 5)
Objektanzahl	23.770	23.770
Achsenabschnitt	30,788	19,717
Steigung	-1,594	-1,383
Pseudo r^2	0,091	0,025
p-Wert	$2,01 \cdot 10^{-68}$	$2,88 \cdot 10^{-178}$
ρ / λ	-2,536	0,598
p-Wert ρ / λ	$6,612 \cdot 10^{-27}$	$2,88 \cdot 10^{-178}$
Moran's I der Residuen	0,695	0,167

4 Diskussion

Die negativen Zusammenhänge zwischen der Energieeffizienz der Gebäude beziehungsweise der Gebäudedächer und den LSTs bestätigen die Hypothese, dass schlecht isolierte Dächer mehr Wärme abgeben und daher höhere LSTs aufweisen.

Die sehr geringen Bestimmtheitsmaße der Modelle deuten darauf hin, dass die LSTs nur einen minimalen Anteil der Variation der Energieeffizienz erklären. Trotz der statistischen Signifikanz der Ergebnisse, die insbesondere aufgrund der großen Stichprobe auftritt, erlauben die beobachteten Thermaldaten keine Rückschlüsse auf die Energieeffizienz der Gebäude.

Die Modelle zur Erklärung der Energieeffizienz der Gebäudedächer zeigen geringfügig bessere Anpassungen. Dies ist nachvollziehbar, da die thermischen Infrarotaufnahmen aus einer senkrechten Blickrichtung ausschließlich die LSTs der Gebäudedächer erfassen. Im Vergleich dazu ist die Energieeffizienz der gesamten Gebäude von diversen Faktoren abhängig. Neben dem Wärmedurchgangskoeffizienten des Daches wird die Energieeffizienz auch von den Außenwänden, den Fenstern und Türen, der Bodenplatte sowie der technischen Gebäudeausrüstung wie Heizung, Lüftung und Warmwasseraufbereitung beeinflusst.

Die geringe Standardabweichung der LSTs aus beiden Messungen zeigt, dass die Werte über beide Messungen hinweg stabil sind. Der Mittelwert der LSTs aus beiden Messungen verringert den Einfluss von Ausreißern, die in einer Einzelmessung auftreten können, und führt infolgedessen zu besseren Modellergebnissen.

Viele Gebäuden im Untersuchungsgebiet weisen eine geringe Grundfläche auf. Dies führt dazu, dass der Geoverarbeitungsprozess zur Auswahl der relevanten Rasterzellen je Gebäude in der überwiegenden Zahl der Fälle nur eine oder zwei Rasterzellen pro Gebäude selektiert. Diese Tatsache kann die Zuverlässigkeit der LSTs beeinträchtigen. Die linearen Regressionen, bei denen die Gebäude nach ihrer Grundfläche gefiltert wurden, weisen in den meisten Fällen eine bessere Modellanpassung auf, wenn Gebäude mit kleinen Flächen ausgeschlossen werden. Für den vorliegenden Datenbestand in einem urbanen Gebiet scheint eine Bodenauflösung von 3,5 m demnach nicht optimal zu sein. Dahingegen haben verschiedenste Interessenvertreter im Rahmen einer Studie eine Bodenauflösung von 3 m bis 6 m als notwendige Auflösung für thermische Aufnahmen erachtet, um diese in Bezug zur Energieeffizienz von Gebäuden bringen zu können (Ben et al., 2021). Es ist zu beachten, dass Wohngebäude mit geringer Grundfläche üblicherweise Reihenhäuser sind. Reihenhäuser sind im Normalfall nicht quadratisch, sondern länglich gebaut, wodurch ihre Breite entsprechend gering ausfällt. Trotz dieser Überlegungen bleibt das Bestimmtheitsmaß für die Modelle, selbst bei der exklusiven Verwendung von Gebäuden mit einer Fläche von mehr als 70 m², sehr gering.

Die Betrachtung verschiedener räumlicher Ebenen hat gezeigt, dass die Modelle bei Aggregation auf größere räumliche Bezugseinheiten bessere Anpassungen an die Daten liefern. Die beste Modellanpassung wurde bei der Aggregation auf Hexagonraster mit einer Seitenlänge von 500 m erzielt, mit einem r^2 Wert von 0,061. Diese Verbesserung der erklärten Varianz impliziert, dass die Bodenauflösung von 3,5 m für die Gebäude in dem Untersuchungsgebiet nicht ausreichend ist. In der Forschungsarbeit von Jazizadeh & Taleghani (2016) wurden Jahresenergieverbräuche von Gebäuden pro Flächeneinheit in verschiedenen Baublöcken genutzt, welche stark mit den LSTs korrelieren ($r^2 = 0,63$). Insofern ist fraglich, warum die Aggregation auf größere Bezugseinheiten in dieser Arbeit nicht zu ähnlichen Ergebnissen führen.

Die geringen Modellgüten erschweren die explizite Berücksichtigung räumlicher Effekte. Das Spatial Lag Modell erweist sich als instabil, da der geschätzte Wert für den räumlichen Parameter ρ außerhalb des zulässigen Bereichs liegt. Das Spatial Error Modell liefert trotz signifikanter Autokorrelation keine Verbesserung gegenüber der globalen Regression. Diese Ergebnisse können einerseits mit den Problemen zusammenhängen, die bereits in den globalen Regressionen zu den geringen Anteilen der erklärten Varianz führen. Andererseits ist fraglich, ob ein Spillover-Effekt tatsächlich vorliegt. Ein solcher Effekt erscheint bei unmittelbar angrenzenden Gebäuden plausibel, da ein Gebäude weniger Wärme an ein energieeffizientes Nachbargebäude abgibt. Dieser Effekt ist in Anbetracht der vielen Einflussfaktoren auf die Energieeffizienz eines Gebäudes zu vernachlässigen. Die räumlichen Muster lassen sich vielmehr durch strukturelle Rahmenbedingungen erklären. Die Entwicklung von Bauland erfolgt üblicherweise für größere abgegrenzte Gebiete. Diese Baugebiete werden im Anschluss innerhalb weniger Jahre bebaut, wodurch benachbarte Gebäude meist ein ähnliches Baujahr aufweisen. Das Baujahr wiederum stellt einen entscheidenden Faktor für die Energieeffizienz eines Gebäudes dar. Zudem führen örtliche Bauvorschriften sowie zeitlich befristete Förderprogramme für energieeffizientes Bauen zu gleichen Energieeffizienzklassen der Gebäude innerhalb eines Gebietes. Die LSTs von benachbarten Gebäuden unterliegen ähnlichen Einflüssen, da sie denselben mikroklimatischen Bedingungen ausgesetzt sind und die umgebende Topografie homogen ist.

