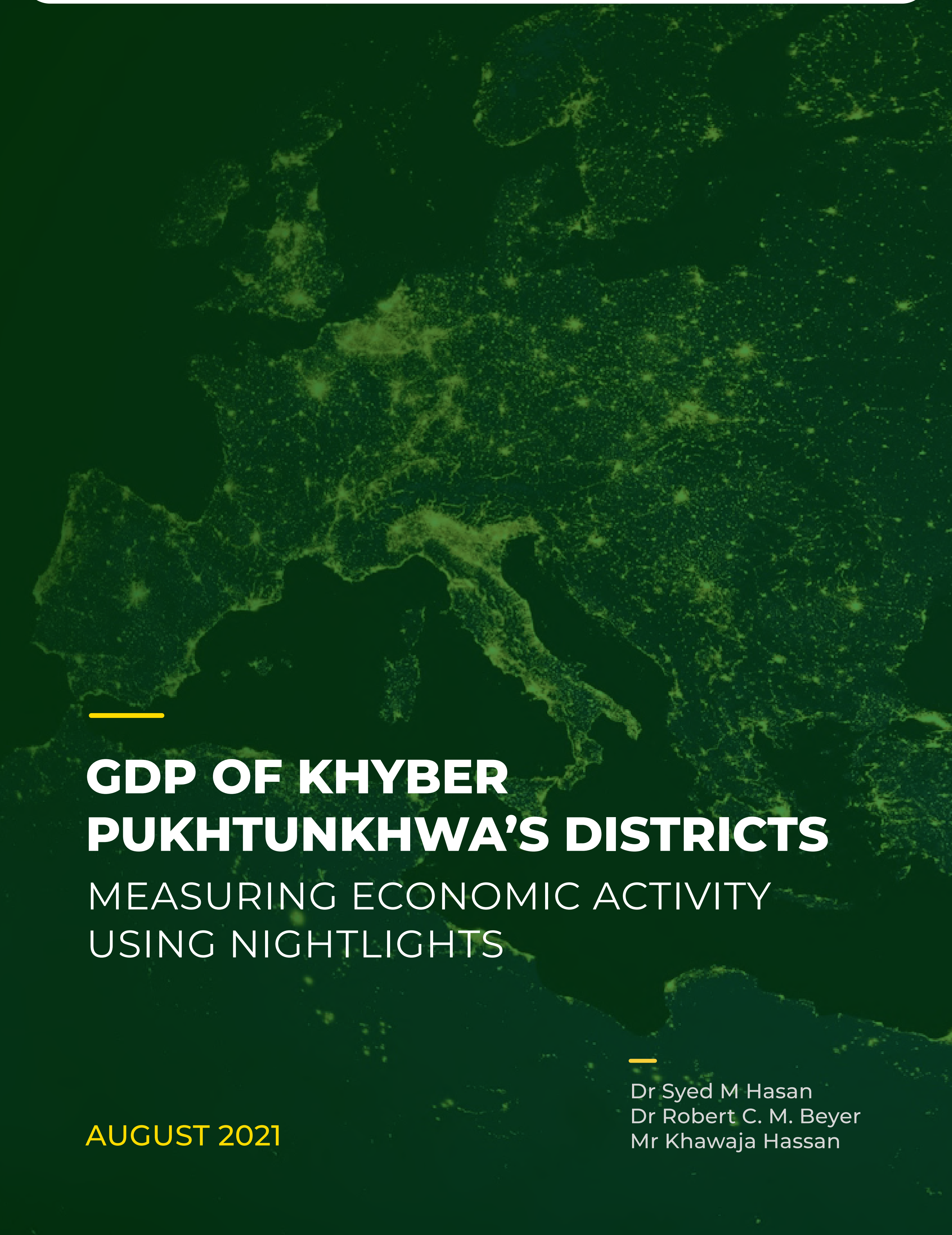


GDP OF KHYBER PUKHTUNKHWA'S DISTRICTS

MEASURING ECONOMIC ACTIVITY USING NIGHTLIGHTS

AUGUST 2021

Dr Syed M Hasan
Dr Robert C. M. Beyer
Mr Khawaja Hassan

About SEED

Sustainable Energy and Economic Development (SEED) is a £37.5 million programme funded and managed by the UK's Foreign Commonwealth and Development Office (FCDO). The first component - Improved Economic and Urban Planning in Khyber Pakhtunkhwa (KP) – with a budget of £15 million, is delivered by Adam Smith International (ASI). The SEED programme, in close collaboration with the Government of Khyber Pakhtunkhwa (GoKP), aims to improve economic and urban planning in Khyber Pakhtunkhwa (KP) to help the province plan and finance investments needs for growth, jobs and prosperity.

This nightlights study has been undertaken in collaboration with and using support from, the Regional Accounts Wing at the KP Bureau of Statistics.

TABLE OF CONTENTS

Acknowledgments v
*Message from **Mr Mahmood Khan*** vi
*Message from **Mr Syed Shahab Ali Shah*** vii
*Message from **Mr Hasaan Khawar*** viii
Executive Summary ix

01. Introduction 01
02. Research Objectives 05
03. Review of Literature 07
Urban Clusters 12
04. Datasets 15
05. Methodology 19
5.1 *Measuring Economic Activity using NTL* 20
5.2 *Extracting Cities or Urban Areas* 23
5.3 *Land Use and Land Cover Maps for KP* 27
06. Results 29
6.1 *Economic Growth and NTL* 30
6.2 *Detecting Urbanisation and Urban Markets* 45
6.3 *Comparing Urban Areas in KP using MODIS and NTL* 54
6.4 *GDP of FATA / Newly Merged Districts of KP Province* 58
6.5 *Limitations* 59
07. Using the NTL Data 60
7.1 *Possible Uses Learnt from the Study* 61
7.2 *Future Extensions* 62
08. References 64
09. Appendix A: Tables 69
10. Appendix B: KP District Level Urban Area (hectares) using MODIS 73
11. Appendix C: Provincial GDP Estimates by KP's BOS - Available at BOS KP Website 75

LIST OF

FIGURES

Figure 1: NTL trend among provinces of Pakistan 10
Figure 2: Nightlights in KP Province 14
Figure 3: Oil and Gas Fields in Kohat and Karak Districts 23
Figure 4: Sub-region Partitioning based on the Quadratic Model Source: Ma et al. (2015) 26
Figure 5: Pakistan Map with NTL and District Boundary 28
Figure 6: DN values Distribution 33
Figure 7: Actual and Estimated GDP for Industries and Services Sector-KP Province (Source: Author's estimation) 37
Figure 8: District GDP Ranking for 2019-20 43
Figure 9: Calculating Urban Population using CIESIN (2018) 46
Figure 10: DN Distribution for Selected Districts 49
Figure 11: Annual Percentage Change for Peshawar, Abbottabad, and Mardan 50
Figure 12: Peshawar BG Map 52
Figure 13: Mardan and Abbottabad BG Maps 53
Figure 14: Urban Area Extraction Using NTL-KP Province 54
Figure 15: KP land use maps using Landscan (A and B) 55
Figure 16: Peshawar Urban Growth Using MODIS Data 56
Figure 17: Peshawar Urban Boundary - Using NTL and BG 56
Figure 18: Periphery Growth in Abbottabad (2014 - 2020) 57

LIST OF TABLES

Table 1 : Sectoral Contribution of Industries in KP Province GDP (%) 12
Table 2: Population and urban areas 13
Table 3: Data Related Information.......... 18
Table 4: Regression results for GDP on SOL per area - Global Replication at Country level 31
Table 5: Regression results for GDP on SOL per area (South Asia - Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka) 32
Table 6: Regression results for GDP on SOL per area for South Asia (South Asia - Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka) 34
Table 7: Sub-national Regression 35
Table 8: Provincial Share in National GDP 36
Table 9: Provincial Share in National GDP (in percentage) ... 37
Table 10: Provincial GDP (NTL and Offical in Billion PKR) 38
Table 11: Economic Ranking of KP Districts 40
Table 12: Economic ranking of districts 41
Table 13: District Level MPI (2014-15) 42
Table 14: Province Level Agriculture Model Estimation 44
Table 15: DN value distribution - district level 47
Table 16: Urban Area Growth by Type for Select Districts (Using BG Approach) 51

LIST OF ACRONYMS

BOS:	Bureau of Statistics
BRI:	Belt and Road Initiative
CAREC:	Central Asia Regional Economic Cooperation Program
CPEC:	China-Pakistan economic corridor
FCDO:	Department for International Development
DMSP:	Defense Meteorological Satellite Program
DN:	Digital Number
FATA:	Federally Administered Tribal Area
GDP:	Gross Domestic Product
KP:	Khyber Pakhtunkhwa
LCU:	Local Currency Units
LSM:	Large-scale manufacturing
LULC:	Land Use Land Cover
MCDA:	Multiple Criteria Decision Analysis
MODIS:	Moderate Resolution Imaging Spectroradiometer
NGDC:	National Geophysical Data Center
NOAA:	National Oceanic and Atmospheric Administration
NTL:	Night-time lights data
PBS:	Pakistan Bureau of Statistics
SEED:	Sustainable Energy & Economic Development
SOL:	Sum of Lights
SSM:	Small-scale manufacturing
VIIRS:	Visible Infrared Imaging Radiometer Suite
WDI:	World Development Indicators

ACKNOWLEDGEMENTS

We would like to thank the Government of Khyber Pakhtunkhwa for providing us with data and guidance. We are especially grateful to Syed Shahab Ali Shah, Additional Chief Secretary, Mr Amer Tareen, Secretary P&D, Mr Ali Asghar, Chief Economist, Mr Khaliq-ur-Rehman, Director Bureau of Statistics and Mr Yahya Bakhtiar, Head of Regional Accounts wing at Bureau of Statistics in this regard.

We would also like to thank the SEED team for their support throughout the project, especially Ms Nazish Afraz, Programme Economist SEED, for reviewing the report multiple times, and Mr Omar Mukhtar, SEED Strategic Planning Advisor, for his valuable technical input.

MESSAGE FROM

Mr Mahmood Khan

Chief Minister
Khyber Pakhtunkhwa

The Government of KP is committed to evidence-based policymaking, the backbone of which is data on KP's economy. We are also committed to ensuring that our policies are making the intended impact and equitable distribution of resources among various districts and cities of the province. Better and reliable data is a key to achieving these goals.

This report has multifaceted bearings. It provides important insights, particularly on spatial disparities and urban centres, which we will be looking at closely in the days to come. This is a great starting point for the Government of KP to continue the use of nightlights to track the pulse of the economy.

We appreciate the effort and the report. This is an important part of our wider engagement with SEED to improve public investment efficiency in KP. We hope this will allow us to target our expenditures better to generate economic growth. I would also like to congratulate our colleagues at the KP Bureau of Statistics for their support in this report and am encouraged by their receptivity to using new methods to supplement traditional data collection.

We look forward to using the findings of this report.

MESSAGE FROM

Mr Syed Shahab Ali Shah

Additional Chief Secretary
Khyber Pakhtunkhwa

On many aspects of the economy, most notably the informal sector, it remains a challenge to collect rigorous, primary data. Without that data, it isn't easy to gauge the size and health of the economy. While the Government of Khyber Pukhtunkhwa is committed to encouraging formalization, and in parallel, to expand our collection of primary data on the services and industrial sectors more generally, it is refreshing to have this alternative window of nightlights through which to view economic activity in the province.

The report offers a host of fresh insights, and it is also heartening to note that it validates and confirms our own official statistics where both are available. I welcome the report and the introduction of this new methodology in government. This is an essential step on the path that the Government of Khyber Pukhtunkhwa has already embarked on, leveraging innovative technologies to inform and provide feedback on our policies.

+ + +

MESSAGE FROM

Hasaan Khawar

Team Leader
Sustainable Energy & Economic Development (SEED) programme

This report is an important part of SEED's commitment to support the Government of Khyber Pukhtunkhwa in increasing the effectiveness of public investments. The inexpensive and innovative approach used is a particularly useful source of information on KP's urban centres. These urban centres can become drivers of growth and hubs of innovation. Yet insufficient or poor planning can not only dampen the potential growth, but also contribute to pollution, congestion, and bottlenecks in the movement of goods and labour to the city, and within the city i.e., the efficiency of converting inputs into outputs. The report provides useful insights on how KP's cities are growing, for example, identifying pockets of rapid growth, the extent and pace of urban sprawl, low levels of urbanization overall, and the slow rise of secondary cities. Cities that are growing need infrastructure to ensure that the growth is sustained and is sustainable. Regions and cities that are lagging will need regenerative efforts. Each of these findings has important policy and planning implications, and we hope that they can find their way into KP's urban planning.

The Government of KP has made substantial inroads into expanding revenue generation. This report should contribute to the Government of KP's and SEED's larger efforts to ensure that this revenue is spent such that it maximizes the impact of each rupee spent.

I would like to thank the authors for this excellent report, and the Government of KP for lending their support wholeheartedly to this report.

Congratulations!

EXECUTIVE SUMMARY

Introduction

This research study has been commissioned by FCDO's SEED program to support the Government of KP in understanding the spatial distribution of economic activity within the province. It uses nightlights data to provide insights into:

- District-level GDP, GDP growth and income per capita
- For KP's major cities, patterns of growth and the distribution of economic activity within the city
- Province-level estimates of GDP

These insights can be utilized by the Government of KP to plan effective, well-targeted public investments that support growth and address regional disparities. They can be used, for example, to plan road infrastructure to improve the connectivity of existing economic clusters. In addition, the data reveals low and stagnating levels of urbanization in KP overall, indicating that the growth potential of cities is not being utilized fully. This can be useful in planning secondary cities and in addressing pockets of stagnation to ensure that development is inclusive and that all areas of KP are able to become a part of KP's growth story.

In addition, insights on the existing cities can feed into better urban planning, for example to address issues of urban sprawl and plan for the provision of urban services and infrastructure to accommodate the pace and direction of growth. This in turn will ease the bottlenecks that prevent cities from becoming efficient and environmentally sustainable contributors to the provincial economy.

This is the first study in Pakistan to use this novel approach to estimate district- and city-level economic activity. It will remain a practical complement to official data as it allows for live feedback on the state of the economy at levels of disaggregation as fine as 1 km^2, in addition to being readily available, easy to use and inexpensive.

Methodology

This study uses harmonized nightlight data to estimate non-farm economic activity. At the first stage, cross-country data is used to estimate the elasticity between GDP and nightlights growth, to help establish the strength of nightlights data in estimating GDP. The second step involves a similar exercise at the national level for South Asian countries and the subnational (province/state) level for urban areas in Pakistan and India. We find that while the relationship still holds at the subnational level, it is weaker than that in the cross-country regression, indicating lower strength in estimating GDP in the South Asian economies. At the third stage, the estimated coefficients from the sub-national model are used to estimate the non-agricultural component of GDP for the KP province of Pakistan. At the fourth stage, to arrive at the GDP of the districts and cities of KP, official provincial GDP is distributed using nightlights to determine the share of non-agricultural GDP in each district and city, and the share of rural population to distribute the share of agricultural GDP. Finally, the distribution of nightlights at the district level along with the daytime satellite imagery is used to identify urban growth trends.

Key Insights

District-wise distribution

District-level estimates of economic activity are arrived at by distributing official provincial GDP numbers in proportion with nightlights for non-agricultural activity, and in proportion with the rural population for agricultural activity. Several interesting insights can be gleaned from the data:

- While the overall size of the economy in Peshawar district is the largest in the KP province, Haripur district has the highest per-capita income followed by Nowshera and Abbottabad. In terms of per-capita GDP, the lowest figure is for Kohistan district preceded by Tor Ghar district. These results are consistent with the per-capita estimates from district-level Multidimensional Poverty Index (MPI) estimates made by UNDP Pakistan.

- The share of Peshawar in the provincial economy has grown from 16% to 19% from 2005-06 to 2019-20.

- The extent of regional disparity within KP is considerable whether measured through the overall size of the local economy or through per capita figures.

- Focusing on the lagging districts, Kohistan and Tor Ghar districts continue to be the lowest end of the spectrum in terms of overall economic activity. In terms of contribution towards the provincial economy, the combined share of eight lagging districts—Hangu, Buner, Upper Dir, Shangla, Battagram, Chitral, Kohistan, and Tor Ghar—is lower than the individual share of more economically prosperous districts such as Peshawar, Nowshera, Mardan, and Haripur

- Per capita GDP estimates at the district level reveal a similar story where there is a staggering difference among districts. The per capita GDP of Kohistan district is thirty times lower than that of Haripur district. The situation of the poorest four districts identified as Shangla, Upper Dir, Tor Ghar, and Kohistan has also been validated using multiple indicators such as literacy rates, Enrollment Gender Parity Index, and the UNDP developed district-level multidimensional poverty indices for the year 2014-15.

- All the newly merged districts except Khyber, Mohmand and North Waziristan have a nearly null share in the manufacturing/services sector as assessed by the NTL. As such we can safely assume that most economic activity in the region relates to the farming sector.

- Using the NTL share and the rural population share, the share of economic activity in the NMDs is around 1 % of national GDP.

- Nightlight maps overlaid on GDP maps show that economic growth and prosperity are strongly correlated with urban development. The districts with high GDPs—Peshawar, Nowshera, and Mardan—are the ones with the brightest NTL as well. Thus, it is evident that the agriculture sector's potential to bring major changes in the local economy is limited. This finding is in line with the urban economic theory that states that high density of population and economic activity in cities unleash agglomeration economies and benefits of scale.

- Following the high economic value block of Peshawar, Mardan, and Nowshera, the map shows that Haripur and Abbottabad districts form another emerging agglomeration. Economic spillover benefits can be observed in the case of Swabi, Mansehra, and Kohat districts as they are geographically contiguous to one or more high GDP districts. Thus, a pattern of growing and lagging districts is visible and hence could be useful for future public sector investment decisions. The NTL data also indicates that around 85% of non-farm economic activity is currently located in seven districts of KP: Peshawar, Nowshera, Haripur, Mardan, Abbottabad, D. I. Khan, and Kohat.