In den Ausgangsdaten ergeben sich insbesondere drei wesentliche Probleme bei der Modellierung der Energieeffizienz von Gebäuden anhand der LSTs:

1. Die thermischen Infrarotaufnahmen von SatelliteVU dienen nicht der Untersuchung des Zusammenhangs zwischen der Energieeffizienz von Gebäuden und den Thermaldaten. Vielmehr sollten die Aufnahmen der Wirtschaft und Wissenschaft ermöglichen, die Eignung der abgeleiteten Bodenauflösung für verschiedene Anwendungsbereiche – etwa das Monitoring von Infrastrukturen und Industrieanlagen oder für Klimaanalysen – zu bewerten. SatelliteVU ist ein Satellitenunternehmen, welches sich unter anderem auf

hochaufgelöste thermische Infrarotaufnahmen mittels Kleinstsatelliten spezialisiert hat. Die in dieser Arbeit verwendeten flugzeuggestützten Aufnahmen waren als Testmessungen konzipiert, um die Möglichkeiten zukünftiger Produkte aufzuzeigen. Aufgrund dessen waren die äußeren Bedingungen zum Zeitpunkt der Aufnahmen suboptimal für die Zwecke dieser Untersuchung. Obwohl die Daten im November und damit innerhalb der Heizperiode erhoben wurden, lagen die Außentemperaturen mit 10 °C bis 13 °C deutlich über dem von K. Mayer et al. (2023, S. 6) genannten Grenzwert von 5 °C, der erforderlich ist, um Wärmeverluste zuverlässig zu detektieren. Bitelli et al. (2015, S. 2157) bekräftigen diese Aussage, indem sie eine kalte Winternacht als ideale Bedingung zur Maximierung der Temperaturdifferenzen zwischen Gebäudedächern und der Umgebung nennen. Die Spannweiten der mittleren LSTs je Gebäude betragen nur 3,3 °C, wobei die meisten Werte im Bereich zwischen 11 °C und 12,5 °C liegen und somit nur eine Varianz von 1,5 °C aufweisen. Solche geringen Temperaturunterschiede erschweren die Modellierung. Eine Studie aus Turin zeigt im Vergleich dazu bereits zwischen zwei benachbarten Energieeffizienzklassen Temperaturdifferenzen von 3,55 °C (Anselmo et al., 2023, S. 7). Ein weiterer Aspekt ist der Aufnahmezeitpunkt am Abend. Die kälteste Tageszeit ist die Morgendämmerung. Zu diesem Zeitpunkt lassen sich Wärmeverluste deutlich besser erkennen, da der Einfluss durch Sonnenerwärmung nahezu ausgeschlossen werden kann. Im Verlauf des Tages erwärmen sich die Oberflächen unterschiedlich stark durch die Sonneneinstrahlung – bedingt durch die umgebende Topographie. Dieser Effekt kann bei einer Messung am Abend noch wirksam sein (Kuenzer et al., 2015, S. 268).

2. Die Berechnung der LST aus der BT beruht auf einer Vielzahl von Annahmen, da das Material der Dachoberflächen unbekannt ist. Wie in Abschnitt 2.2.4 erläutert, ist die Berücksichtigung der Emissionsgrade ein entscheidender Faktor für die Genauigkeit der berechneten LSTs. Die Emissionsgrade wurden auf Grundlage der verfügbaren Daten bestmöglich geschätzt. Diese Schätzung stellt eine potenzielle Fehlerquelle dar. Laut Chen et al. (2017, S. 83) ist eine Genauigkeit von 0,01 für den Emissionsgrad erforderlich, um zuverlässige LSTs ableiten zu können. Bereits ein Fehler von 0,01 führt bei den verwendeten Thermaldaten zu einem LST-Fehler von ca. 0,26 °C. Angesichts der geringen Varianz der LSTs ist ein solcher Fehler von großer Relevanz und kann die Modellierung signifikant beeinträchtigen.
3. Die Energieeffizienz der Gebäude ist nicht gleich mit der Energieeffizienz der Dächer. Auch wenn hochwertige Dächer häufig mit energieeffizienten Gebäuden korrelieren, gilt dies nicht zwangsläufig auf der Ebene einzelner Gebäude (Bitelli et al., 2015, S. 2156). Zudem liegen für die Dächer nur ordinalskalierte Effizienzwerte in fünf Kategorien vor, wohingegen für die Energieeffizienz der Gesamtgebäude kontinuierliche und deutlich differenzierte Werte verfügbar wären. Aus den zuvor ausgeführten Gründen und auf Basis der

Ergebnisse der ersten Regressionsanalysen wurden dennoch die Werte für die Dachenergieeffizienz verwendet. Die Umwandlung der ordinalen Werte in numerische Werte erfolgte unter der Annahme gleicher Abstände zwischen den Kategorien. Die Aggregation der Werte bei der Betrachtung des Modells im Hinblick auf verschiedene räumliche Ebenen liefert Fließkommazahlen für die Dachenergieeffizienz und verbessert damit die Kardinalität der Skala. Der einheitliche linearere Regressionsansatz gewährleistet zudem die Vergleichbarkeit der verschiedenen Modelle. Trotz dessen bleibt die ursprüngliche grobe Einteilung eine Einschränkung, da jede Kategorie eine breite Spanne an Wärmedurchgangskoeffizienten umfasst.

Neben den drei Hauptfehlerquellen wird die Dachoberflächentemperatur durch zahlreiche weitere Parameter beeinflusst. Hierzu zählen die Oberflächenfeuchtigkeit, die Wärmeleitfähigkeit, die Wärmebilanz der Außenhülle, die Belüftung des Dachraums, die umgebende Topografie, das Mikroklima sowie lokale klimatische Schwankungen (Bitelli et al., 2015, S. 2156; Hemachandran et al., 2018, S. 242; Voogt & Oke, 2003, S. 371). Darüber hinaus unterliegt die Modellierung weiteren Unsicherheiten, etwa durch unbewohnte oder unbeheizte Gebäude zum Aufnahmezeitpunkt, nicht deckungsgleiche Datensätze zwischen den Gebäudepolygonen und den Thermaldaten sowie Qualitätsmängel in den zugrunde liegenden Energieausweisdaten.

5 Schlussfolgerungen und Ausblick

Die Modellierung der Energieeffizienz von Gebäuden mit thermischen Fernerkundungsdaten mit einer Bodenauflösung von 3,5 m konnte trotz einer umfangreichen Stichprobe von etwa 43.000 Energieausweisdaten keinen belastbaren statistischen Zusammenhang aufzeigen.

Die Modellierung unterlag jedoch mehreren potenziellen Fehlerquellen. Zum Zeitpunkt der Aufnahmen lagen die Lufttemperaturen mit 10 °C bis 13 °C deutlich über dem für zuverlässige Detektionen empfohlenen Schwellenwert von 5 °C, wodurch die Erfassung der Abwärme von Gebäuden erheblich erschwert wurde. Darüber hinaus ist das dicht bebaute Untersuchungsgebiet überwiegend durch kleine Reihenhausstrukturen geprägt, sodass die räumliche Auflösung der Thermaldaten die Gebäude unzureichend abbildet. Zusätzlich fehlten wichtige Informationen zu den Dachmaterialien sowie zu den Oberflächentemperaturen der Fassaden, was die Modellierung weiter einschränkt.

Trotz diverser Fehlerquellen in den zugrunde liegenden Daten und der daraus resultierenden geringen Modellgüte, zeigen alle berechneten Regressionen einen leichten negativen Trend zwischen der Energieeffizienz der Gebäude beziehungsweise der Gebäudedächer und den LSTs. Dies bestätigt die Hypothese, dass die Abwärme von Gebäuden mittels thermischer Infrarotmessungen prinzipiell erfassbar ist und der Wärmefluss von innen nach außen auf den Wärmebildern sichtbar wird. Zudem konnten gewisse Effekte – wenn auch in begrenztem Umfang – nachgewiesen werden, indem die Filterung der Eingangsdaten nach den Gebäudeflächen oder die Betrachtung unterschiedlicher räumlicher Ebenen zu verbesserten Modellanpassungen führten. Aufgrund der unzureichenden Qualität der verfügbaren Daten kann keine belastbare Aussage zur generellen Genauigkeit der Modellierung der Energieeffizienz von Gebäuden auf Basis thermischer Fernerkundungsdaten mit einer Bodenauflösung von 3,5 m getroffen werden. Lediglich für die vorliegenden Thermaldaten lässt sich festhalten, dass kein ausreichender statistischer Zusammenhang festgestellt werden konnte. Die beobachteten räumlichen Muster konnten mit den vorliegenden Daten nicht ausreichend modelliert werden.

Um zuverlässige Ergebnisse im Hinblick auf die wissenschaftlichen Fragestellungen dieser Arbeit zu erhalten, wäre ein Thermaldatensatz mit entsprechender Bodenauflösung notwendig, der gezielt für die Modellierung der Gebäudeenergieeffizienz erhoben wird. Die thermischen Infrarotaufnahmen sollten in der Morgendämmerung an einem kalten Wintertag erfolgen. Des Weiteren sind präzise Informationen über die Materialien der Dachoberflächen essenziell, um die Berechnung der LSTs aus BTs mit hoher Genauigkeit durchführen zu können. Solche Informationen können auch mittels hochauflösender Hyperspektraldaten abgeleitet werden. Eine auf diese Arbeit aufbauende Forschung, bei der die zuvor genannten Daten für eine ausreichend große Stichprobe vorhanden sind, könnte die hier

zugrundeliegenden Fragestellungen verlässlich beantworten. Eine solche Untersuchung hätte das Potenzial, großflächige Aussagen zur Energieeffizienz zu liefern und die Einsatzmöglichkeiten thermischer Fernerkundungsdaten für die Energieplanung und strategische Entscheidungsfindung zu erweitern.

Die in dieser Arbeit gewonnenen Erkenntnisse bilden eine methodische Grundlage für weiterführende Studien, da sie sowohl übertragbare Analyseverfahren als auch zentrale Herausforderungen der Modellierung aufzeigt.

Literaturverzeichnis

- Ali, U., Shamsi, M. H., Bohacek, M., Purcell, K., Hoare, C., Mangina, E., & O'Donnell, J. (2020). A data-driven approach for multi-scale GIS-based building energy modeling for analysis, planning and support decision making. *Applied Energy*, 279, 115834. <https://doi.org/10.1016/j.apenergy.2020.115834>
- Anselmo, S., Ferrara, M., Corgnati, S. P., & Boccardo, P. (2023). Aerial urban observation to enhance energy assessment and planning towards climate-neutrality: A pilot application to the city of Turin. *Sustainable Cities and Society*, 99, 104938. <https://doi.org/10.1016/j.scs.2023.104938>
- Ben, H., Mackie, E., Parry, I., Shuckburgh, E., Hawker, G., & Jenkins, M. G. (2021). Targeting building energy efficiency using thermal infrared earth observation telescopes. *Journal of Physics: Conference Series*, 2042, 12001. <https://doi.org/10.1088/1742-6596/2042/1/012001>
- Bitelli, G., Conte, P., Csoknyai, T., Franci, F., Girelli, V. A., & Mandanici, E. (2015). Aerial Thermography for Energetic Modelling of Cities. *Remote Sensing*, 7(2), 2152–2170. <https://doi.org/10.3390/rs70202152>
- Boomsma, C., Goodhew, J., Goodhew, S., & Pahl, S. (2016). Improving the visibility of energy use in home heating in England: Thermal images and the role of visual tailoring. *Energy Research & Social Science*, 14, 111–121. <https://doi.org/10.1016/j.erss.2016.01.005>
- Chen, F., Yang, S., Yin, K., & Chan, P. (2017). Challenges to quantitative applications of Landsat observations for the urban thermal environment. *Journal of Environmental Sciences*, 59, 80–88. <https://doi.org/10.1016/j.jes.2017.02.009>
- Coutts, A. M., Harris, R. J., Phan, T., Livesley, S. J., Williams, N. S. G., & Tapper, N. J. (2016). Thermal infrared remote sensing of urban heat: Hotspots, vegetation, and an assessment of techniques for use in urban planning. *Remote Sensing of Environment*, 186, 637–651. <https://doi.org/10.1016/j.rse.2016.09.007>