- The process of urbanization in the KP province has largely remained stagnant. As such, KP's share in the national economy has also remained around 9–10% in the period under study.

Cities

Since urban areas generate mostly industrial and service sector economic activity, nightlights are the best fit for generating city-level insights. The following insights emerge from the analysis:

- Urbanization in the province, when analysed using either the census data or the nightlights data, indicates a stagnating trend. According to the census, only one city in the province has a population of more than 1 million and hence all other cities are relatively small. Comparing across provinces, Sindh has two such cities (including Karachi, the largest city of the country) and Punjab has five such cities (including Lahore, the second-largest city of the country).

- In the case of Peshawar, the high-density city core and low-density city periphery have significantly increased over time. The extent of high intensity lights has not only increased around the 2014 area but also encompasses a new area in the east.

- Despite having an increase in the high-density core region, Peshawar still experiences urban sprawl as well.

- In the case of Abbottabad, there is little change in the city core. However, most of the recent growth has taken place on the city periphery which is again indicative of increasing urban sprawl.

- Both the tehsils of Abbottabad district – Abbottabad and Havelian, as well as locations in the Galiyat region, are experiencing new low density urban development which is often regarded as a sub-optimal use of land leading to multiple inefficiencies in resource allocation and service delivery.

- In Mardan, Nowshera and Abbotabad, there is dimming of nightlights at the core of the city and more growth on the periphery which signals horizontal urban growth and potential issues related to efficient provision of municipal services and public transportation. Urban growth on the periphery of cities is sometimes not truly captured by population census. Hence, this form of urban growth could lead to “messy” and “hidden” urbanization, as highlighted by the 2015 World Bank report on the South Asian cities.

Estimated GDP for KP Province

To ensure the reliability of results, this study uses two methodologies to estimate province-level GDP for the KP province. In the first approach, coefficients from the regression on subnational GDP in India and Pakistan are used to estimate province-level GDP (manufacturing and services industries only) for the KP province. Due to data limitations, this is done for the period 2004-05 to 2013-14. The estimated and actual values for the non-agricultural component are very close, which is an indication of the robustness of our methodology, and the official data produced by KP BOS. The difference between the actual and estimated GDP can indicate the presence of an informal economy, although such differences can also reflect differences in methodology. Between 2009-2014, this difference has varied in the range of 5-11%.

In the second method, official national GDP numbers are distributed into provincial shares using their shares in nightlights and rural population numbers. To make up for the missing data in the earlier approach and to avoid weaknesses of nightlights data, this exercise is carried out for the years 2013-14 to 2019-20.

Focusing on recent years, it is observed that the difference between official and estimated GDP figures was very small in 2018-19 but has again increased in 2019-20. These differences between official and estimated GDP figures arise because of (a) differences in methodologies and (b) annual variations in the scale of informal economic activity. The presence of these differences highlights the need to focus more on capturing the informal economic activity in the official estimates.

Using the Data

This study provides several insights that can be utilized for better planning and more effective targeting of public investments.

Firstly, it gives a clear understanding of the economic growth and development of districts in the KP province. This information is a prerequisite for planning infrastructure (for example to connect vibrant economic clusters and industrial parks) and is vital for designing policies to address regional disparities.

Secondly, our intra-city analysis gives a clear picture of the growth of cities, in particular the pace and direction of new development. This information can feed into urban planning, for example addressing issues of urban sprawl, planning for new infrastructure, and provision of urban services and local amenities (roads, transportation, utilities, crime control, waste disposal, mitigation of environmental degradation) to accommodate growing cities, realize their growth potential and help make them sustainable and efficient. It can also be used to develop a system of cities with plans for secondary cities, that so far appear to be less visible in the case of KP. The report also documents a low and stagnating level of urbanization in KP overall, which means that the agglomeration and growth potential of cities is currently highly underutilized and is an area where policymakers in KP can devote more attention.

In addition to these immediate insights, once the capacity to understand and leverage nightlights data is developed in KP, it opens the doors to a reliable, live, and inexpensive source of information on future topics where official data is usually silent.

The spatial information embedded in the nightlights data gives it the unique advantage of providing granularity to the level of 1 km^2. Estimation of sub-national GDP to the level of a district or a tehsil/town yields an economic indicator that can be extensively used for policy planning and evaluation. Nightlights have been used to (a) obtain reliable estimates of poverty; (b) study the impact of cash transfers; (c) create wealth index; (d) create development indices; (e) validate district per capita income values obtained from Labor Force Surveys; (f) calculate regional inequality measures (g) understand urban crime patterns, etc. Most of these studies have been carried out by combining gridded nightlights data, gridded population data, and socioeconomic variables from other available datasets.

Recently, nightlight data has been used to study the economic impact of various events, for example, restrictive government measures such as the imposition of lockdowns following the COVID-19 outbreak. The argument is that non-pharmacological measures to contain disease transmission while mitigating their economic impact require an assessment of the economic situation in near real-time and at high spatial granularity. A comparison of changes in light intensity before and after lockdowns indicates the impact on local economic activity and hence provides useful information for policymakers. This approach could be useful not just to measure COVID impacts but also to assess the impact of natural disasters or other policy measures, in a sense allowing the policymaker a way to continually monitor the pulse of economic activity in the province.

The utilization of satellite imagery, in particular nightlights, is a relatively new concept and is still evolving as researchers find novel ways to use it to understand development. It is an opportune time for KP to develop the expertise to analyze and utilize this rich source of information to help achieve its development targets.

01 INTRODUCTION

During the last century, a greater part of mainstream research in economics – both theoretical and empirical – has discounted the significance of geographical space. Although economic issues invariably involve questions concerning the place specificity of activities as well as concerns related to overcoming distance constraints, due to the non-tractability or non-availability of granular data, variables capturing space or location had been excluded from economic models in the past. However, with recent technological advancements, the availability of spatial data on a global level and improvement in computational skills have opened new avenues of research. Coupled with advances in technology that have led to the development of new tools that help overcome this challenge of spatial analysis, economists have started to frequently utilize this new information and better explain the spatial element of economic activity.

Consider the case of Gross Domestic Product (GDP) – a well-known economic measure that for a specific period, provides the total value of all goods and services produced within an economy. The GDP of a country is an all-inclusive indicator of a country's economic wellbeing. Although being an aggregate, the use of GDP as an economic pointer creates some opacity yet it is widely relied upon to provide an overall snapshot of an economy and used to estimate its size and growth over time. The GDP of an economy can be calculated through expenditures made, the output produced, or incomes generated and can be adjusted for inflation effect and population numbers to provide meaningful insights. Callen (2012) elaborates that GDP measurement can be done in three different ways:

The expenditure approach adds up the value of purchases made by final users—for example, the consumption of food, televisions, and medical services by households; the investments in machinery by companies; and the purchases of goods and services by the government and foreigners.

The production approach sums the "value-added" at each stage of production, where value-added is defined as total sales less the value of intermediate inputs into the production process.

The income approach sums the incomes generated by production—for example, the compensation employees receive and the operating surplus of companies.

Typically, the national and provincial statistical agencies[1] calculate the national and provincial GDPs using surveys of all industries and trade flows. However, despite the effort undertaken, sampling and data collection in the case of developing countries is often undertaken at the second tier of administrative units' hierarchy that is at the province level. This approach limits the depth of information about economic activity at a more spatially disaggregated level. The use of nationally or provincially aggregate GDP is not able to explain the welfare distribution of economic growth or output as economic agents and entities are spatially spread across a country.

The spatial heterogeneity in location choices of economic agents as well as the presence of place-specific natural advantages means that contribute to and share in the final economic aggregate will vary across space. Hence it is important to have consistent data on economic activity for smaller geographical units to understand growth at sub-national levels. Besides, reliable impact assessment of previous policies and robust design of future growth strategies hinges on the availability of such estimates. Nonetheless, the challenges associated with GDP estimation at sub-national levels such as provinces or districts become

[1] In Pakistan, the national statistical agency is the Pakistan Bureau of Statistics. The BOS KP is the relevant provincial agency.

more pronounced due to deficiency of prior experience for this, and constraints faced in data collection and compilation at that level. In Pakistan, provinces have recently gained extensive functional autonomy after the passage of the 18th constitutional amendment in 2010. Following this, they are gradually building their institutional capacities. So far due to various political exigencies, the local governments at the district level have rarely operated as a functional tier of government and hence lack trained personnel and necessary resources. Thus, both provincial and district governments need to improve their capacity which shall bear long-run benefits for the society through improved governance. It is added for the reader’s information that districts are at the third level of the administrative division below provinces and divisions but hold special significance as they form the top tier of local government structure.

The study of economic activity at a granular level is also important because much of the success of public policy instruments depends on economic conditions immediately surrounding the targeted group of people. The current growth of development economics and policy experiments focusing on health, governance, local finance, etc. demands statistics that are not aggregated at the country level but bifurcated along the lines of various smaller administrative and geographical units. To overcome the constraints and potential errors in obtaining reliable estimates of economic output at the national and sub-national level, the economic literature has several examples where the researchers have resorted to the use of various proxies generated from satellite imagery. One such proxy that is being most widely used in recent times is the amount of light that can be observed from outer space (Henderson, et al., 2009). Besides nightlights that are a good proxy for economic activity in the industrial and services sector, another measure used is the Land Use Land Cover data which has been found useful in literature (see Keola, 2015) to capture the contribution of the agriculture sector.

A region-based analysis of economic activity further reveals that urban centers are the major contributors to the national GDP. The structure of an urban center eases out the constraints in achieving high productivity, implying that spatial allocation and temporal growth of urban clusters are good indicators of how successful a nation has been in achieving high economic growth. The rapid pace of urbanization in South Asian countries, especially Pakistan, and the growth of cities has the potential to unleash future growth. This is subject to certain caveats as highlighted by Ellis & Roberts (2015) wherein they have characterized the current state of South Asian urbanization as "messy" and "hidden". It is therefore very important to study and analyze cities' economy but unfortunately, very little information or data is available for this.

The methodology used for this study comprises the following broad steps. We use harmonized nightlights data that has been developed using two different satellite streams to allow access to a longer span covering the period from 1992 to 2020[2] to estimate the country-level measures of elasticity between GDP and nightlights growth at a global level. This step is a replication of Henderson et al., (2012), but augmented for an extended period to validate the methodology and determine the need for any modification. The second step involves a similar exercise at the national level for South Asian countries. In the third step, a similar estimation of coefficients has been done at the subnational level- the second administrative level for Pakistan and India. For both the second and third steps, we introduced a small modification in methodology by estimating the model for urban areas only. These estimates shall be used to estimate the non-agricultural component of GDP for the KP province of Pakistan. The district-level GDP is then estimated using Beyer et al., (2018)'s methodology for KP province where district-level lights' hare, and rural population share are used as weights to apportion non-agricultural and agricultural activity respectively. Here an agriculture productivity index can be used to account for variation in yields across districts. Finally, the distribution of nightlights at the district level along with the daytime satellite imagery is then used to identify urban markets following the methodology of Baragwanath et al., (2019).

This report is structured as follows. The document begins by stating the research objectives along with their relevant justifications. This is followed by a review of relevant literature. We then discuss the datasets to be used in the study and the proposed methodology for various steps of the study. Results for various steps along with explaining figures and heat maps are presented in the next section. The document concludes by highlighting the significance of the results and pointing out the limitations and avenues for extending this research.

[2] Harmonized NTL is periodically updated and is available at https://figshare.com/articles/dataset/Harmonization_of_DMSP_andVIIRS_nighttime_light_data_from_1992-2018_at_the_global_scale/9828827

02 RESEARCH OBJECYIVES

The **first objective** of this research is to benchmark economic activity at the provincial level and to measure the economic output of KP province and its districts. GDP figures and other national account statistics have been available in Pakistan at the national level but limited official counterparts of such reports have been published at the provincial level. The BOS KP, based on available data, published a review of the provincial economy including provincial GDP estimates in 2018 (see Appendix C). The report acknowledges data deficiencies and the analysis at the district level is largely descriptive. This study aims to estimate the sub-national economic development – our first objective – primarily by using the intensity of nighttime lights (hereafter NTL) as captured by the satellite images of the earth. Nighttime luminosity is a variable that has proven to be a very useful source of information in representing economic and demographic conditions of administrative units of varying sizes in several developing countries wherever such data is missing.

The **second objective** of the study is to analyze the information about KP's economy at provincial and district levels, to study the economic growth rates, districts' contribution to provincial GDP, and quantify regional disparities. Here we have also used the NTL data along with spatial distribution of population to distribute the economic activity among the third administrative tier that is the districts, following Beyer et al., (2018) approach. This helps determine the district economic share in the province as well as the balance between agricultural and non-agricultural components at the local level. Analyzing this over time provides deeper insights for policy analysis and design.

The **third objective** of the study is to identify the growth of urban centers in general and city core areas. In this part, an inter-temporal analysis of NTL distribution at various thresholds has been done to determine the overtime growth of cities and urban markets. Our report includes urban growth maps for the three most urbanized districts of the province.

The **fourth objective** of the study is to use land use/land cover maps along with NTL across the province for the last two decades. These maps can then be used to determine inter-city and intra-city growth patterns across more urbanized districts of KP province.

03 REVIEW OF LITERATURE

The current decade has witnessed a tremendous increase in the use of the intensity of lights emitted at night to estimate economic activity. Despite the sensitivity of this measure depending upon the geographical extent of analysis and the type of economic variables – output, poverty, population density, etc. it can be safely said that nighttime lights data add to our understanding of regional economic development especially in the case of developing countries, which typically lack availability of reliable data – see Chen et al. (2011). Henderson et al. (2012) is probably the best illustration in this stream of literature that appropriately explains the context around the utility of NTL in economics research. According to Google Scholar,[3] this paper has been cited 1757 times since its publication, but its popularity mainly lies in the derivation and implementation of an elegant methodological framework that brings growth in NTL and income together. Using data on more than 180 countries, the authors estimate the elasticity between the growth of GDP and growth of lights to be around 0.3 concluding that nighttime lights are best utilized when augmenting official statistics of low and medium developed countries.

[3] As of February 2021

The empirical estimation of Henderson (2012) has been performed using Defense Meteorological Satellite Program (DMSP) operated satellites using Operational Linear Scan (OLS) light sensors. This stream of night light data is available for the years 1992 till 2013. In this study, we have replicated the above estimation for an extended period till 2020 using a harmonized global nightlight dataset that combines DMSP data with Visible Infrared Imaging Radiometer Suite (VIIRS). The details about the two sources of NTL are mentioned in the Data and Methodology section.

Since the time of Henderson et al.'s (2012)'s landmark paper, a profusion of studies has emerged that have tried to capture economic activity using nighttime light data. One of the prime instances of similar methodological application is Bickenbach et al. (2016) which tries to replicate it at sub-national levels focusing on Brazil and India. Other research endeavors associating output and luminosity at sub-national levels include Bhandari et al. (2011) that finds a 0.34 percent increase in district-level GDP of India following a 1 percent increase in nighttime lights and Sutton & Costanza (2002) that focuses on gross state product of US states. Many other researchers have also commented on the power of nighttime lights for studying GDP (both level and growth) in terms of parameters often reported for checking statistical significance. For example, Doll et al. (2006) estimate an R-squared of around 0.9 on average between lights and GDP aggregated at different scales using data on 11 European countries. Beyer et al., (2018) show that the relationship between GDP levels and nightlight intensity observed elsewhere in the world also holds in South Asia's case where inverse Henderson elasticity for all countries outside of South Asia is 0.267 and that for the countries in South Asia is 0.248. Beyer et al., (2018) further point out that as intuitively understandable, nightlight intensity is more strongly correlated with economic activity in manufacturing and services industries than in the agriculture sector. They show that in South Asia the relationship for the agricultural sector is statistically insignificant whereas for manufacturing and services sectors the elasticities are 0.25 and 0.35 respectively.

The heterogeneity in the composition of economies, whether national or at the sub-national level, demands caution in the use of NTL intensity for bridging the existing data gaps. Foregoing studies highlight the importance of nighttime lights in economics literature but in the case of developing nations, other sources of information are necessary to complement the power of this unique data in estimating local economic activity. In the developing world, agriculture contributes a greater part of GDP, and agriculture, unlike manufacturing or services, is an activity that emits very lights.

The utility of NTL as a proxy for economic growth also depends on the efficiency and capacity of relevant statistical offices. Chen et al. (2011) analyze the utility of nighttime light data in estimating local economic development for various countries by assigning them to different quality levels concerning capacity for generating correct official statistics. These quality categories run from A to E with countries tagged A as ones with the highest quality data collection systems placed. The authors conclude that light intensity is a good proxy for D or E tagged countries but warn that luminosity might not give good estimates for places that have low-output-density (low-density observations have been defined as having output density less than $8,100/sq.km in the year 2000)[4]. Under these grades, Pakistan and the KP province are not graded in the low output density category.

The greater share of agriculture in the economic outlook of developing countries like Pakistan does put limitations on the use of nighttime light data. However, the availability of other sources of information has paved the way for alternate methodologies that could overcome such limitations. Ghosh et al. (2010) uses Landscan based rural population data and percentage contribution of agriculture towards national GDP along with nighttime lights and creates a gridded dataset of economic activity. On the other hand, Keola et al. (2015) employs Henderson et al. (2012) estimation strategy but runs separate regressions for agriculture and non-agriculture output growth rates. Employing a Moderate Resolution Imaging Spectroradiometer (MODIS) land cover dataset, these authors use various land classification categories as inputs in the regression explaining agriculture growth rate. It has been suggested by these authors that combined land area classified as Grassland, Cropland, and natural vegetation is a better determinant of agriculture output whereas luminosity measure should only be used for non-agriculture growth rate. Land cover data was also used by Yue et al. (2014) as ancillary to nighttime lights for studying the regional GDP of China's Zhejiang province. Keola et al. (2015) highlight a major limitation of the use of land cover data for estimating growth in agricultural output. According to them, in this method expanding agricultural land is required to increase production but it is often possible to change production levels without affecting the amount of agricultural land.