- European Commission. (2020). In Focus: Energy Efficiency in Buildings. Zugriff am 23.01.2024. Verfügbar unter: https://commission.europa.eu/news/focus-energy-efficiency-buildings-2020-02-17_en
- European Parliament. (2010). Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast). *Official Journal of the European Union*, L153/13.
- Gulbe, L., Caune, V., & Korats, G. (2017). Urban area thermal monitoring: Liepaja case study using satellite and aerial thermal data. *International Journal of Applied Earth Observation and Geoinformation*, 63, 45–54. <https://doi.org/10.1016/j.jag.2017.07.005>
- Hemachandran, B., Geoffrey, J. H., Rahman, M. M., Couloigner, I., Zhang, Y., Karim, B., Fung, T. S., & Kyle, C. D. (2018). Developing Multiscale HEAT Scores from H-Res Airborne Thermal Infrared Imagery to Support Urban Energy Efficiency: Challenges Moving Forward. In *Urban Remote Sensing* (2. Aufl., S. 235–270). CRC Press.
- Jazizadeh, F., & Taleghani, M. (2016). Towards urban facilities energy performance evaluation using remote sensing. *Procedia Engineering*, 145, 916–923. <https://doi.org/10.1016/j.proeng.2016.04.119>
- Kamaraj, N. P., Shekhar, S., Sivashankari, V., Balasubramani, K., & Prasad, K. A. (2021). Detecting heat-inducing urban built-up surface material with multi remote sensing datasets using reflectance and emission spectroscopy. *Remote Sensing of Environment*, 264, 112591. <https://doi.org/10.1016/j.rse.2021.112591>
- Kotthaus, S., Smith, T. E. L., Wooster, M. J., & Grimmond, C. S. B. (2014). Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94, 194–212. <https://doi.org/10.1016/j.isprsjprs.2014.05.005>
- Kuenzer, C., Zhang, J., & Dech, S. (2015). Thermal Infrared Remote Sensing: Principles and Theoretical Background. In *Remotely Sensed Data Characterization, Classification, and Accuracies*. CRC Press.

- Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., & Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, 131, 14–37. <https://doi.org/10.1016/j.rse.2012.12.008>
- Liu, K., Su, H., Zhang, L., Yang, H., Zhang, R., & Li, X. (2015). Analysis of the Urban Heat Island Effect in Shijiazhuang, China Using Satellite and Airborne Data. *Remote Sensing*, 7(4), 4804–4833. <https://doi.org/10.3390/rs70404804>
- López-González, L. M., López-Ochoa, L. M., Las-Heras-Casas, J., & García-Lozano, C. (2016). Energy performance certificates as tools for energy planning in the residential sector. The case of La Rioja (Spain). *Journal of Cleaner Production*, 137, 1280–1292. <https://doi.org/10.1016/j.jclepro.2016.08.007>
- Martin, M., Chong, A., Biljecki, F., & Miller, C. (2022). Infrared thermography in the built environment: A multi-scale review. *Renewable and Sustainable Energy Reviews*, 165, 112540. <https://doi.org/10.1016/j.rser.2022.112540>
- Mayer, K., Haas, L., Huang, T., Bernabé-Moreno, J., Rajagopal, R., & Fischer, M. (2023). Estimating building energy efficiency from street view imagery, aerial imagery, and land surface temperature data. *Applied Energy*, 333, 120542. <https://doi.org/10.1016/j.apenergy.2022.120542>
- Mayer, Z., Epperlein, A., Vollmer, E., Volk, R., & Schultmann, F. (2023). Investigating the Quality of UAV-Based Images for the Thermographic Analysis of Buildings. *Remote Sensing*, 15(2), 301. <https://doi.org/10.3390/rs15020301>
- Oltra-Carrió, R., Sobrino, J. A., Franch, B., & Nerry, F. (2012). Land surface emissivity retrieval from airborne sensor over urban areas. *Remote Sensing of Environment*, 123, 298–305. <https://doi.org/10.1016/j.rse.2012.03.007>
- Pandey, A., & Pati, U. C. (2019). Image mosaicing: A deeper insight. *Image and Vision Computing*, 89, 236–257. <https://doi.org/10.1016/j.imavis.2019.07.002>
- Parlow, E. (2021). Regarding Some Pitfalls in Urban Heat Island Studies Using Remote Sensing Technology. *Remote Sensing*, 13(18), 3598. <https://doi.org/10.3390/rs13183598>

- Pasichnyi, O., Wallin, J., Levihn, F., Shahrokni, H., & Kordas, O. (2019). Energy performance certificates—New opportunities for data-enabled urban energy policy instruments? *Energy Policy*, *127*, 486–499. <https://doi.org/10.1016/j.enpol.2018.11.051>
- Rüdisser, D., Posch, T., & Sulzer, W. (2024). The GRAZ Method—Determination of Urban Surface Temperatures from Aerial Thermography Based on a Three-Dimensional Sampling Algorithm. *Remote Sensing*, *16*(21), 3949. <https://doi.org/10.3390/rs16213949>
- UK Government. (2012). The Energy Performance of Buildings (England and Wales) Regulations 2012. (SI 2012/3118). Verfügbar unter: <https://www.legislation.gov.uk/ukSI/2012/3118/contents>
- Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. *Remote Sensing of Environment*, *86*(3), 370–384. [https://doi.org/10.1016/S0034-4257\(03\)00079-8](https://doi.org/10.1016/S0034-4257(03)00079-8)
- Wederhake, L., Wenninger, S., Wiethe, C., Fridgen, G., & Stirnweiß, D. (2022). Benchmarking building energy performance: Accuracy by involving occupants in collecting data - A case study in Germany. *Journal of Cleaner Production*, *379*, 134762. <https://doi.org/10.1016/j.jclepro.2022.134762>
- Zhao, C., Jensen, J., Weng, Q., & Weaver, R. (2018). A Geographically Weighted Regression Analysis of the Underlying Factors Related to the Surface Urban Heat Island Phenomenon. *Remote Sensing*, *10*(9), 1428. <https://doi.org/10.3390/rs10091428>
- Zhu, Z., Zhou, Y., Seto, K. C., Stokes, E. C., Deng, C., Pickett, S. T. A., & Taubenböck, H. (2019). Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sensing of Environment*, *228*, 164–182. <https://doi.org/10.1016/j.rse.2019.04.020>

Anhang

Alle im Rahmen dieser Masterarbeit erstellten Python-Skripte wurden in einem GitHub-Repository veröffentlicht: https://github.com/pbl128/energy_efficiency_thermal.