To tackle the issues just highlighted, economics literature relates NTL data with not only the aggregate GDP but also breaks it into two broad categories: agriculture and non-agriculture that comprises manufacturing and services. This exercise provides estimates of elasticities of nighttime light intensity with different types of sectoral outputs. The initial hypothesis that non-agricultural output has greater elasticity (in magnitude) relative to agriculture output has been found correct (Keola et al., 2015). The authors then employ land use and land cover data (spatial

[4] Pakistan is given a grade of C.

extent of croplands and pastures) along with national/provincial data to divide the sub-national estimates of economic activity into agriculture and non-agriculture.

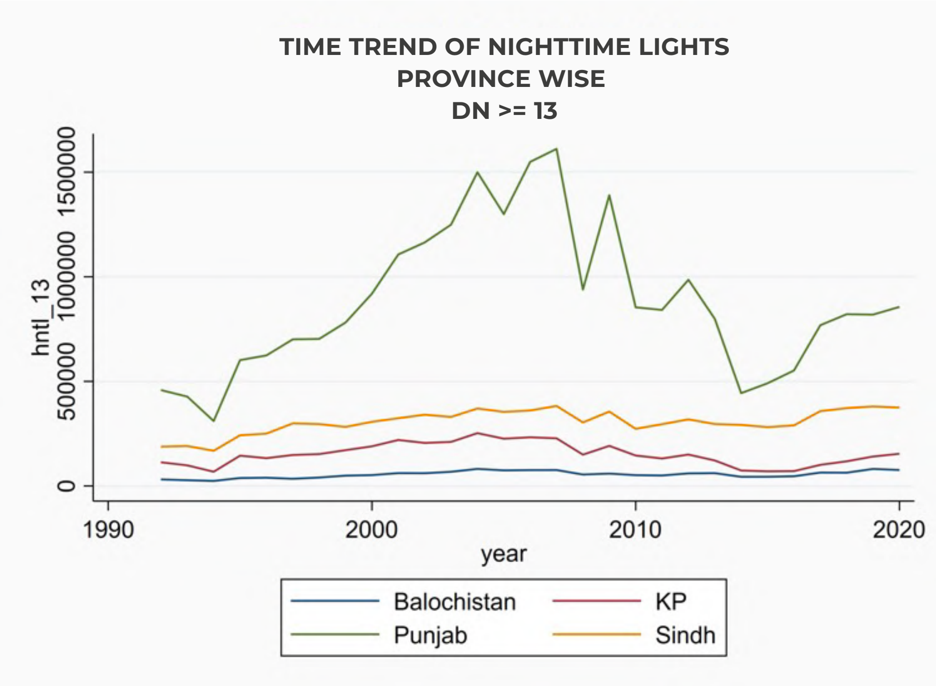


Figure 1: NTL trend among provinces of Pakistan

Available estimates of provincial GDPs for Pakistan are usually obtained by disaggregating national accounts figures at lower spatial levels and are less focused on explaining the trend or reasoning behind such values. The proxy variables used for such disaggregation are derived from datasets that are not updated frequently. Furthermore, the available analysis covers periods that are now more than a decade old – in the cases of Arby, M. F. (2008) and Bengali and Sadaqat. (2005) – or concentrate only on one region of the country – see Burki, S. J. et al. (2012). Nevertheless, previous literature – despite being outdated – provides insights that are useful for further studies on the topic.

Arby (2008) calculates provincial GDP for all four provinces of Pakistan and quantifies the contributions of labor, capital, and technology towards the output. In his estimations, the provincial gross value added in the case of primary activities that include major and minor crops, livestock, fishing, mining has been directly estimated using province-wise production data and base year prices. However, gross value added in the case of manufacturing and services has been estimated by disaggregating the national GDP based on relevant allocators. Table 1A (Appendix A) shows the share of agriculture, industry, and services sectors in all the four provinces' GDP over 45 years from 1971-2005 (Arby, 2008), and Table 2A provides the within industry sectoral shares for agriculture, manufacturing, and services. Pasha (2015) provides these estimates for more recent years, extending the work of Bengali and Sadaqat (2005). Their methodology essentially involves the identification of appropriate regional allocators for the value-added in different sectors/ sub-sectors of the economy. Three approaches adopted by them include estimation of factor incomes, production output, and expenditure incurred. With a total number of 17 sub-sectors, the factor income approach has been adopted for four sub-sectors, the output method for seven sub-sectors, and the expenditure approach for the remaining six sub-sectors.

Given the focus of the current study, Table 1 provides more recent and detailed accounts for the KP province. Two major sectoral changes that stand out in the case of KP's economy are (a) agriculture activity's contribution towards overall output has declined considerably, and (b) throughout the study, the services sector has become a major driving force of the provincial economy. However, despite the downward trend, agriculture continues to be a large proportion of the overall economy. Initially, industries and services sectors showed a growing trend but have remained stagnant over the last five years. It may be mentioned that a comparison of sectoral figures in Table 1 and Table 1A might indicate some inconsistencies, but these mainly arise due to differences in GDP estimations.

To deal with some of the above-mentioned challenges and improve the estimate of economic activity by enhancing the scope of application of remote sensing data, NTL data shall be supplemented with other relevant sources of information for grasping the local economic environment. This is because NTL data is largely considered a reliable proxy for economic activity in the industrial and services sector -both formal and informal, and to arrive at the overall economic activity it is important to add the share of agriculture. The argument regarding the use of complementary data covering agricultural output is important as this sector continues to hold a significant portion of the province's economic portfolio.

Table 1 : Sectoral Contribution of Industries in KP Province GDP (%)

	Description	2005-06	2010-11	2015-16	2019-20
A	***GVA of Agriculture Sector***	***25.05***	***22.63***	***20.09***	***19.81***
1	**Crops**	5.63	4.14	4.27	3.39
i	**Major Crops**	3.31	2.55	2.77	2.12
ii	**Minor Crops**	2.32	1.59	1.50	1.27
2	**Livestock & Poultry**	18.93	17.27	15.55	16.04
3	**Fisheries**	0.06	0.23	0.07	0.24
4	**Forestry**	0.43	0.99	0.20	0.13
B	***GVA of Industries Sector***	***21.95***	***23.18***	***24.24***	***24.38***
1	**M&Q**	2.73	5.81	6.71	6.53
2	**Manufacturing**	10.63	10.73	10.44	10.49
i	**LSM**	8.34	8.26	7.79	7.50
ii	**SHMI**	1.36	1.60	1.85	2.16
iii	**Slaughtering**	0.93	0.87	0.80	0.84
3	**Construction**	5.09	4.66	4.87	4.75
4	**EGD&GD**	3.50	1.98	2.21	2.61
i	**EGD**	3.50	1.98	2.06	2.39
ii	**GD**	0.00	0.00	0.16	0.21
C	***GVA of Services Sector***	***53.00***	***54.19***	***55.67***	***55.82***
1	**Wholesale, retail, Hotel & Restaurants**	21.76	20.69	24.00	22.70
2	**Transport Storage and Communication**	10.11	12.74	10.25	9.11
3	**Finance and Insurance**	2.19	1.98	2.26	2.88
4	**Housing**	4.23	3.87	3.50	3.40
5	**General Government**	5.71	6.01	6.83	8.22
6	**Other Private Services**	9.01	8.91	8.84	9.52
	GDP of Khyber Pakhtunkhwa	100.00	100.00	100.00	100.00

Source: KP BOS

Urban Clusters

The use of lights emitted at night is not limited to cross-country analysis for economic growth measurement. Indeed, a strand of literature is growing that targets a narrower geographical region in the country-city or district and uses nighttime lights for quantifying various social and economic indicators at this level. The research topics covered at this level are plenty but drawing inferences from NTL about measurement and dynamics of urbanization stands out– see Zhou et al. (2015); Ma et al. (2015); Tan, M. (2015); Baragwanath et al. (2019); Harari, M. (2020) among many others.

The focus of available economics literature and this study on the urban clusters is due to the immense economic potential of cities. Belleflamme et al., (2000) underscore the traditional division of these gains into two: ones resulting from the production of similar goods by firms locating nearby – localization economies; and ones coming from conducting overall economic activity in a specific geographic area –

urbanization economies. These benefits are so important that they are responsible for employment growth, technological change, innovation, and knowledge spillovers into other areas – see Glaeser et al. (1992); Henderson et al. (1995); Feldman et al. (1999); Smith, P. J. (1999).

A snapshot of the number of people and mega-urban centers – areas with a population greater than 1 million – concerning four provinces of Pakistan provides interesting insights as shown in table 2. Interestingly, KP has only 1 mega-urban center even though it houses more than 30 million people. Many countries in Western Asia and Eastern Africa have populations less than this figure. The lack of multiple cities in KP can potentially have a negative effect on the poverty levels across the province as it has been found that incidence of poverty is inversely related to the size of town or city – S. R. Hashim (2014). The NTL map (figure 2) for KP province again shows that most lights relate to the low-intensity rural category with only a few large urban centers with bright lights. Like many developing countries, Pakistan is going through the rural-to-urban migration process and a limited number of urban areas mean that the existing cities will face pressure on the supply of various amenities like housing, water, etc. Such distortions in demand and supply could result in negative social and economic outcomes for both rural and urban residents. Keeping these realities in mind, we have extracted the spatial extent of the urban centers of the province and quantify the growth of these over time. This exercise is useful in analyzing how the growth in cities matched with the growth in the local population. This information can potentially be used in determining the extent of urban sprawl or compactness- information required for efficient urban planning and city management.

Table 2: Population and Urban Areas

Province	Population	Proportion of total population	Urban Population (%)	Rural Population (%)	# Mega Urban Areas
Punjab	110,012,442	52.95	36.36	63.63	5
Sindh	47,886,051	23.05	52.02	47.98	2
KP	30,523,371	14.69	16.52	83.47	1
Balochistan	12,344,408	5.94	27.55	72.45	1
FATA	5,002,000	2.41	-	-	-
Islamabad	2,001,579	0.97	50.45	49.54	-

Source: Pakistan population census 2017

In the case of Pakistan, not many studies have used nighttime lights to explore economic and social phenomenon. To the best of our knowledge, only a few research works have employed such a novel dataset. Mahmood et al. (2017) calculate the elasticity of luminosity measure concerning large-scale manufacturing output for various districts of Pakistan covering all the provinces whereas Ali (2016) uses light at night in the context of inferring welfare implications of the power outages. More recently Beyer et al. (2018) have estimated district-level economic activity using nightlights data while capturing agricultural components through the rural population.

Our research not only augments growing literature on nighttime lights and economics but also helps to explain sub-national development in Pakistan. Our work further adds to the existing literature as we use land use/ land cover data for segregating agricultural activity. We hope that our work will help in designing better economic and public policy reforms by uncovering areas or districts in the province that lag in terms of share in modernization.

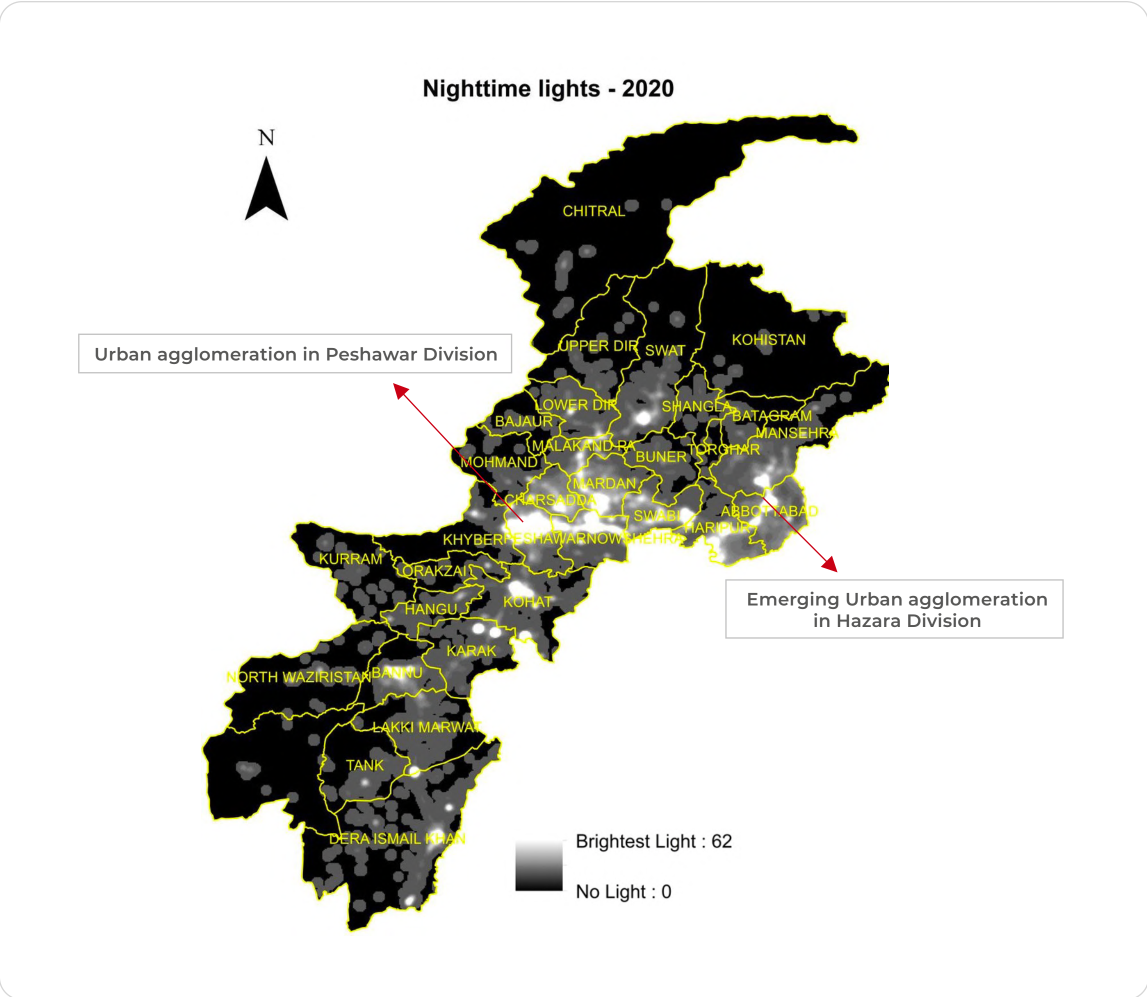


Figure 2: Nightlights in KP Province

04 DATASETS

The study has employed a variety of datasets originating from satellite imagery for nightlights and land use/ land cover, along with population census and crop surveys conducted by statistical agencies at the federal and provincial levels.

Nightlight Data

The global nightlight data used for this study is derived from two sources- nightlight imagery collected by the Defense Meteorological Satellite Program (DMSP/OLS) during the period 1992-2013, and the satellite data from Visible Infrared Imaging Radiometer Suite (VIIRS) which is available since 2012. The DMSP satellites follow an oscillating orbit with nightlight captured daily between 8:30 pm and 9:30 pm. The OLS sensor has the capability of detecting city lights, gas flares, shipping fleets, and fires. Since 1992, satellite images have been systematically digitized at the NOAA National Geophysical Data Centre (NGDC). Using these, a stable dataset cleaned from sunlight, moonlight, glare, and observations with clouds has been made publicly available. The data observation unit is a pixel with a spatial resolution of 30 arc-seconds (0.86 sq. km at the Equator) with each pixel encoded with 'digital number'(DN) measuring annual brightness on a relative scale ranging from 0 – 63.

Because of its global coverage and a temporal span extending over two decades, the DMSP data have been extensively used in studies as highlighted in the literature review section. However, the DMSP data are not publicly available after 2013 which limits its use for the current period. The subsequent nightlight data streamed from VIIRS is considered of better quality than DMSP-OLS, due to its finer spatial resolution, wider radiometric detection range, and onboard calibration (Elvidge et al. 2013). Importantly the advent of VIIRS makes it possible to continue the use of NTL after 2012. Further VIIRS data have fewer over-glow effects and are spatially more explicit in detecting light intensity variations within a city. Despite these quality advantages, the available temporal span of VIIRS is from 2012 to the present, resulting in a relatively short period for exploring the dynamics of human activities. Hence, a harmonization of NTL observations from DMSP and VIIRS data was needed and is the focus of several studies, though these efforts remained largely limited as these concentrated on local or regional scales and did not provide a global perspective (Li et al., 2020).

An annual composite at a global scale of DMSP and VIIRs datasets for the period 1992-2020 has been released and made publicly available by Li et al., (2020)[5] . This study generated an integrated and consistent NTL dataset at the global scale by harmonizing the inter-calibrated NTL observations from the DMSP data and the simulated DMSP-like NTL observations from the VIIRS data. The generated global DMSP NTL time-series data (1992–2020) shows consistent temporal trends.

[5] Available at https://www.nature.com/articles/s41597-020-0510-y.

Which Nightlights Data Should be used? Implications for Empirical Estimations

The sources and use of nightlight data have been discussed so far. Here we briefly mention the issues pointed out in the literature for each of the two data sources namely DMSP and VIIRS. Gibson et al. (2020) highlight that despite the overwhelming use of DMSP data in economics studies, certain issues need to be kept into consideration. According to them, major flaws in DMSP data include blurring, coarse resolution, top coding(saturation), and unrecorded variation in sensor amplification that limits comparability over time and space. It must be kept in mind that the original purpose of DMSP was to detect clouds for short-term Air Force weather forecasts. In comparison, the VIIRS Day-Night Band (DNB) was designed to help researchers consistently measure the radiance of light coming from the earth, in a wide range of lighting conditions with high spatial accuracy and with temporally comparable data.

Based on these considerations, Gibson et al. (2020) highlight that DMSP data do not appear to be a suitable proxy for GDP outside of cities, with a negative relationship between DMSP lights and the real GDP of non-urban spatial units, while VIIRS data positively relate to these. The negative lights-GDP relationship is especially for rural areas of low population density, in line with the conclusion of Chen and Nordhaus (2011) that luminosity data do not provide reliable estimates of economic activity for low output density regions. Gibson et al. (2020) show that in the case of Indonesia's urban sector, the lights-GDP relationship is positive irrespective of the type of night lights data used, but the relationship is twice as noisy if the DMSP data are used rather than the VIIRS data.

Further related literature shows that satellite data on night lights are not a very appropriate source of information to study low population density rural areas. The reason is that the sort of lights typically used in rural villages is not easily detectable from space, especially from DMSP data. An experiment (Tuttle et al., 2014) on testing accuracy of DMSP data, lit up previously dark areas, required a bank of 1000-W high-pressure sodium lamps to be seen with the DMSP sensors. Such lights are not found in rural villages but are more like light from concentrated streetlamps and industrial facilities, which are typically found in urban areas. While VIIRS can better detect dimly lit areas, the use of it or a harmonized NTL is likely to affect the magnitude of the regression coefficient as it shall lower the correlation between lights and economic activity. In our analysis, we find that (i) on a global scale the size of coefficient lowers in magnitude when we use VIIRS compared to DMSP, and (ii) for the KP province we observe that several DMSP pixels have zero values whereas none of the VIIRS pixels had a zero value. Further discussion on modification in our methodology to deal with the issues arising out of harmonization of two different streams of data is in the methodology and results section.

| Other Data Sources

Apart from the nighttime lights data, the other datasets used in the study are listed in Table 3 that indicates the information about the relevant agencies and the period of data used.

Table 3: Data Related Information

S. No	Dataset	Area coverage	Time coverage
1	**Population Census**	National, Provincial, District	1998, 2007
2	**Labor Force Survey**	Provincial	2003-04 to 2017-18
3	**Agriculture census** **Crop Survey**	National, Provincial, District	2010 Multiple Years
4	**National Accounts**	National	1992 to 2020
5	**Regional Accounts**	Provincial	1992 to 2005
6	**Landscan (GIS)**	100-meter grid	2000-2019
7	**MODIS (GIS)**	National, Provincial, District	2001 to 2020
8	**Census Population Grid (GIS)**	National, Provincial, District	1998
9	**World Development Indicators**	National	Multiple years
10	**CIESIN Gridded Population**	Spatial-1 km	Multiple Years

| PBS and BOS Data

The Pakistan Bureau of Statistics (PBS) which is the national statistics agency has data on population censuses, labor force statistics, and National accounts for multiple years. The provincial statistical agency that is BOS KP has data on regional accounts and the province's crop surveys.

Landscan. This is a spatial dataset representing global population distribution and is available on request for the academics from https://landscan.ornl.gov/landscan-datasets. The time covered is 2000-2019. For the census year 1998, we not only use the census figures but also use the corresponding 100-meter GIS gridded dataset made by Azar et al. (2013).

MODIS (i.e., MCD12Q1) datasets that provide classification values of vegetation cover of KP. We will use these datasets to extract land-use and land-cover states of KP for the years for which the data is available. MODIS is freely available but, unlike nighttime light data, is downloaded in terms of tiles that cover specific portions of land area. Further GIS-related processing would be required to make such tiles into a format from which we can make inferences about land cover information.

Shapefile data: This refers to GIS data in digital vector format that contains information about geometric objects. In our case, these objects correspond to the boundaries of administrative units of Pakistan at various levels i.e., national, provincial, divisions, districts, and tehsils. We use two sources of shapefile data: first is publicly accessible data available for downland at US Census Bureau's website (https://www.census.gov/geographies/mapping-files/time-series/demo/international-programs/demobase.html) and second data source comes from the Pakistan Bureau of Statistics (PBS).

World Development Indicators (WDI). We will use these to extract information about Pakistan's GDP, manufacturing output, agricultural output, and other national accounts statistics.

05 METHODOLOGY

5.1 Measuring Economic Activity Using NTL

In the **first step**, we run the basic regression of Henderson et al. (2012) as shown in equation (1) below. The idea behind this exercise is to replicate the already available empirical strategy to (a) confirm if there are any variations in the estimates when the analysis is spread over a longer period, and (b) to test the validity of the harmonized nightlights data that is derived from both DMSP and VIIRS. In the review of literature, we highlighted the output-luminosity relationship uncovered by Henderson et al. (2012) but this research work is limited in terms of the period it covers which is from 1992-2008, and its reliance on a single source of NTL data -DMSP. By re-doing the same regression analysis for the period 1992-2018 and segregating the results over data source and periods as using (i) DMSP only, (ii) VIIRS only, and (iii) harmonized data, we test the variation in elasticity between growth rates of GDP and nightlights over a longer period. Following Beyer et al., (2018) we run a similar analysis for the South Asian region to have a comparison of coefficients between the region and countries across the globe.

$$y_{jt} = \beta x_{jt} + c_j + d_t + e_{jt} \qquad (1)$$

where $y_{jt}, x_{jt}, c_j, d_t,$ and e_{jt} refer to GDP of country j at time t, NTL of country j at time t, country effects, time (year) effects, and the error term, respectively.

In the **second step** of our methodology, we shift our focus from cross-country to sub-national analysis. In other words, we employ the same regression equation (1) but with a unit of observation at the second tier of administrative hierarchy that is a province or a state. To have enough observations for the regression, we shall be using sub-national GDP estimates and nightlights data for India and Pakistan. The state-level GDP estimates are officially available for India[6] so with base year corrections, we can generate a time series for our analysis. As provincial GDP estimates are not officially published, we use data generated by Arby, M. F. (2008) and along with provincial GDP estimates prepared by BOS KP and Planning and Development Department, Punjab. Results from step 2 provide us with an elasticity between output and luminosity but now a relationship factoring in local conditions concerning electricity consumption, sectoral composition, etc. at the sub-national level of India and Pakistan.

In the **third step,** we go further into dissecting the association between output and intensity of nightlights by bifurcating economic activities into two parts: agriculture and non-agriculture. The idea behind such a division is to make sure that we link nightlights with only those activities that are responsible for such radiation. As indicated by Ghosh et al. (2010), the underlying assumption behind the use of nightlights in the economics literature is that industry and service sectors make the greatest chunk in output produced. However, this assumption might not be true in the case of Pakistan or KP province. We have seen in the estimates of Arby, M. F. (2008) that, despite showing a declining trend, the share of the agriculture sector in the provincial GDP of KP has remained above 20% during the 1971-2005 period. Insights from a recent labor force survey (see table 2 A in appendix A) also point to the importance of the agriculture sector in KP province as around 32% of employed people aged 10 and above are involved in this sector of the local economy. Keeping these facts in mind and inspired by the technique of Keola et al. (2015), we estimate two different regression equations shown as (2) and (3).

$$y_{jt}^{na} = B^{na} x_{jt}^{na} + c_j^{na} + d_t^{na} + e_{jt}^{na} \qquad (2)$$

$$y_{jt}^{a} = B^{a} l_{jt}^{na} + c_j^{a} + d_t^{a} + e_{jt}^{a} \qquad (3)$$

where y_{jt}^{na}, y_{jt}^{a}, and x_{jt}^{na} refer to non-agriculture GDP of province ***j*** at time ***t***, agriculture GDP of province ***j*** at time ***t***, and nighttime lights emitted by province ***j*** at time ***t***, respectively. l_{jt}^{na} refers to the cropland area as derived from MODIS dataset. ***c, d,*** and ***e*** are defined as before. It can be seen from equation (3) that we have excluded nightlights information in a relationship that explains the agricultural activity. Instead, we have introduced a variable

[6] http://mospi.nic.in/data

to measure land cover types that includes croplands, grasslands, and natural vegetation (Keola et al.,2015). In case results from (3) do not show a reliable estimation, we shall use the BOS-based numbers for the agriculture component of provincial GDP as is done by Arby (2008).

Our **fourth step** involves the spatial division of (a) the national GDP into second administrative level provincial figures and (b) the provincial GDP of KP province into shares of the third tier of administrative units i.e., districts. We achieve this goal by using the formula introduced by Beyer et al. (2018) for breaking down known output figures of a larger or national geographical unit into output estimates corresponding to smaller sub-national units. This approach involves retrieving estimated GDP values from regression (1) and using equation (4) for estimating district-level values.

$$\ln(GDP_{it}) = \left(\frac{light_{it}}{\sum_{i=0}^{i=I} light_{it}} * \frac{MAN_t + SER_t}{GDP_t} + \frac{rpop_{it}}{\sum_{i=0}^{i=I} rpop_{it}} * \frac{AGR_t}{GDP_t}\right) * PGDP_t \qquad (4)$$

Where GDP_t, $PGDP_t$, $light_{it}$, and $rpop_{it}$ refers to actual national (provincial) GDP, official/estimated national (provincial) GDP, night-time lights emitted by province (district) ***i***, and rural population the in province (district) ***i***, respectively. On the other hand, MAN_t, SER_t, and AGR_t are the manufacturing, services, and agriculture components in the official/estimated provincial GDP. The relevant time frequency for all calculations here is the annual basis.

5.1.1 Masking Gas Flares

District level estimates for GDP obtained through (4) can return spurious results if the local NTL sum is influenced by the presence of gas flares from oil and gas fields. Examining the NTL maps and list of operational gas fields in KP province, we find that results in Kohat and Karak districts are influenced by additional lights due to gas flares. Using the gas flares database shapefile (Elvidge et al., 2016) we mask the additional lights shown in figure 3 and replace the NTL with mean district lights.

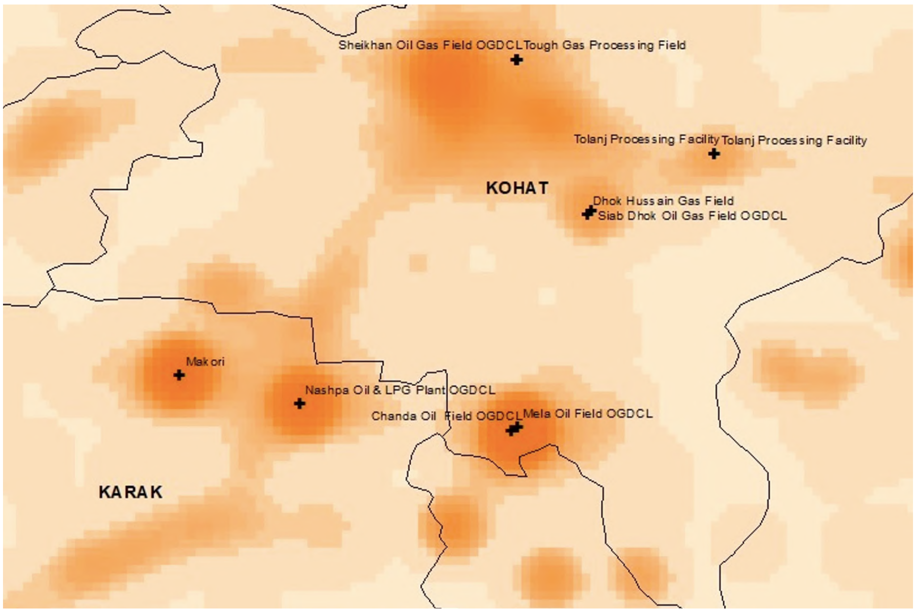


Figure 3: Oil and Gas Fields in Kohat and Karak Districts

5.2 Extracting Cities or Urban Areas

The utility of satellite data is not just limited to the study of the size and growth of economic activity. Another important application of such a novel dataset is related to extraction of geographical boundaries of areas i.e., urban boundaries of districts, etc. considered important concerning various social-economic roles they perform in a larger system of cities. Previously applied methodologies in the literature differ on some levels but all boil down to inferring the spatial extent of an area from the spatial extent of lights emitted at night.

To determine the extent of urban markets, we use two approaches namely the threshold approach and the brightness gradient approach. We shall discuss each of these here in detail. In the “thresholding” technique - see Imhoff et al. (1997); Amaral et al. (2005); Gibson et al. (2020), and more recently Baragwanath et al. (2019); Harari, M. (2020)

contiguous pixels with DN values exceeding the chosen threshold are taggeiss urban and hence constitute a city. Keeping the same DN threshold value during the whole period, the technique estimates the size of each targeted city and derives the annual growth rates of expansion. The quantitative connections between night-time light brightness and urbanization parameters for a given target region are usually constructed by counting the area covered by lit pixels (i.e., the area of lighting) or summing total radiances of pixel light (i.e., the sum of night-time radiance) exceeding the pre-chosen threshold. Both these indices enable us to examine overall responses of night-time light to regional socioeconomic activity over time and to compare the interregional differences of urban development (Doll et al., 2006; Elvidge et al., 2001; Ghosh et al., 2010; Zhang & Seto, 2011).

The thresholding technique using NTL data shows great potential in urban mapping as historical records of NTL dating back to 1992 have been made publicly available. However, the harmonized NTL (1992-2020) used in this study gets the first component of data based on DMSP for the years 1992-2013. NTL data from 2014-2020 is purely based on VIIRS and the joint use of DMSP and VIIRS for urban detection can potentially generate inconsistent estimates. The use of NTL stream based on DMSP for urban extent mapping needs to be adjusted for two important factors that are also considered as a weakness of the DMSP series: saturation of luminosity and the blooming effect. The DMSP-OLS based NTL indicates artificial lights at night in clear weather conditions, however, the data value of urban centres is too bright and tends to saturate due to the limitation of radiometric range (Imhoff et al., 1997). Additionally, the light scattering effect makes some places luminous without light (Small and Elvidge, 2013). Therefore, urban mapping using NTL data with threshold methodology can return spurious results unless the saturation and blooming effects have been duly accounted for in the estimation. These two practical drawbacks pose challenges for partitioning and classifying night-time light images in association with different degrees of human activity over space. Besides to deal with the inter-calibration of time series and their temporal pattern adjustment require the use of dynamic optimal thresholds to avoid varying exaggeration of NTL data in the large-scale urban mapping. This is even more important if we are interested in understanding intraregional spatial variations of night-time light signals, which are likely correlated with the fluctuations of spatial patterns of urbanization dynamics at a local scale.

To depict the pixel-level fluctuations of night-time light (NTL) across human settlements, we use the brightness gradient (BG) method proposed by Ma et al. (2015) to measure spatial changes in artificial night-time lighting signals. BG is defined as the rate of maximum change in NTL from the site to its neighbors (rise) across the corresponding geographical span (run). Thus, the output BG can be calculated as:

$$BG = \frac{NTL\ Rise}{Geographical\ Run} \tag{5}$$

For gridded NTL, the pixel-level BG is regarded as a measure of the maximum change in NTL over one-pixel size between the grid cell and its neighbor cells. Ma et al. (2015) use the average maximum technique (Burrough &McDonell, 1998) to calculate BG for each grid cell. This algorithm applies weighting coefficients that are proportional to the reciprocal of the square of the distance from the processing grid cell for the nearer NTL values. Mathematically, the pixel-level BG for a pixel with reference to its neighboring pixels can be estimated by using the rates of change of NTL in the horizontal (dNTL/dx) and vertical (dNTL/dy) directions from the central grid cell to its eight adjacent grid cells as follows:

$$BG = \sqrt{(\frac{dNTL}{dx})^2 + (\frac{dNTL}{dy})^2} \tag{6}$$

Ma et al. (2015) analyze the brightness gradient across the Chinese city Huainan for 2012. The intra-urban comparisons reveal that grid cells with low BG values are likely to be found in both the central region of the urban area with intensified human activity and high NTL and the rural region with less human activity and low NTL. In contrast grid cells with high BG values commonly appear in the urban-rural transition zone with medium NTL.

The authors further find that the quantitative relationship between BG and NTL is not linear for cities that are not small, isolated, and circular and there is a wide range of brightness gradients associated with the brightest and dimmest pixels. Therefore, the relationship between the pixel-level NTL and BG for a given city can be fitted by a quadratic polynomial:

$$BG = aNTL^2 + bNTL + c \tag{7}$$

where a, b and c represent fitting coefficients. To examine the prevalence of this statistical relationship, we performed quadratic regressions for the pixel-level NTL and BG at a local urban scale using the least-squares method. The results indicate that the quadratic relation is valid for large cities where there is considerable core urban area but does not represent the correct model for smaller cities with scattered urbanization.

As such we shall apply the threshold method to the top eight urban districts but restrict the BG approach to three/four districts where robust quadratic relationships between the pixel-level NTL and BG across KP districts allow us to quantitatively classify the urban area according to conspicuous differences in NTL at a local scale.

For the BG approach, following Ma et al. (2015) we spatially subdivide a city into five different sub-regions involving: low (night-time lights range from NTL_0 to NTL_1), medium-low (from NTL_1 to NTL_2), medium (from NTL_2 to NTL_3), medium-high (from NTL_3 to NTL_4) and high (from NTL_4 to NTL_5) night-time lighting areas as shown in figure 4. The turning point of the parabola in which the maximum BG and medium NTL, typically occur in the transition zone of urban area and suburban.

This methodology implies that for a given city with a significant quadratic relationship between pixel-level night-time light and spatial brightness gradient, the partition intervals for night-time light imagery are determined by the range of NTL and the fitted coefficients of the quadratic function. This also means that no empirical threshold and parameterization of the split points are required for partitioning night-time light imagery across different cities. Moreover, the primary advantage of the quadratic curve-based partitioning method for DMSP/OLS night-time light imagery is that it reduces the impacts of various urban development patterns because the partition method is based upon the relative quantitative relationship between the pixel-level NTL and BG at a local scale.