Anhang A: Python-Skript zur Berechnung des zeitlichen Trends und zur Normalisierung der Daten

```
import matplotlib.dates as mdates
import numpy as np
import os
import pandas as pd
import rasterio as rs
import re
import seaborn as sns
import statsmodels.api as sm
import zipfile
from datetime import datetime

# get date, time and mean temperature per tile
directory = r"C:\Users\p_bl\Desktop\Studium_Master\15_Thesis\data\unzipped"
data_temp = []
for subfolder in os.listdir(directory):
    subfolder_directory = os.path.join(directory, subfolder)
    for file in os.listdir(subfolder_directory):
        if file.endswith("TABI.tiff"):
            # extract date and time
            match = re.search(r"\d{8}\T\d{6}", file)
            datetime_string = match.group()
            time_object = datetime.strptime(datetime_string, '%Y%m%dT%H%M%S') # convert string to datetime-object
            # extract mean temperature
            with rs.open(os.path.join(subfolder_directory, file), "r") as tile:
                thermal = tile.read(1)
                thermal_masked = np.where(thermal == 0, np.nan, thermal) # mask nodata
                thermal_mean = np.nanmean(thermal_masked)
                thermal_median = np.nanmedian(thermal_masked)
            # append to list
            data_temp.append({
                "filename": file,
                "datetime": pd.Timestamp(time_object),
                "date": pd.Timestamp(time_object).date(),
                "time": pd.Timestamp(time_object).time(),
                "pxls": (~np.isnan(thermal_masked)).sum(),
                "temp_mean": thermal_mean,
                "temp_median": thermal_median,
            })
temp = pd.DataFrame(data_temp)

# calculate weighted linear regression (wls)
temp['time_seconds'] = temp['time'].apply(lambda x: x.hour * 3600 + x.minute * 60 + x.second)
data_reg = []
for date in temp['date'].unique():
    # calculate wls
    subset = temp[temp['date'] == date]
    X = np.array(subset['time_seconds']).astype("int")
    X = sm.add_constant(X) # add constant to calculate intercept
    y = np.array(subset['temp_median']).astype("float")
    w = np.array(subset['pxls']).astype("int")
    wls_mod = sm.WLS(y, X, weights=w)
    wls_res = wls_mod.fit()
    # append wls parameters to list
    data_reg.append({
        "date": date,
        "intercept": wls_res.params[0],
        "slope": wls_res.params[1],
        "w_mean": (y * w).sum()/w.sum(),
        "pearson_r": subset['time_seconds'].corr(subset["temp_median"]),
        "r_squared": wls_res.rsquared,
        "p_value_slope": wls_res.pvalues[1],
    })
reg = pd.DataFrame(data_reg)
print(reg)
```

```

# plot results
for date in temp['date'].unique():
    subset = temp[temp['date'] == date]
    g = sns.relplot(
        data=subset,
        x='datetime',
        y='temp_median',
        col='date',
        marker="o",
        size="pxls",
        color="#008888",
        col_wrap=None,
        facet_kws=dict(sharex=False, sharey=False),
        legend='auto'
    );
    # calculate regression line
    intercept, slope = reg.loc[reg['date'] == date, ['intercept', 'slope']].iloc[0]
    x_pred = np.linspace(19 * 3600 + 50 * 60, 21 * 3600 + 30 * 60, 100) # start = 19:50, end = 21:30
    y_pred = intercept + slope * x_pred
    # adaptation of the plot output
    x_pred_datetime = pd.to_datetime(x_pred, unit='s', origin=pd.Timestamp(date))
    ticks = pd.date_range(start=pd.Timestamp(f'{date} 20:00:00'), end=pd.Timestamp(f'{date} 21:30:00'), freq="15min")
    for ax in g.axes.flat:
        ax.set_xticks(ticks)
        ax.xaxis.set_major_formatter(mdates.DateFormatter('%H:%M'))
        ax.set_ylabel("Brightness Temperature (K)", fontweight='bold')
        ax.set_xlabel("Time", fontweight='bold')
        ax.set_title("")
        ax.plot(x_pred_datetime, y_pred, color="red")
    g.legend.set_title('Pixelanzahl')
    # save plot as png
    g.savefig(f'{date}.png", bbox_inches='tight')

# apply regression factor to data from 12.11.2022
read_dir = r"C:\Users\p_bl\Desktop\Studium_Master\15_Thesis\data\unzipped"
write_dir = r"C:\Users\p_bl\Desktop\Studium_Master\15_Thesis\data\time_adjusted"
for subfolder in os.listdir(read_dir):
    subfolder_directory = os.path.join(read_dir, subfolder)
    for file in os.listdir(subfolder_directory):
        if file.endswith("TABI.tiff") and file.startswith('20221112'):
            # calculate temperature delta
            temp_subset = temp[temp['filename'] == file]
            date_for_file = temp_subset['date'].iloc[0]
            reg_subset = reg[reg['date'] == date_for_file]
            min_time_seconds = temp[temp['date'] == date_for_file]['time_seconds'].min()
            time_delta = temp_subset['time_seconds'].iloc[0] - min_time_seconds
            temp_delta = time_delta * reg_subset['slope'].iloc[0]
            # subtract temperature delta and save files
            with rs.open(os.path.join(subfolder_directory, file)) as tile:
                prod = tile.read()
                prod_adj = np.where(prod == 0, 0, prod - temp_delta)
                with rs.open(os.path.join(write_dir, file), 'w', **tile.meta) as tile_adj:
                    tile_adj.write(prod_adj)

```

Anhang B: Python-Skript zur Berechnung der gewichteten Mittelwerte für die Brightness Temperatures

```
import arcpy
import pandas as pd
import numpy as np

fc = "Buildings_buffer_mosaic_12_intersect_075"
fields = ["id", "grid_code", "area_share"]
data = []

with arcpy.da.SearchCursor(fc, fields) as cursor:
    for row in cursor:
        data.append(row)

df = pd.DataFrame(data, columns=fields)

result_20221112 = df.groupby("id").agg(
    BT_wMean_12=("grid_code", lambda x: np.average(x, weights=df.loc[x.index, "area_share"])),
    BT_wMean_Count_Grids_12=("id", "count")
).reset_index()

csv_path=r"C:\Users\p_bl\Desktop\Studium_Master\15_Thesis\ArcGIS_MT\ArcGIS_MT_Thermal\result_20221112.csv"
result_20221112.to_csv(csv_path, index=False)
```