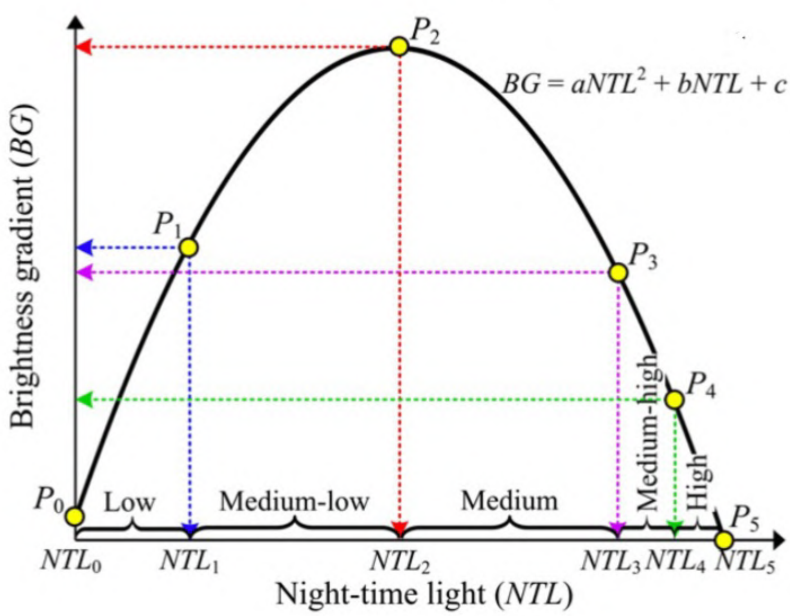


Figure 4: Sub-region Partitioning based on the Quadratic Model

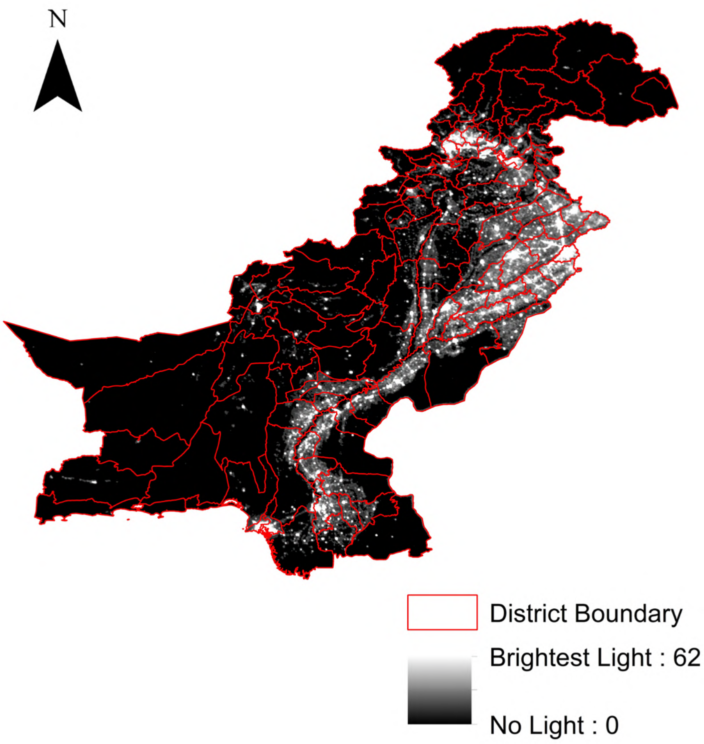
Source: Ma et al. (2015)

5.3 Land Use & Land Cover Maps for KP

Finally, we compare the spatial extent of NTL-based markets with those formed by daytime satellite imagery. The daytime satellite imagery is classified to detect built-up urban land using the MODIS Land Cover Type product (MCD12Q1) developed by Sulla-Menashe and Friedl (2018). Combining daytime imagery to measure the boundary of markets with NTL to measure the intensity of economic activity is a robust approach as the two datasets are publicly available for around two decades. Land cover is the physical material at the surface of the Earth whereas land use is a description of how people utilize the land. There are two primary methods for capturing information on the land cover: field surveys which might be limited in terms of area covered and frequency of data collection and from remotely sensed imagery. For provincial-level analysis, remote sensing appears to be a feasible way to complement the field surveys.

The MODIS product that we use here is MDC12Q1which is a yearly land cover product. Data are presented in tiles of approximately 1200 X1200 km with 500 m nominal spatial resolution. As we aim to study the changes in land use over time, we focus on its four broad categories: (i) agriculture, (ii) water bodies, (iii) forests and (iv) built-up areas. For the agricultural sector, we have extracted and combined the classes of most relevance as highlighted by Keola et al. (2015). These classes are numbered as 10 (grasses/cereals), 12 (croplands), and 14 (cropland/natural vegetation mosaic). The built-up area signifies urban land use and over time its variation can be used to measure urban growth.

NIGHTTIME LIGHTS IN PAKISTAN 1998



NIGHTTIME LIGHTS IN PAKISTAN 2020

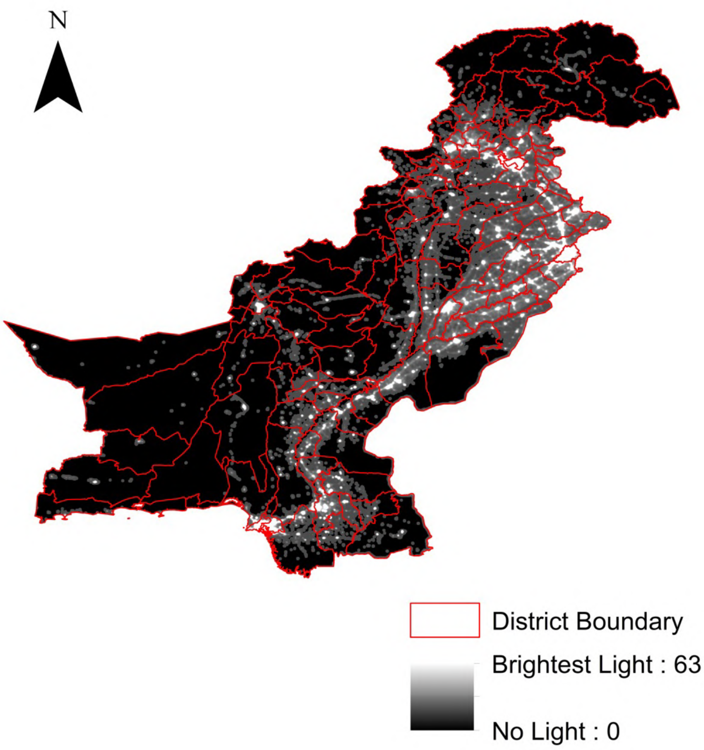


Figure 5: Pakistan Map with NTL and District Boundary

06 RESULTS

6.1 / 6.1.1

Economic Growth & NTL

Replication of Henderson (2012)'s Approach at Country-level, Global Scale

Using the econometric model specified in (1) the relationship between economic activity and nightlight intensity is typically estimated assuming a constant elasticity called the inverse Henderson elasticity. As can be seen from Table 4, the relationship between GDP levels and nightlight intensity (SOL is the sum of lights) has been estimated using different data sources while breaking up the periods. Column 1 shows simple replication of Henderson et al.'s (2012) results using the data provided by the same authors. Column (2) then uses World Development Indicator's country-level real GDP in local currency units (LCU) which has the base year of 2010 (compared to the base year 2000 for Henderson et al.) and harmonized NTL. Subsequent columns (3-5) differ in terms of NTL pixel area adjustment[7] as suggested by Henderson et al. (2012) and time-period focus. Our analysis indicates that this is because of differences in variation of satellite scanners used to collect DMSP and VIIRS data.

[7] Data for lights are reported on a latitude-longitude grid. An arc-second is one-sixtieth of an arc-minute, which is one-sixtieth of a degree of latitude or longitude. Because of the curvature of the Earth, grid cell size varies in proportion to the cosine of the latitude. Thus, all grid cell sizes are reported at the equator; sizes at other latitudes can be calculated accordingly (Elvidge et al. 2004).

As VIIRS data has been collected through more sensitive sensors, it picks rural lights as well where otherwise economic activity is much lower than urban areas. As such use of NTL for years when the data source is VIIRS lowers the coefficient size. We will cover more discussion on this in subsequent estimations where we try control for this issue.

Table 4: Regression results for GDP on SOL per area - Global Replication at Country level

	(1)	(2)	(3)	(4)	(5)
	1992 - 2008	**1992 - 2008**	**1992 - 2008**	**2009 - 2018**	**1992 - 2018**
(Log) SOL/Area	0.28*** (0.03)	0.23*** (0.03)	0.23*** (0.03)	0.04*** (0.01)	0.11*** (0.02)
Observations	2961	2893	2893	1794	4711
Countries	184	181	181	181	184
R-squared	0.77	0.78	0.78	0.62	0.79

Column 1 uses Henderson's GDP and Henderson's SOL per area
Column 2 uses recent WDI's GDP and harmonized SOL per area
Columns 3, 4, 5 uses recent WDI's GDP and harmonized SOL per area derived using Henderson et al. (2012) pixel area adjustment

Columns 1-3 confirm that our coefficient is very close to Henderson et al. and hence validates the data and estimation process. As mentioned earlier, the use of VIIRS for the harmonized NTL lowers the size of the coefficient as can be seen in columns 5 and 6. This is discussed in more detail in the following sub-section.

6.1.2 Replication of Henderson's Approach at regional-level, South Asian Focus

In this sub-section, we replicate the global analysis while restricting the countries' samples to the South Asian region. As mentioned earlier we can see from table 5 that the coefficient magnitude falls drastically when VIIRS data is being used to generate NTL (see columns 4 and 5).

Table 5: Regression results for GDP on SOL per area (South Asia - Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka)

	(1)	(2)	(3)	(4)	(5)
	1992 - 2008	1992 - 2008	1992 - 2008	2009 - 2018	1992 - 2018
ln lights/area	0.20**	0.20*	0.20*	-0.01	0.09
	(0.06)	(0.09)	(0.09)	(0.02)	(0.08)
Observations	116	116	116	70	186
Countries	7	7	7	7	7
R-squared	.97	.97	.97	.97	.96
Time F.E.	Yes	Yes	Yes	Yes	Yes
Country F.E.	Yes	Yes	Yes	Yes	Yes

To deal with this issue we segregate the rural area from the urban area as this was causing the issue in the correct model estimation. As discussed earlier in the data section, this is because of the difference in the sensitivity of the sensors and here the coefficient value dips as VIIRS assigns a positive value to rural areas as opposed to DMSP that usually assigns a zero (DN=0) value to rural area pixels. Hence there is a need to delineate urban from rural areas and restrict the regression to urban areas only. Following Roberts (2018) which in turn refers Ellis and Roberts (2016), we use a DN threshold such that pixels with DN > 12 are classified as urban and all other pixels are classified as rural. Roberts (2018) justifies this choice using figure 6 which shows this for South Asia where histograms of observed DN values in areas classified as urban (red) and agricultural (blue) in a conventional land-use map for the region. As can be seen, DN values greater than 13 tend only to be observed in urban areas. Using DN=13 we segregate the urban and rural areas and then restrict our model estimation to urban areas only.

AVERAGE NTL DN: CROPLAND & URBAN NTL

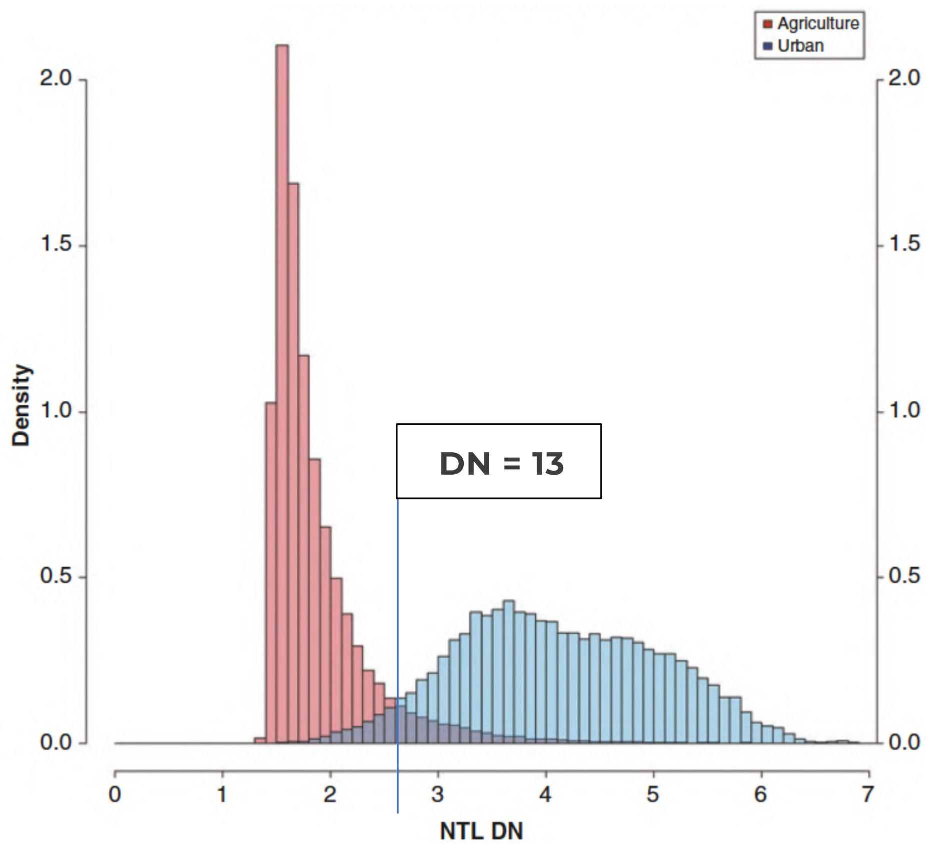


Figure 6: DN values Distribution

(Source: Roberts (2018) DN values are measured on a natural log scale.

The regression results are shown in Table 6 where the coefficient for the entire period matches that of Henderson et al. (see columns 2 and 5).

Table 6: Regression results for GDP on SOL per area for South Asia (South Asia - Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka)

	(1)	(2)	(3)	(4)	(5)
		1992 - 2008		**1992 - 2018**	
	0≤DN≤63	**0≤DN≤63**	**DN ≥ 13**	**0≤DN≤63**	**DN ≥ 13**
In lights/area	0.16* (0.07)	0.21* (0.09)	0.17** (0.05)	0.09 (0.08)	0.21*** (0.03)
Observations	102	102	102	162	162
Countries	6	6	6	6	6
R-squared	.97	.97	.97	.96	.98
Time F.E.	Yes	Yes	Yes	Yes	Yes
Country F.E.	Yes	Yes	Yes	Yes	Yes

Column 1 uses Henderson's GDP and Henderson's SOL per area
Columns 2 and 3 use recent WDI's GDP and harmonized SOL per area
Columns 4 and 5 use recent WDI's GDP and harmonized SOL per area

6.1.3 Sub-National Regression for India and Pakistan

As the intent of this study is to estimate sub-national GDP at province and district level, we run the regression using state /province level GDP (SGDP) for India and Pakistan as using just Pakistan data was not sufficient in terms of number of observations. Table 7 provides the regression coefficients where column 5 is of particular interest as it estimates the model for the non-agriculture economic activity taking place in urban areas. The coefficient of night lights is 0.10 and it is statistically significant at the 10 percent level. The magnitude of the coefficient is only half of what we found at the global level and the South Asian region. Our results match what has been found by Prakash et al. (2019). The columns with DN values larger than 25 have been added to indicate that arbitrarily increasing the threshold will lead to information loss.

Table 7: Sub-National Regression

	(1)	(2)	(3)	(4)	(5)	(6)
	SGDP			Industry+ Services		
	0 ≤ DN ≤ 63	**DN ≥ 13**	**DN ≥ 25**	**0 ≤ DN ≤ 63**	**DN ≥ 13**	**DN ≥ 25**
ln lights/area	0.05 (0.06)	0.07 (0.06)	0.06 (0.04)	0.08 (0.06)	0.10* (0.05)	0.07** (0.03)
Observations	606	605	592	606	605	592
States	31	31	31	31	31	31
R-squared	0.95	0.95	0.95	0.96	0.96	0.97
Time F.E.	Yes	Yes	Yes	Yes	Yes	Yes
State F.E.	Yes	Yes	Yes	Yes	Yes	Yes

6.1.4 Benchmarking Economic Activity at Province Level

The focus of the current study is on the estimation of economic activity in KP province, but we also use the NTL data to estimate province level GDP using Beyer et al. (2018) methodology. In this approach, we use the province-level share of Sum of Lights to estimate the manufacturing and services share and provincial rural population share to estimate the agricultural contribution of a province towards the national GDP.

In table 8 we present province-level shares along with other similar estimates based on studies of Arby (2008), Pasha et al. (2015) as well as figures obtained from BOS of Punjab and KP Province. The figures used here are based on 2005-06 constant prices. Our numbers indicate a higher share of Punjab, and Balochistan and a lower share of Sindh, whereas the KP share in national GDP is largely consistent with other studies and estimates. The provincial GDP figures 2005-06 constant prices also match with the provincial BOS estimates. Province level GDP estimates using NTL and shares in national GDP are reported in tables 9 and 10.

Table 8: Provincial Share in National GDP

Source	Punjab		Sindh		KP		Balochistan	
	1999-00 (%)	2017-18 (%)	1999-00 (%)	2017-18 (%)	1999-00 (%)	2017-18 (%)	1999-00 (%)	2017-18 (%)
NTL Data (Beyer, 2018)	58	56	24	27	14	10	4	6
Arby (2008)	53	-	30	-	12	-	6	-
Pasha (2015) [8]	55	54	30	30	11	13	4	3
PBS, BOS Punjab, and KP	-	48	-	-	-	10	-	-

Source: Author's Estimation

6.1.5 Estimated GDP for KP Province

To ensure the reliability of results, this study uses two methodologies to estimate province-level GDP for the KP province. In the first approach, we use coefficients from the regression on subnational GDP in India and Pakistan to estimate province-level GDP (manufacturing and services industries only) for the KP province. Due to data limitations, this is done for the period 2004-05 to 2013-14. Plotting estimated and actual values for the non-agricultural component, Figure 7 shows that both estimates are very close, and which is an indication of the robustness of our methodology, and the official data produced by KP BOS. The difference between the

[8] Figures mentioned against Pasha(2015) in 2017-18 column pertain to 2014-15

actual and estimated GDP can indicate the presence of an informal economy, though differences in methodology can also be contributing to differences. Between 2009-2014, this difference has varied in the range of 5-11%.

In the second method, we distribute official national GDP into provincial shares using their shares in nightlights luminosity and rural population numbers. To make up for the missing data in the earlier approach and to avoid weaknesses of nightlights data, this exercise is carried out for the years 2013-14 to 2019-20. The results are summarized in tables 9 and 10.

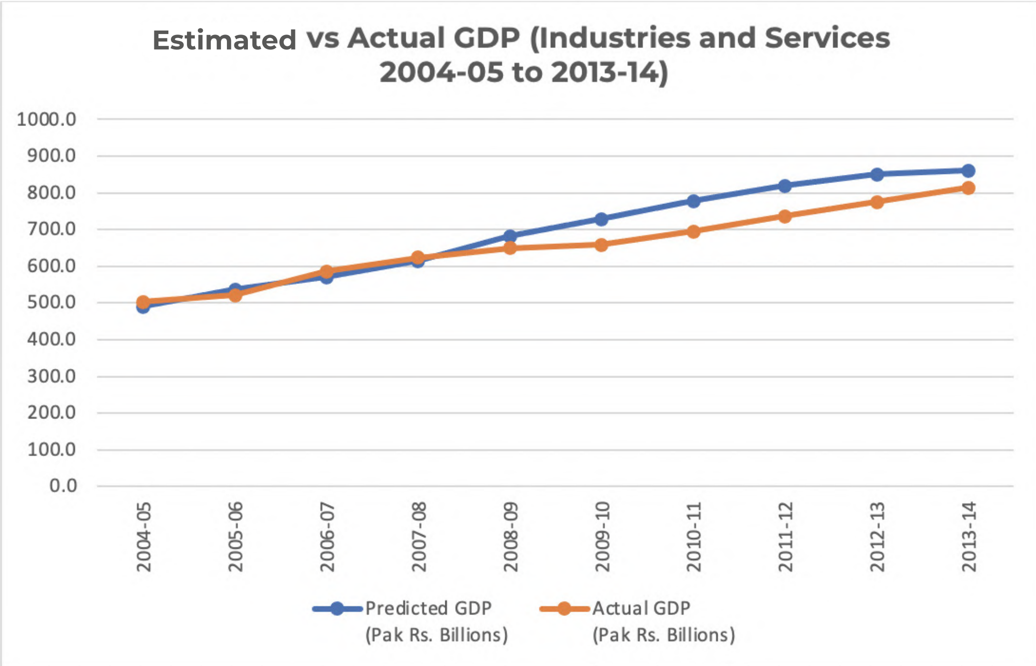


Figure 7: Actual and Estimated real GDP at constant prices for Industries and Services Sector-KP Province

Source: Author's Estimation

Table 9: Provincial Share in National GDP (in percentage)

	Using the distribution of national data in proportion to NTL and rural population shares				Official BOS data
Fiscal Year	**Baluchistan**	**Sindh**	**Punjab**	**KP**	**KP official**
2013-2014	5.31	27.69	56.58	10.42	9.70
2014-2015	5.27	31.19	54.41	9.13	9.67
2015-2016	5.17	29.75	56.37	8.71	9.87
2016-2017	5.23	27.22	58.47	9.09	9.80
2017-2018	5.10	25.98	59.36	9.57	10.05
2018-2019	5.32	24.84	59.92	9.93	9.88
2019-2020	5.33	23.70	60.58	10.39	9.66

Source: Author's Estimation

Focusing on recent years, it is observed that the difference between official and estimated GDP figures was very small in 2018-19 but has again increased in 2019-20. These differences between official and estimated GDP figures arise because of (a) differences in methodologies and (b) annual variations in the scale of informal economic activity. The presence of these differences highlights the need to focus more on capturing the informal economic activity in the official estimates.

Table 10: Provincial GDP (NTL and Official in billion PKR)

	Using the distribution of national data in proportion to NTL and rural population shares					
Fiscal Year	**Baluchistan**	**Sindh**	**Punjab**	**KP**	**KP Official GDP**	**Difference (NTL – official) (%)**
2013-2014	542.89	2,828.90	5,780.96	1,064.31	991.54	7.3
2014-2015	560.62	3,316.10	5,784.33	970.60	1,028.10	-5.5
2015-2016	574.62	3,306.78	6,266.65	968.75	1,097.50	-11.7
2016-2017	611.27	3,183.47	6,839.49	1,062.70	1,146.20	-7.2
2017-2018	629.07	3,206.73	7,327.12	1,181.35	1,240.04	-4.7
2018-2019	669.99	3,129.43	7,550.00	1,251.24	1,245.44	0.4
2019-2020	667.86	2,971.89	7,598.40	1,303.69	1,211.20	7.6

Source: Author's Estimation

6.1.6 District-level GDP for KP Province Using Beyer et al. (2018)

Applying the methodology of Beyer et al. (2018), the provincial GDP has been distributed among the districts based on NTL and rural population data. The results in table 11 are based on the district's proportion of SOL (calculated using DN above 12) in the total provincial SOL along with the district's rural population share which is used to proxy for the agriculture's contribution to the local GDP. It may be noted that the rural population numbers have been obtained by overlapping NTL pixels with the Landscan data and counting people in rural pixels (ones with DN values up to 12).

Table 11 indicates that whereas the overall economy of Peshawar district is the largest in the KP province, Haripur district has the highest per-capita income followed by Nowshera and Abbottabad. In terms of per-capita GDP, the lowest figure is for Kohistan district preceded by Torgarh district. We aim to validate the per-capita estimates with district-level Multidimensional Poverty Index (MPI) estimates made by UNDP[9] Pakistan. For reference, Multidimensional Poverty Indices use a range of indicators to calculate an overall poverty figure for a specific population, whereby a larger number indicates a higher level of poverty. This figure considers both the proportion of the population that is deemed poor, and the 'extent' of poverty experienced by these 'poor' households, using the Alkire and Foster (2011) 'counting method'.

The latest district ranking for the incidence of Multidimensional Poverty is given in table 13 A and there is much similarity there with our findings- Haripur has the lowest MPI and Kohistan has the highest MPI. Similar figures indicating district-wise educational outcomes (see table 13-B) in terms of literacy rates and enrollment gender parity index (GPI) validate our findings. Literacy rates are defined as proportion of the population with the ability to read and write. GPI is the ratio of the female to male gross enrollment rate at the primary level. When the GPI shows a value equal to one, boys' and girls' enrollment are equal in terms of gross enrolment rates. A value less than one indicates that proportionately fewer girls than boys are enrolled at the primary level. Further, the share of Peshawar in the provincial economy has grown from 16% to 19% between 2005-06 to 2019-20. The district-wise ranking is also shown using a heatmap in figure 8. The district economies have also been ranked based on their overall size. The last column of the table indicates if there is a change in the rank of the district between 2005-06 and 2019-20.

[9] https://www.pk.undp.org/content/pakistan/en/home/library/development_policy/Multidimensional-Poverty-in-Pakistan.html

Table 11: Economic Ranking of KP Districts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
District	**GDP (Rs. in millions)**		**GDP per capita**	**GDP share (%)**		**Rank**		**Change**
	2005-06	**2019-20**	**2019-20**	**2005-06**	**2019-20**	**2005-06**	**2019-20**	
Peshawar	106853	228583	53543	16	19	1	1	Same
Nowshera	55821	142301	93708	8	12	4	2	Up
Mardan	69816	113350	47765	10	9	2	3	Down
Haripur	35655	104021	103707	5	9	7	4	Up
Abbottabad	23240	81088	60835	3	7	11	5	Up
Swabi	53070	74478	45843	8	6	5	6	Down
D. I. Khan	33709	72156	44345	5	6	8	7	Up
Charsada	57788	56273	34817	9	5	3	8	Down
Kohat	29184	52264	52586	4	4	9	9	Same
Mansehra	23115	51039	32791	3	4	12	10	Up
Bannu	36221	45886	39289	5	4	6	11	Down
Swat	28001	40085	17355	4	3	10	12	Down
Malakand	12973	33430	46411	2	3	15	13	Up
Lakki Marwat	16627	28166	32146	2	2	14	14	Same
Lower Dir	16976	23418	16308	3	2	13	15	Down
Karak	12491	14650	20741	2	1	16	16	Same
Tank	7328	8617	21988	1	1	21	17	Up
Hangu	7418	8301	16000	1	1	20	18	Up
Buner	10547	8107	9034	2	1	17	19	Down
Upper Dir	8457	6665	7042	1	1	19	20	Down
Shangla	3927	6403	8448	1	1	24	21	Up
Batagram	4758	4932	10348	1	0	23	22	Up
Chitral	6551	4632	10355	1	0	22	23	Down
Kohistan	8947	1544	3032	1	0	18	24	Down
Torgarh	1221	811	4733	0	0	25	25	Same

Source: Author's Estimation

Further, table 12 indicates the district-wise share in the industry and manufacturing sectors of the economy (based on the NTL share) and agriculture component of provincial GDP. The choice of years is restricted to ones with official population censuses to ensure better reliability. The NTL share indicates the proportion of non-farm economic activities situated in a particular district whereas the rural population share indicates the proportion of agricultural activities in that district.

Table 12: Economic ranking of districts

District	Nightlights Share (%)	Nightlights Share (%)	Rural Population Share	Rural Population Share
	1998	2017	1998	2017
Peshawar	23.13	27.48	5.93	7.85
Nowshera	11.02	13.84	3.68	4.03
Haripur	6.46	8.94	3.46	2.97
Mardan	12.23	8.88	6.62	6.60
Abbottabad	3.06	7.57	4.10	3.55
D. I. Khan	3.60	5.79	4.13	4.32
Kohat	3.44	4.37	2.33	2.47
Bannu	2.49	4.30	3.58	3.82
Swabi	10.92	3.82	4.81	4.61
Charsada	12.77	3.48	4.71	4.60
Mansehra	0.74	2.14	5.21	4.82
Swat	1.14	1.75	6.16	5.51
Khyber	3.40	1.47	2.80	3.04
Malakand	1.49	1.22	2.32	2.23
Hangu	0.65	1.15	1.42	1.42
Lakki Marwat	0.75	1.11	2.52	2.69
Karak	0.24	1.03	2.29	2.24
South Waziristan	0.06	0.92	2.44	2.30
North Waziristan	1.05	0.33	2.02	1.84
Mohmand	0.44	0.20	1.90	1.59
Tank	0.51	0.11	1.15	1.18
Lower Dir	0.27	0.08	3.82	4.77
Chitral	0.00	0.00	1.64	1.36
Bajaur	0.12	0.00	3.38	3.73
Kohistan	0.00	0.00	2.68	2.68
Batagram	0.00	0.00	1.75	1.63
Buner	0.06	0.00	2.87	3.06
Upper Dir	0.00	0.00	3.14	3.08
Torgarh	0.00	0.00	0.99	0.59
Shangla	0.00	0.00	2.47	2.59
Orakzai	0.00	0.00	1.28	0.87
Kurram	0.00	0.00	2.40	1.98

Source: Author's Estimation

Table 13 A: District Level MPI (2014-15)

District	MPI
Haripur	0.110
Peshawar	0.148
Abbottabad	0.149
Mardan	0.153
Nowshehra	0.168
Malakand	0.171
Chitral	0.194
Lower Dir	0.194
Mansehra	0.204
Swabi	0.210
Charsadda	0.213
Kohat	0.238
Karak	0.253
Hangu	0.271
Swat	0.271
Bannu	0.289
Lakki Marwat	0.32
D.I. Khan	0.362
Buner	0.373
Tank	0.385
Batagram	0.422
Shangla	0.438
Upper Dir	0.443
Torgarh	0.571
Kohistan	0.581

Source: UNDP Pakistan

Table 13-B District Level Literacy Rates and GPI

District	TOTAL			GPI
	Male	Female	Total	
Khyber Pakhtunkhwa	68	32	50	0.5
Abbottabad	78	54	66	0.7
Bannu	70	24	46	0.3
Batagram	55	13	31	0.2
Buner	57	19	37	0.3
Charsada	66	28	48	0.4
Chitral	66	38	50	0.6
D. I. Khan	59	22	41	0.4
Dir Lower	69	30	48	0.4
Dir Upper	53	19	36	0.4
Hangu	72	20	42	0.3
Haripur	77	51	64	0.7
Karak	87	40	63	0.5
Kohat	79	37	56	0.5
Kohistan	54	3	31	0.1
Lakki Marwat	70	17	48	0.2
Malakand	71	39	55	0.5
Mansehra	78	46	62	0.6
Mardan	70	33	52	0.5
Nowshera	73	37	55	0.5
Peshawar	66	34	50	0.5
Shangla	40	6	24	0.2
Swabi	72	35	53	0.5
Swat	70	30	50	0.4
Tank	67	21	44	0.3
Tor Ghar	25	2	12	0.1

Source: Development Statistics 2019

RANKING OF KP DISTRICTS W.R.T GDP
YEAR 2019/20

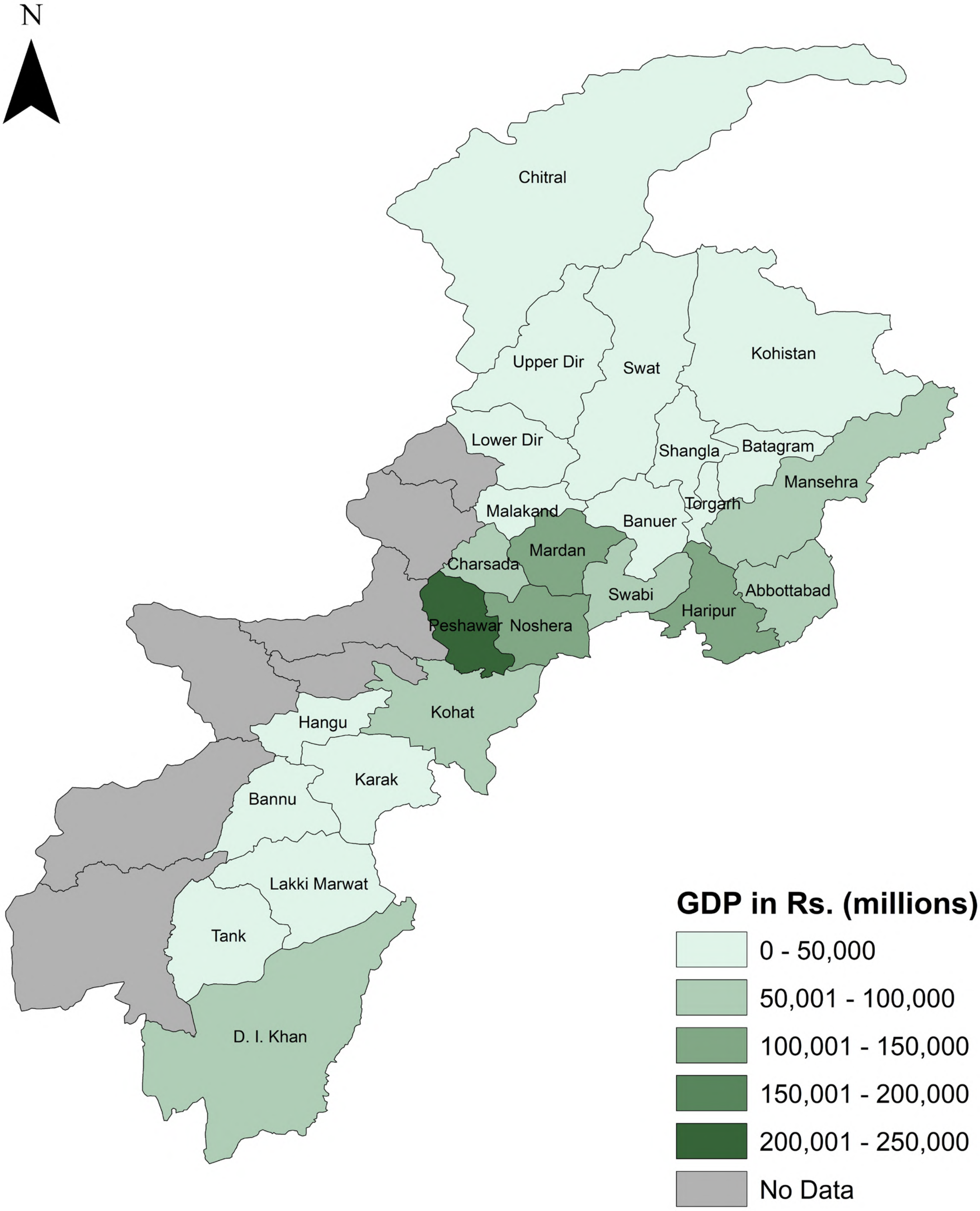


Figure 8: District GDP Ranking for 2019-20

6.1.7 Regression results for Agriculture -GDP Using Modis Datas

Table 14: Province Level Agriculture Model Estimation

Agricultural Area (log)	0.60*** (0.08)	0.61*** (0.08)	0.68*** (0.02)
Observations	232	232	232
Countries	29	29	29
R-Squared	0.16	0.22	0.90
Time F.E.	No	Yes	Yes
State F.E.	No	No	No

Agriculture GDP at 2011-12 base year.

6.2 Detecting Urbanization and Urban Markets

6.2.1 Threshold Method

The two-pronged methodology for detecting urban growth and urban markets has been discussed in section 5. In this study, we applied the “thresholding” method on eight urbanized districts of KP using DMSP data for the years 1998 and 2013 and the “Brightness Gradient” approach on the more urbanized districts employing the high-resolution VIIRS data for the years 2014 and 2020. The choice of the years was made to ensure that NTL stream is based on a single source and hence the analysis does not suffer on account of data source inconsistencies. Moreover, these years nearly coincide with the national population census so that numbers could be conveniently cross validated. Table 15 provides the results using which we can determine the growth of urban markets in the eight most urbanized districts of the KP province (based on the population census of 2017). The NTL numbers provide useful insights in terms of threshold DN values, overall urban growth, growth of core regions in the districts, and population proportions inhabiting the urban and core regions. The figures have been derived using the district-specific threshold values of NTL distribution at 75th, 85th, 90th, 95th, and 99.5th percentiles of NTL distribution. The spatial population distribution for the relevant pixels has been taken from CIESIN (2018) where the population for urban pixels is counted as shown in figure 9. The sum of the population for all pixels tagged as urban gives the respective population share. For each of the percentiles, the table provides the threshold DN (minimum DN value), the number of pixels underlying there in the corresponding year, and the underlying pixels based on 1998 threshold DN values. Columns 16-20 provide the percentage of urban population corresponding to the district-specific percentile of DN distribution.