Anhang C: Python-Skript zur Erstellung der Diagramme für die deskriptiven Statistiken

```

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns

file = "Buildings_LST_mean.csv"

df = pd.read_csv(file, sep=";")
df["geometry_area_m2"] = df["geometry_area_m2"].str.replace(",",".").astype(float)
df["CURRENT_ENERGY_EFFICIENCY"] = df["CURRENT_ENERGY_EFFICIENCY"].astype(int)
df["LST_Celsius_13"] = df["LST_Celsius_13"].str.replace(",",".").astype(float)
df["LST_Celsius_12"] = df["LST_Celsius_12"].str.replace(",",".").astype(float)
df["LST_Celsius_mean"] = df["LST_Celsius_mean"].str.replace(",",".").astype(float)

# plot diagramm for building areas
bins = [30, 50, 60, 70, 80, 90, 100, 110, 120, 130, 5000]
labels = ['< 50', '> 50 - 60', '> 60 - 70', '> 70 - 80', '> 80 - 90', '> 90 - 100', '> 100 - 110', '> 110 - 120', '> 120 - 130', '> 130']
df["Gebaeudeflaechen"] = pd.cut(df["geometry_area_m2"], bins=bins, labels=labels, right=True)
mean_value = df["geometry_area_m2"].mean()
median_value = df["geometry_area_m2"].median()
min_value = df["geometry_area_m2"].min()
max_value = df["geometry_area_m2"].max()

ax = df["Gebaeudeflaechen"].value_counts().sort_index().plot(kind="bar", color="#008B8B", edgecolor="black")
plt.xlabel("Gebaeudeflaeche in m²", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')

x_pos = len(labels) - 4
y_pos = max(df["Gebaeudeflaechen"].value_counts())
text = (
    f"Minimum: {min_value:.1f} m²\n"
    f"Maximum: {max_value:.1f} m²\n"
    f"Mittelwert: {mean_value:.1f} m²\n"
    f"Median: {median_value:.1f} m²"
)

plt.text(x_pos, y_pos, text, fontsize=10, color="black",
         verticalalignment="top", horizontalalignment="left",
         bbox=dict(facecolor="white", alpha=0.8, edgecolor="black"))
plt.xticks(rotation=45)
plt.savefig("Statistik_Gebaeudeflaeche.png", bbox_inches='tight')
plt.show()

# plot diagramm for energy efficiency of the buildings (A to G)
order = ["A", "B", "C", "D", "E", "F", "G"]
category_counts = df["CURRENT_ENERGY_RATING"].value_counts().reindex(order)
category_counts.plot(kind="bar", color=["#4575b4", "#91bfb3", "#e0f3f8", "#ffffff", "#fee090", "#fc8d59", "#d73027"], edgecolor="black")

plt.xlabel("Energieeffizienzklasse der Gebäude", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.xticks(rotation=0)
plt.savefig("Statistik_Energieeffizienzklasse.png", bbox_inches='tight')
plt.show()

# plot diagramm for energy efficiency of the building roofs (Very Poor to Very Good)
order = ["Very Good", "Good", "Average", "Poor", "Very Poor"]
category_counts = df["ROOF_ENERGY_EFF"].value_counts().reindex(order)
category_counts.plot(kind="bar", color=["#2c7bb6", "#abd9e9", "#ffffff", "#fdae61", "#d7191c"], edgecolor="black")

plt.xlabel("Energieeffizienz der Gebäudedächer", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.xticks(rotation=0)
plt.savefig("Statistik_Dach_Energieeffizienz.png", bbox_inches='tight')
plt.show()

# plot histogram for energy efficiency of the buildings (1 to 100)
mean_value = df["CURRENT_ENERGY_EFFICIENCY"].mean()
median_value = df["CURRENT_ENERGY_EFFICIENCY"].median()

sns.histplot(df["CURRENT_ENERGY_EFFICIENCY"], bins=100, kde=False, color="#008B8B", edgecolor="black")
plt.axvline(mean_value, color="#7570b3", linestyle='dashed', linewidth=2, label=f'Mittelwert: {mean_value:.1f}')
plt.axvline(median_value, color="#d95f02", linestyle='dashed', linewidth=2, label=f'Median: {median_value:.1f}')
plt.legend()
plt.xticks(range(0, 110, 10))
plt.xlabel("Energieeffizienz der Gebäude", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.savefig("Statistik_Energieeffizienz_1_bis_100.png", bbox_inches='tight')
plt.show()

```

```

# plot histogram for the LST from 12.11.2022
mean_value = df["LST_Celsius_13"].mean()
median_value = df["LST_Celsius_13"].median()

print(f"Mittelwert: {mean_value}")
print(f"Median: {median_value}")
print(f"Minimum: {df["LST_Celsius_13"].min()}")
print(f"Maximum: {df["LST_Celsius_13"].max()}")
print(f"Spannweite: {df["LST_Celsius_13"].max()-df["LST_Celsius_13"].min()}")
print(f"Standardabweichung: {df["LST_Celsius_13"].std()}")
print(f"Q1: {df["LST_Celsius_13"].quantile(0.25)}")
print(f"Q3: {df["LST_Celsius_13"].quantile(0.75)}")
print(f"Skewness: {df["LST_Celsius_13"].skew()}")
print(f"Kurtosis: {df["LST_Celsius_13"].kurt() + 3}")

sns.histplot(df["LST_Celsius_13"], bins=100, kde=True, color="#008888", edgecolor="black")
plt.axvline(mean_value, color='#7570b3', linestyle='dashed', linewidth=2, label=f'Mittelwert: {mean_value:.2f} °C')
plt.axvline(median_value, color='#d95f02', linestyle='dashed', linewidth=2, label=f'Median: {median_value:.2f} °C')
plt.legend()
plt.xticks(np.arange(10.0, 13.5, 0.5))
plt.xlabel("Land Surface Temperature (LST) in °C", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.savefig("Statistik_LST_Celsius_13.png", bbox_inches='tight')
plt.show()

# plot histogram for the LST from 13.11.2022
mean_value = df["LST_Celsius_12"].mean()
median_value = df["LST_Celsius_12"].median()

print(f"Mittelwert: {mean_value}")
print(f"Median: {median_value}")
print(f"Minimum: {df["LST_Celsius_12"].min()}")
print(f"Maximum: {df["LST_Celsius_12"].max()}")
print(f"Spannweite: {df["LST_Celsius_12"].max()-df["LST_Celsius_12"].min()}")
print(f"Standardabweichung: {df["LST_Celsius_12"].std()}")
print(f"Q1: {df["LST_Celsius_12"].quantile(0.25)}")
print(f"Q3: {df["LST_Celsius_12"].quantile(0.75)}")
print(f"Skewness: {df["LST_Celsius_12"].skew()}")
print(f"Kurtosis: {df["LST_Celsius_12"].kurt() + 3}")

sns.histplot(df["LST_Celsius_12"], bins=100, kde=True, color="#008888", edgecolor="black")
plt.axvline(mean_value, color='#7570b3', linestyle='dashed', linewidth=2, label=f'Mittelwert: {mean_value:.2f} °C')
plt.axvline(median_value, color='#d95f02', linestyle='dashed', linewidth=2, label=f'Median: {median_value:.2f} °C')
plt.legend()
plt.xticks(np.arange(10.0, 14.5, 0.5))
plt.xlabel("Land Surface Temperature (LST) in °C", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.savefig("Statistik_LST_Celsius_12.png", bbox_inches='tight')
plt.show()

# plot histogram for the mean values of the LST
mean_value = df["LST_Celsius_mean"].mean()
median_value = df["LST_Celsius_mean"].median()

print(f"Mittelwert: {mean_value}")
print(f"Median: {median_value}")
print(f"Minimum: {df["LST_Celsius_mean"].min()}")
print(f"Maximum: {df["LST_Celsius_mean"].max()}")
print(f"Spannweite: {df["LST_Celsius_mean"].max()-df["LST_Celsius_mean"].min()}")
print(f"Standardabweichung: {df["LST_Celsius_mean"].std()}")
print(f"Q1: {df["LST_Celsius_mean"].quantile(0.25)}")
print(f"Q3: {df["LST_Celsius_mean"].quantile(0.75)}")
print(f"Skewness: {df["LST_Celsius_mean"].skew()}")
print(f"Kurtosis: {df["LST_Celsius_mean"].kurt() + 3}")

sns.histplot(df["LST_Celsius_mean"], bins=100, kde=True, color="#008888", edgecolor="black")
plt.axvline(mean_value, color='#7570b3', linestyle='dashed', linewidth=2, label=f'Mittelwert: {mean_value:.2f} °C')
plt.axvline(median_value, color='#d95f02', linestyle='dashed', linewidth=2, label=f'Median: {median_value:.2f} °C')
plt.legend()
plt.xticks(np.arange(10.5, 13.5, 0.5))
plt.xlabel("Land Surface Temperature (LST) in °C", fontweight='bold')
plt.ylabel("Anzahl der Objekte", fontweight='bold')
plt.savefig("Statistik_LST_Mittelwert.png", bbox_inches='tight')
plt.show()