It can be seen from columns 16-20 that there does not exist any single threshold DN value that can be used to extract urban areas of all the districts considered in Table 15. In the case of Peshawar, threshold DN value corresponding to 75th percentile gives the proportion of urban population estimate i.e., column 13, closer to census proportion estimate i.e., column 21. However, threshold DN value corresponding to 75th percentile does not work well in the case of Swat where threshold DN corresponding to 75th percentile results in a very low DN value but a very high proportion of urban population estimate relative to population census figure. Thus, the extraction of urban areas for each district requires a district-specific threshold DN value.

In the case of Peshawar for the period 1992-2013, at the central core (99.5th percentile) light intensity and the number of pixels has grown. However, at the periphery of the core (75th percentile), light intensity has decreased as well as the number of pixels. In the case of Nowshera, the small core has maintained its area and light intensity, but all periphery has lower light intensity and area.

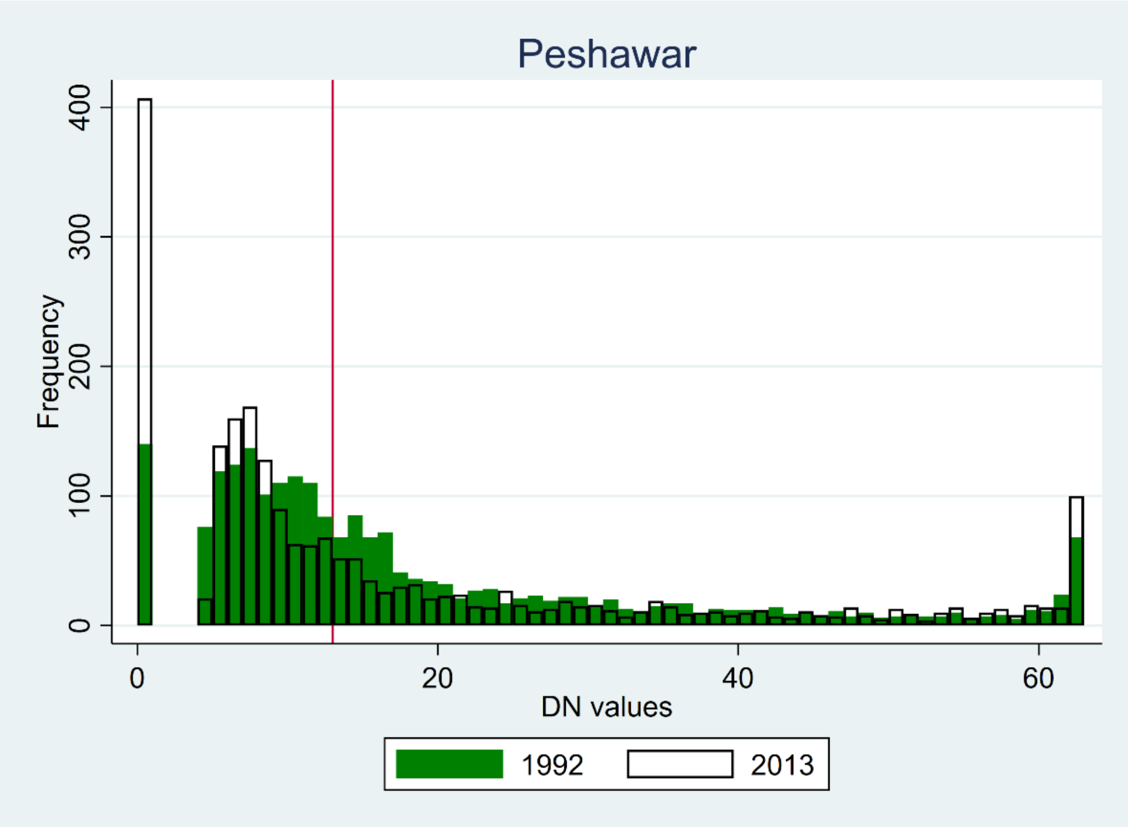
The numbers indicate that the NTL trend follows the census patterns in terms of district-wise urban growth. Figure 10 provides NTL distribution for 1998 overlaid with that of 2013. It can be observed that in the case of Peshawar there is an increase in high-intensity light as well low intensity (that is DN 62 and zero) while a reduction in the frequency of DNs close to the urban-rural boundary (indicated by the red line at DN=13)

0	**0**	**50**	**1253**	**10**	**22**	**1**
1	**34**	**1563**	**156**	**456**	**123**	**23**
34	**102**	**1563**	**5496**	**5463**	**236**	**26**
365	**4587**	**2136**	**4856**	**213**	**488**	**122**
36	**155**	**1225**	**4512**	**6536**	**213**	**34**
26	**515**	**500**	**5156**	**985**	**788**	**32**
26	**26**	**48**	**14**	**31**	**28**	**32**

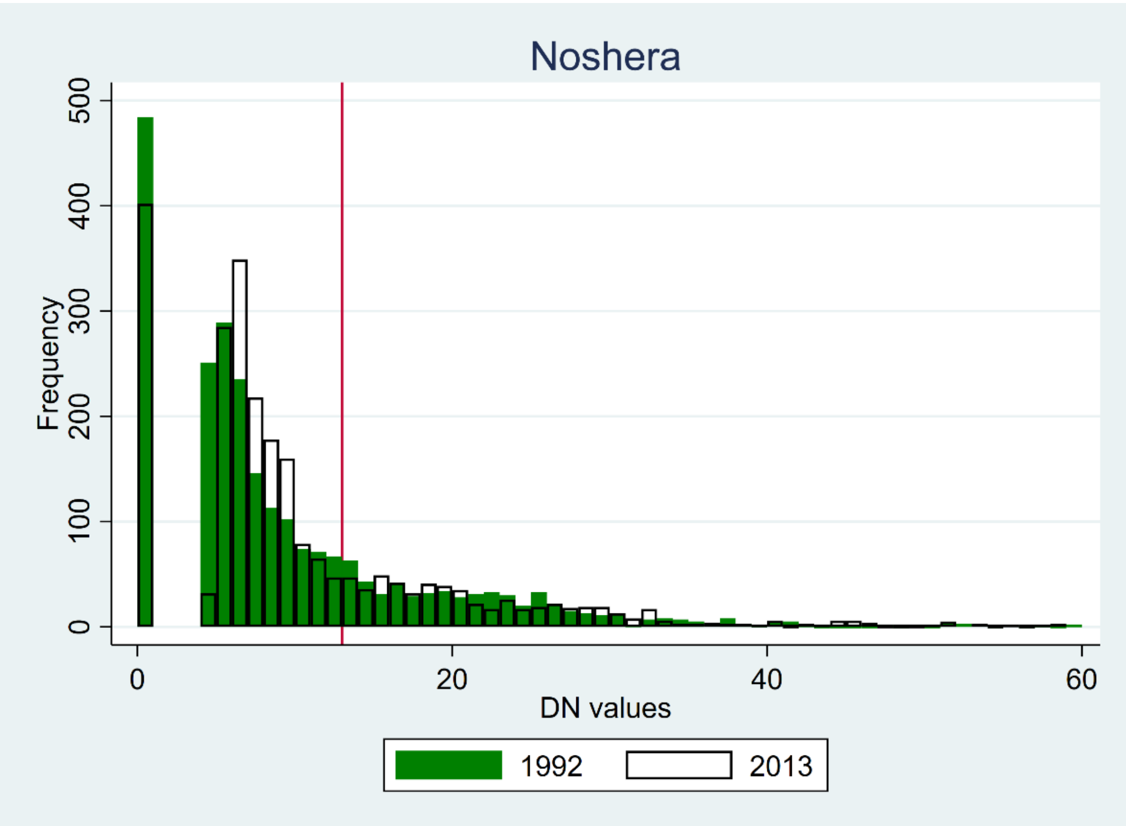
Figure 9: Calculating Urban Population Using CIESIN (2018)

Table 15: DN Value Distribution - District Level

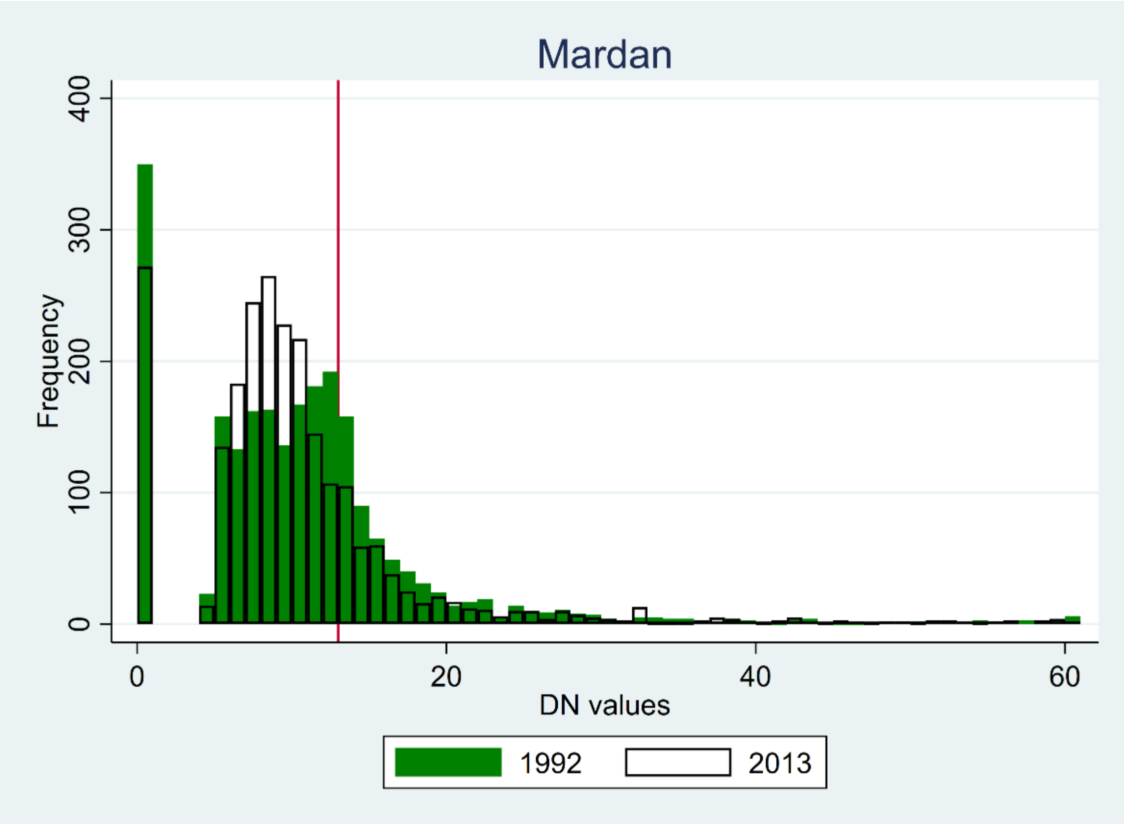
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
	75th pct.			85th pct.			90th pct.			95th pct.			99.5th pct.			Pop. Urban (%)					
	Threshold DN	Urban Pixels	Urban Pixels (1998 threshold.)	Threshold DN	Urban Pixels	Urban Pixels (1998 threshold.)	Threshold DN	Urban Pixels	Urban Pixels (1998 threshold)	Threshold DN	Urban Pixels	Urban Pixels (1998 threshold)	Threshold DN	Urban Pixels	Urban Pixels (1998 threshold)	75th pct.	85th pct.	90th pct.	95th pct.	99.5th pct.	*Population* Census
Peshawar																					
1998	28	562	.	40	345	.	49	219	.	58	125	.	62	37	.	30	19	12	7	2	48
2013	23	552	471	39	327	319	51	218	236	61	114	152	63	63	100	30	18	12	6	3	46
Swat																					
1998	0	7650	.	0	7650	.	4	838	.	5	610	.	22	38	.	100	100	12	8	0	14
2013	0	7650	7650	0	7650	7650	5	845	852	7	389	845	25	38	43	100	100	12	5	0	30
Kohat																					
1998	5	1602	.	8	783	.	9	582	.	13	281	.	49	24	.	33	17	13	6	0	27
2013	6	1209	1860	7	881	665	9	526	526	14	266	294	45	27	17	26	19	11	5	1	27
Nowshera																					
1998	16	616	.	22	374	.	26	267	.	31	128	.	52	12	.	25	15	11	5	0	26
2013	13	612	480	19	365	269	23	252	190	29	131	99	53	12	12	25	15	10	5	0	22
D.I. Khan																					
1998	4	3382	.	4	3382	.	5	1445	.	7	802	.	36	64	.	37	37	16	9	1	15
2013	0	12401	2825	5	2631	2825	6	1573	2631	8	644	942	34	66	59	100	31	19	8	1	22
Abbottabad																					
1998	8	622	.	9	455	.	12	256	.	18.6	123	.	42	14	.	25	18	10	5	1	18
2013	8	800	800	10	445	586	13	264	309	20	129	143	57	12	44	32	18	11	5	0	22
Hangu																					
1998	4	916	.	5	448	.	7	209	.	8	125	.	26	9	.	48	23	11	7	0	20
2013	5	514	557	5	514	514	6	271	156	8	105	105	19	10	0	27	27	14	5	1	20
Mardan																					
1998	16	587	.	18	364	.	22	247	.	30	123	.	58	12	.	26	16	11	5	1	17
2013	12	606	275	14	394	212	17	237	146	25	119	83	56	12	9	26	17	10	5	1	19



Peshawar
Frequency
0
100
200
300
400
0
20
40
60
DN values
1992
2013



Noshera
Frequency
0
100
200
300
400
500
0
20
40
60
DN values
1992
2013



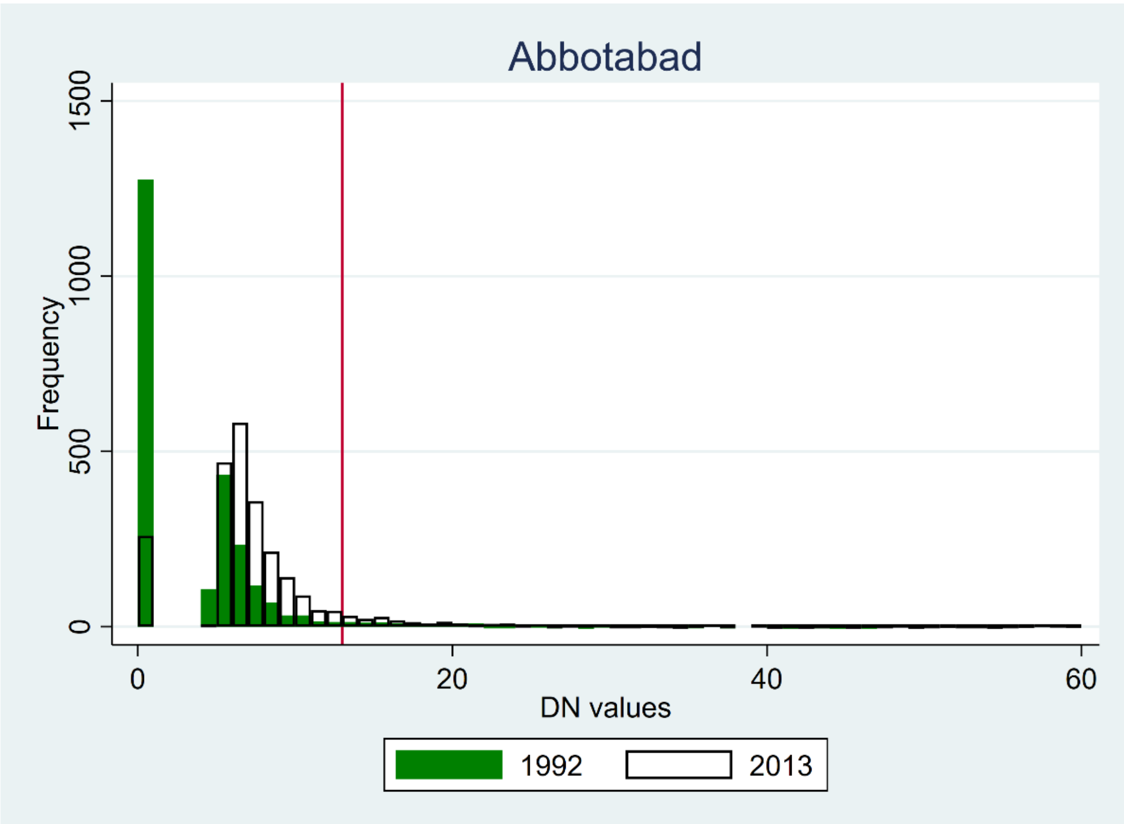


Figure 10: DN Distribution for Selected Districts

6.2.2 Brightness Gradient Method

For the BG methodology, we concentrate on the three KP districts namely Peshawar, Abbottabad, and Mardan where the quadratic model (Equation 7) fits and extract urban area growth between years 2014 and 2020. The geographical extent of these urban areas is shown in Figures 11-13. As discussed in the methodology section, following Ma et al. (2015) we spatially subdivide a city into five different sub-regions involving: low, medium-low, medium, medium-high, and high night-time lighting areas. The high NTL area corresponds to the city core whereas the low NTL is the rural area of the district. For urban extent, high, medium-high and medium-lit areas are of interest to us. However, an increase or decrease in medium-lit area signifies a future urban growth trend. For Peshawar, the percentage growth is maximum for the High, for Abbottabad the percentage growth is maximum for Medium-high and for Mardan percentage growth is maximum for Medium-low areas. These numbers signify the urban growth pattern indicating the densification in the case of Peshawar and urban sprawl in the case of Mardan.