```

Anhang D: Python-Skript zur Modellierung der Energieeffizienz mittels globaler Regressionsansätze und als multiskaliger Ansatz

```

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sns
import statsmodels.api as sm

# import data and adjust data types
file = "Buildings_LST_final.csv"
df = pd.read_csv(file, sep=";")

df["geometry_area_m2"] = df["geometry_area_m2"].str.replace(",",".").astype(float)
df["CURRENT_ENERGY_EFFICIENCY"] = df["CURRENT_ENERGY_EFFICIENCY"].astype(int)
df["LST_Celsius_13"] = df["LST_Celsius_13"].str.replace(",",".").astype(float)
df["LST_Celsius_12"] = df["LST_Celsius_12"].str.replace(",",".").astype(float)
df["LST_Celsius_mean"] = df["LST_Celsius_mean"].str.replace(",",".").astype(float)
df["ROOF_ENERGY_EFF_SCORE"] = df["ROOF_ENERGY_EFF_SCORE"].astype(int)

# define function which calculates regression and creates plot
def ols_regression_and_plot(df, LST, ENERGY_EFFICIENCY, name_energy_efficiency, LST_title, Efficiency_title, name_fig, ax_limit=None, Filter_title=None, Raemuliche_Ebene=None, ):
    # calculate linear regression (ordinary least square regression = ols)
    X = df[LST]
    y = df[ENERGY_EFFICIENCY]
    X = sm.add_constant(X)
    ols_mod = sm.OLS(y, X).fit()

    data_reg = []
    data_reg.append({
        "intercept": ols_mod.params["const"],
        "slope": ols_mod.params[LST],
        "pearson_r": df[LST].corr(df[ENERGY_EFFICIENCY]),
        "r_squared": ols_mod.rsquared,
        "p_value_slope": ols_mod.pvalues[LST],
        "count": df.shape[0]
    })
    reg = pd.DataFrame(data_reg) # create dataframe from list
    print(reg)

# plot results
g = sns.relplot(
    data=df,
    x=LST,
    y=ENERGY_EFFICIENCY,
    marker="o",
    color="#008080",
    col_wrap=None,
    facet_kws=dict(sharex=False, sharey=False),
    legend='auto',
    **{"s": 6}
);

ax = g.ax
sns.regplot(
    data=df,
    x=LST,
    y=ENERGY_EFFICIENCY,
    scatter=False,
    line_kws={"color": "red"},
    ci=90, # 90 % confidence interval (explorative analyses)
    ax=ax
)

text = (
    f"r2: {reg['r_squared'].iloc[0]:.3f}\n"
)
g.set_axis_labels("Land Surface Temperature (LST) in °C", name_energy_efficiency, fontweight='bold')
if(ax_limit):
    ax.set_ylim(0, ax_limit)
if(Filter_title):
    ax.set_title(f"LST: {LST_title}\nEnergieeffizienz: {Efficiency_title}\nFilterung: {Filter_title}\nr2: {reg['r_squared'].iloc[0]:.3f}\nObjektanzahl: {reg['count'].iloc[0]:.0}", loc="left", fontweight='bold')
elif(Raemuliche_Ebene):
    ax.set_title(f"LST: {LST_title}\nEnergieeffizienz: {Efficiency_title}\nRaemuliche Ebene: {Raemuliche_Ebene}\nr2: {reg['r_squared'].iloc[0]:.3f}\nObjektanzahl: {reg['count'].iloc[0]:.0}", loc="left", fontweight='bold')
else:
    ax.set_title(f"LST: {LST_title}\nEnergieeffizienz: {Efficiency_title}\nr2: {reg['r_squared'].iloc[0]:.3f}", loc="left", fontweight='bold')
plt.savefig(name_fig, bbox_inches='tight')
plt.show()

# use the regression function for the different LST values once for the building energy efficiency and once for the roof energy efficiency
ols_regression_and_plot(df, "LST_Celsius_12", "CURRENT_ENERGY_EFFICIENCY", "Energieeffizienz der Gebäude", "Werte vom 12.11.2022", "Gebäude (1 bis 100)", "Reg_Geb_Eff_LST12.png", ax_limit=110)
print()
ols_regression_and_plot(df, "LST_Celsius_13", "CURRENT_ENERGY_EFFICIENCY", "Energieeffizienz der Gebäude", "Werte vom 13.11.2022", "Gebäude (1 bis 100)", "Reg_Geb_Eff_LST13.png", ax_limit=110)
print()
ols_regression_and_plot(df, "LST_Celsius_mean", "CURRENT_ENERGY_EFFICIENCY", "Energieeffizienz der Gebäude", "Mittelwert beider Tage", "Gebäude (1 bis 100)", "Reg_Geb_Eff_LSTmean.png", ax_limit=110)
print()
ols_regression_and_plot(df, "LST_Celsius_12", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Werte vom 12.11.2022", "Dächer (1 bis 5)", "Reg_Dach_Eff_LST12.png", ax_limit=6)
print()
ols_regression_and_plot(df, "LST_Celsius_13", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Werte vom 13.11.2022", "Dächer (1 bis 5)", "Reg_Dach_Eff_LST13.png", ax_limit=6)
print()
ols_regression_and_plot(df, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)", "Reg_Dach_Eff_LSTmean.png", ax_limit=6)

# filter the df to different building areas
df_area_50 = df[df["geometry_area_m2"] > 50]
df_area_60 = df[df["geometry_area_m2"] > 60]
df_area_70 = df[df["geometry_area_m2"] > 70]
df_area_80 = df[df["geometry_area_m2"] > 80]
df_area_100 = df[df["geometry_area_m2"] > 100]
df_area_120 = df[df["geometry_area_m2"] > 120]

```

```

# use the regression function for the filtered df
ols_regression_and_plot(df_area_50, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche50_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 50 m²")
print()
ols_regression_and_plot(df_area_60, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche60_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 60 m²")
print()
ols_regression_and_plot(df_area_70, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche70_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 70 m²")
print()
ols_regression_and_plot(df_area_80, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche80_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 80 m²")
print()
ols_regression_and_plot(df_area_100, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche100_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 100 m²")
print()
ols_regression_and_plot(df_area_120, "LST_Celsius_mean", "ROOF_ENERGY_EFF_SCORE", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Flaeche120_Dach_Eff_LSTmean.png", ax_limit=6, Filter_title="Gebäudefläche > 120 m²")
print()

# calculate the average LST and average energy efficiency values for different spatial levels
df_grouped_100 = df.groupby("GRID_ID_100").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("GRID_ID_100", "size")
).reset_index()
df_grouped_100 = df_grouped_100[df_grouped_100["count"]>3]

df_grouped_200 = df.groupby("GRID_ID_200").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("GRID_ID_200", "size")
).reset_index()
df_grouped_200 = df_grouped_200[df_grouped_200["count"]>5]

df_grouped_300 = df.groupby("GRID_ID_300").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("GRID_ID_300", "size")
).reset_index()
df_grouped_300 = df_grouped_300[df_grouped_300["count"]>10]

df_grouped_500 = df.groupby("GRID_ID_500").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("GRID_ID_500", "size")
).reset_index()
df_grouped_500 = df_grouped_500[df_grouped_500["count"]>20]

df_grouped_1026 = df.groupby("GRID_ID_1026").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("GRID_ID_1026", "size")
).reset_index()
df_grouped_1026 = df_grouped_1026[df_grouped_1026["count"]>20]

df_grouped_ward = df.groupby("Ward_Entity").agg(
    LST_grouped=("LST_Celsius_mean", "mean"),
    ROOF_ENERGY_EFFICIENCY_grouped=("ROOF_ENERGY_EFF_SCORE", "mean"),
    ENERGY_EFFICIENCY_grouped=("CURRENT_ENERGY_EFFICIENCY", "mean"),
    count=("Ward_Entity", "size")
).reset_index()
df_grouped_ward = df_grouped_ward[df_grouped_ward["count"]>20]

# use the regression function for the df grouped by spatial levels
ols_regression_and_plot(df_grouped_100, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Raumlich100_Dach_Eff_LSTmean.png", Raemliche_Ebene="Hexagonraster (100 m Seitenlänge)")
print()
ols_regression_and_plot(df_grouped_200, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Raumlich200_Dach_Eff_LSTmean.png", Raemliche_Ebene="Hexagonraster (200 m Seitenlänge)")
print()
ols_regression_and_plot(df_grouped_300, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Raumlich300_Dach_Eff_LSTmean.png", Raemliche_Ebene="Hexagonraster (300 m Seitenlänge)")
print()
ols_regression_and_plot(df_grouped_500, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Raumlich500_Dach_Eff_LSTmean.png", Raemliche_Ebene="Hexagonraster (500 m Seitenlänge)")
print()
ols_regression_and_plot(df_grouped_1026, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_Raumlich1026_Dach_Eff_LSTmean.png", Raemliche_Ebene="Hexagonraster (1026 m Seitenlänge)")
print()
ols_regression_and_plot(df_grouped_ward, "LST_grouped", "ROOF_ENERGY_EFFICIENCY_grouped", "Energieeffizienz der Dächer", "Mittelwert beider Tage", "Dächer (1 bis 5)",
"Reg_RaumlichWards_Dach_Eff_LSTmean.png", Raemliche_Ebene="Wards")

```

Anhang E: Python-Skript zur Modellierung der Energieeffizienz mittels räumlich expliziter Regressionsansätze

```
import geopandas as gpd
from libpysal.weights import DistanceBand
from spreg import GM_Error
from spreg import GM_Lag

# import data
gdf = gpd.read_file("Geodaten100m/Buildings_LST_Points_100m_oneGroup.shp")
y = gdf[["RoofEff"]].values
X = gdf[["LST"]].values

# calculate weight matrix based on inverse distances with threshold 100 m
w = DistanceBand.from_dataframe(gdf, threshold=100)
w.transform = "r"

# calculate spatial lag modell
slm_model = GM_Lag(y, X, w=w)
data_reg_slm = []
data_reg_slm.append({
    "intercept": slm_model.betas[0][0],
    "slope": slm_model.betas[1][0],
    "pseudo_r_squared": slm_model.pr2,
    "p_value": slm_model.z_stat[1][1],
    "rho": slm_model.rho[0],
    "p_value_rho": slm_model.z_stat[2][1],
})
reg_slm = pd.DataFrame(data_reg_slm)
print("Spatial Lag:")
print(reg_slm)
gdf["slm_Resid"] = slm_model.u.flatten()

# calculate spatial error modell
sem_model = GM_Error(y, X, w=w)
data_reg_sem = []
data_reg_sem.append({
    "intercept": sem_model.betas[0][0],
    "slope": sem_model.betas[1][0],
    "pseudo_r_squared": sem_model.pr2,
    "p_value": sem_model.z_stat[1][1],
    "lambda": sem_model.betas[-1][0],
    "p_value_lambda": sem_model.z_stat[-1][1],
})
reg_sem = pd.DataFrame(data_reg_sem)
print("\nSpatial Error:")
print(reg_sem)
gdf["sem_Resid"] = sem_model.u.flatten()

# export shape-file with residuals
gdf.to_file("GeodatenResid\Buildings_LST_Points_100m_oneGroup_Resid.shp", driver="ESRI Shapefile", encoding="utf-8")
```