ANNUAL PERCENTAGE CHANGE IN URBAN AREA TYPES (2014-2020)

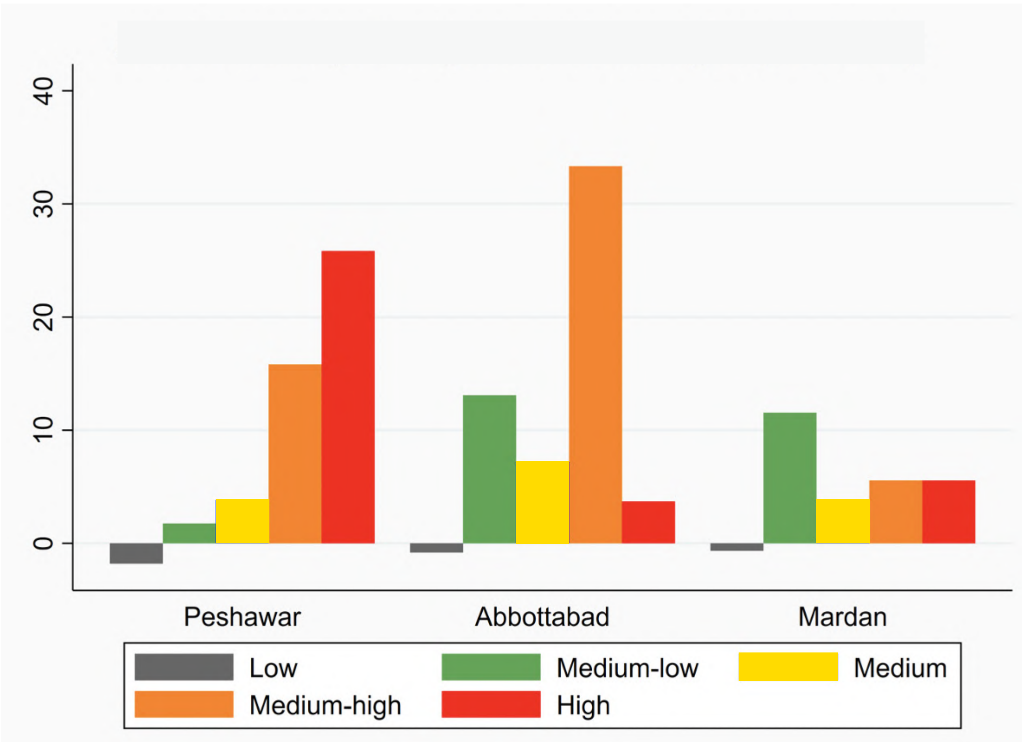


Figure 11: Annual Percentage Change for Peshawar, Abbottabad and Mardan

Table 16: Urban Area Growth by Type for Select Districts (Using BG Approach)

	2014					2020					2014-20					2014-20				
	Type of Urban Area (sq km)					Type of Urban Area (sq km)					Growth rate (%)					Annual growth rate (%)				
District	**Low**	**Medium-low**	**Medium**	**Medium-high**	**High**	**Low**	**Medium-low**	**Medium**	**Medium-high**	**High**	**Low**	**Medium-low**	**Medium**	**Medium-high**	**High**	**Low**	**Medium-low**	**Medium**	**Medium-high**	**High**
Peshawar	1641.9	353.5	162.2	40.3	41.3	1465.2	390.6	199.4	78.5	105.4	-10.8	10.5	22.9	94.9	155.0	-1.8	1.8	3.8	15.8	25.8
Abbottabad	2397.0	110.7	38.3	5.2	9.3	2281.1	197.7	54.9	15.5	11.4	-4.8	78.5	43.2	200.0	22.2	-0.8	13.1	7.2	33.3	3.7
Mardan	2254.9	121.1	13.5	3.1	3.1	2165.8	205.1	16.6	4.1	4.1	-4.0	69.3	23.1	33.3	33.3	-0.7	11.5	3.8	5.6	5.6

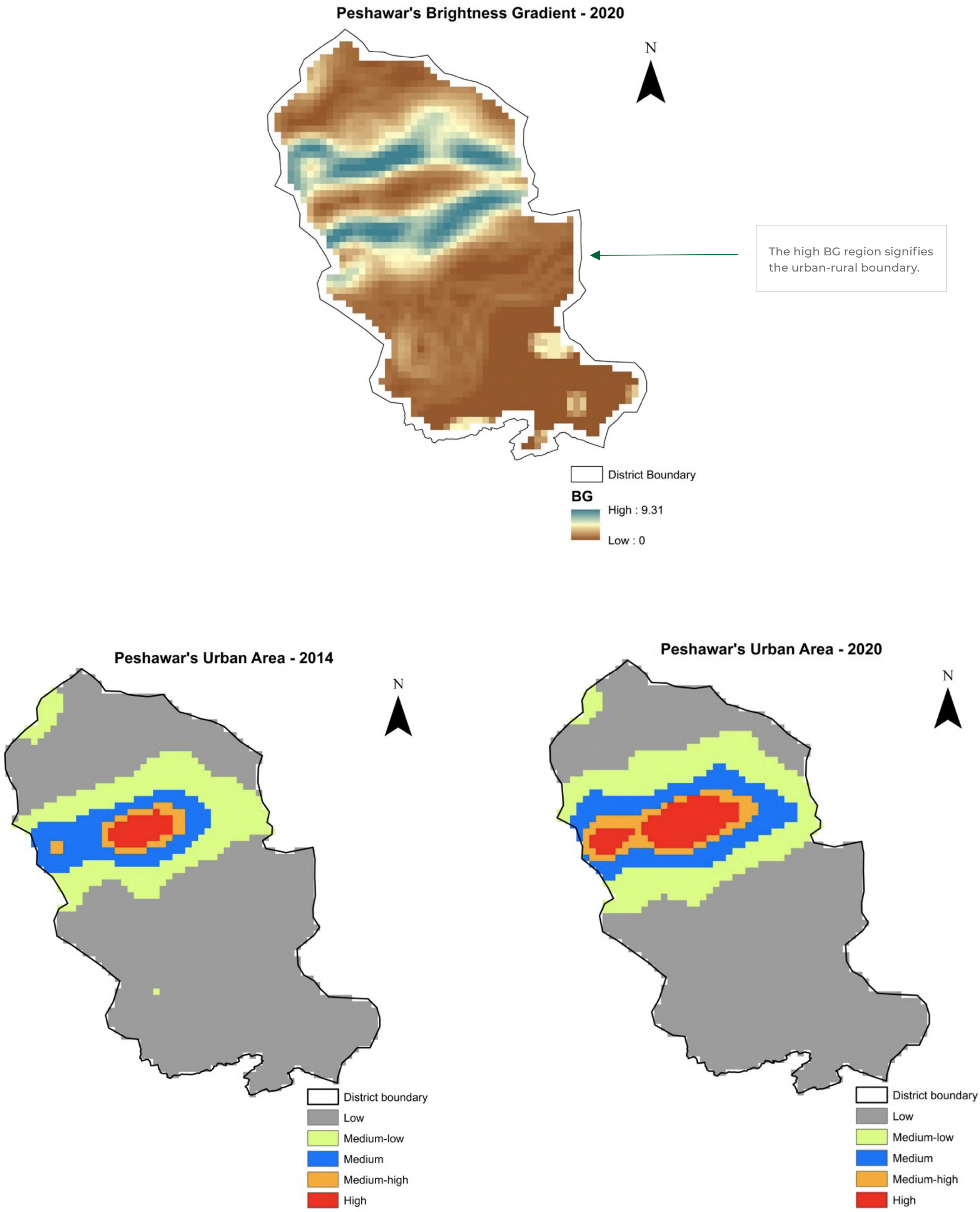


Figure 12: Peshawar BG Map

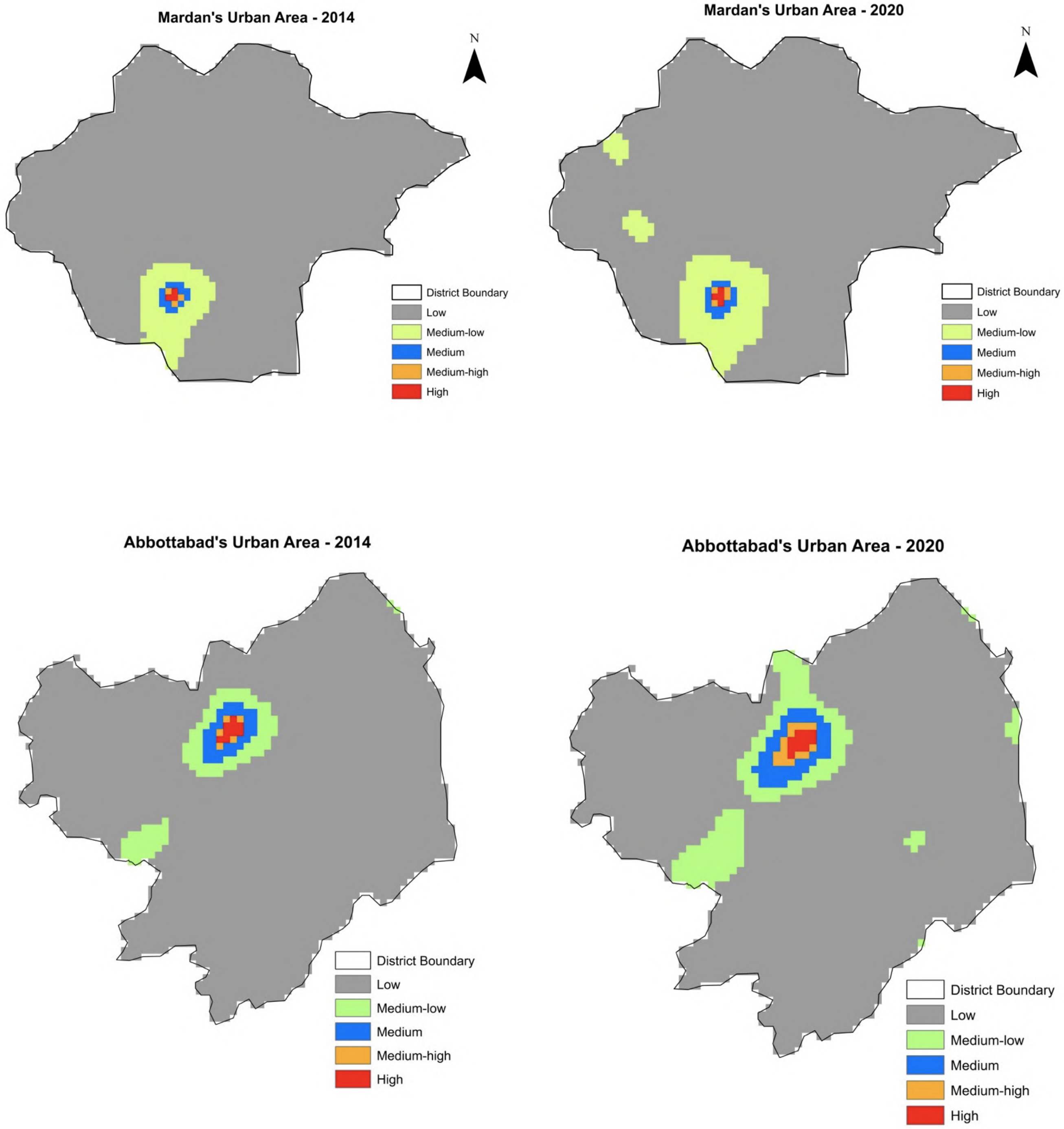


Figure 13: Mardan and Abbottabad BG Map

6.3 Comparing Urban Areas in KP Using MODIS and NTL

Using NTL with thresholding technique in case of small or medium cities or brightness gradient methodology in case of large cities, urban pixels can be extracted. Figure 14 signifies urban growth over the period 1992-2013. The choice of years is made so that NTL is based on DMSP source only. Similarly urban areas can be extracted using Land cover/land data from MODIS. Land use maps from MODIS are shown in figure 14 that correspond to years 2001 and 2019. It appears that land use data is a good resource for identification of urban areas but its use to measure urban growth has limited scope.

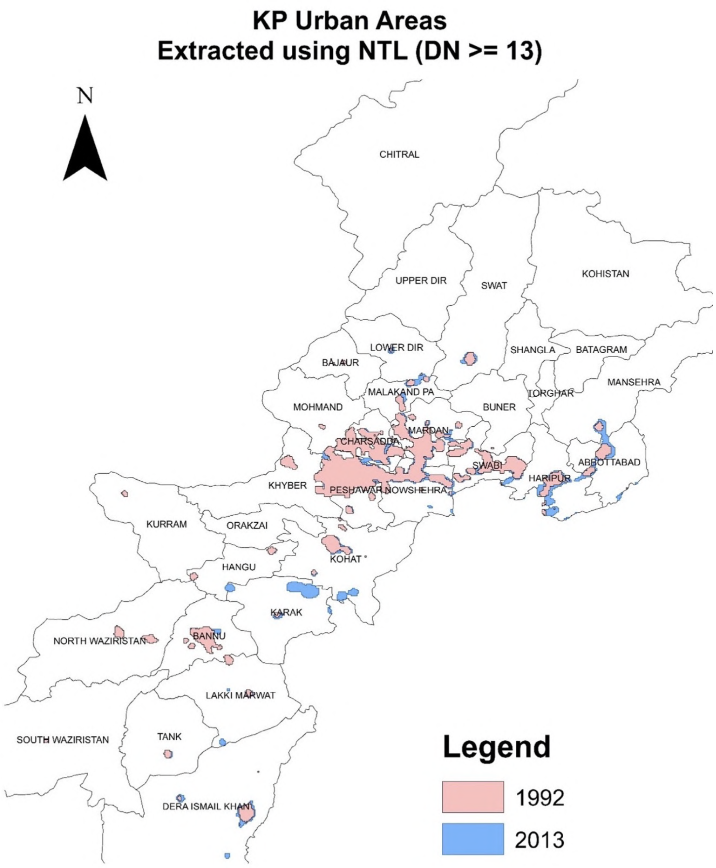
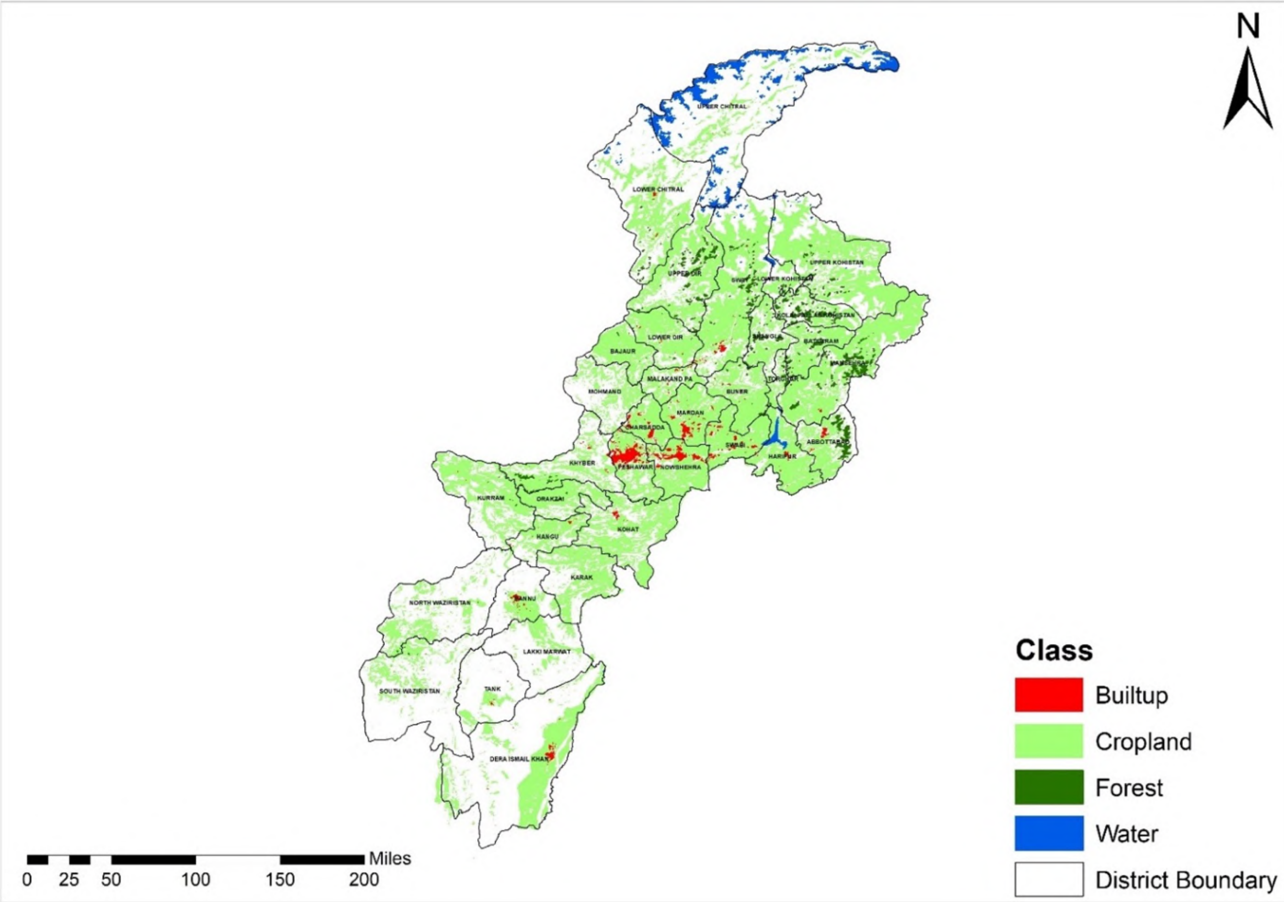
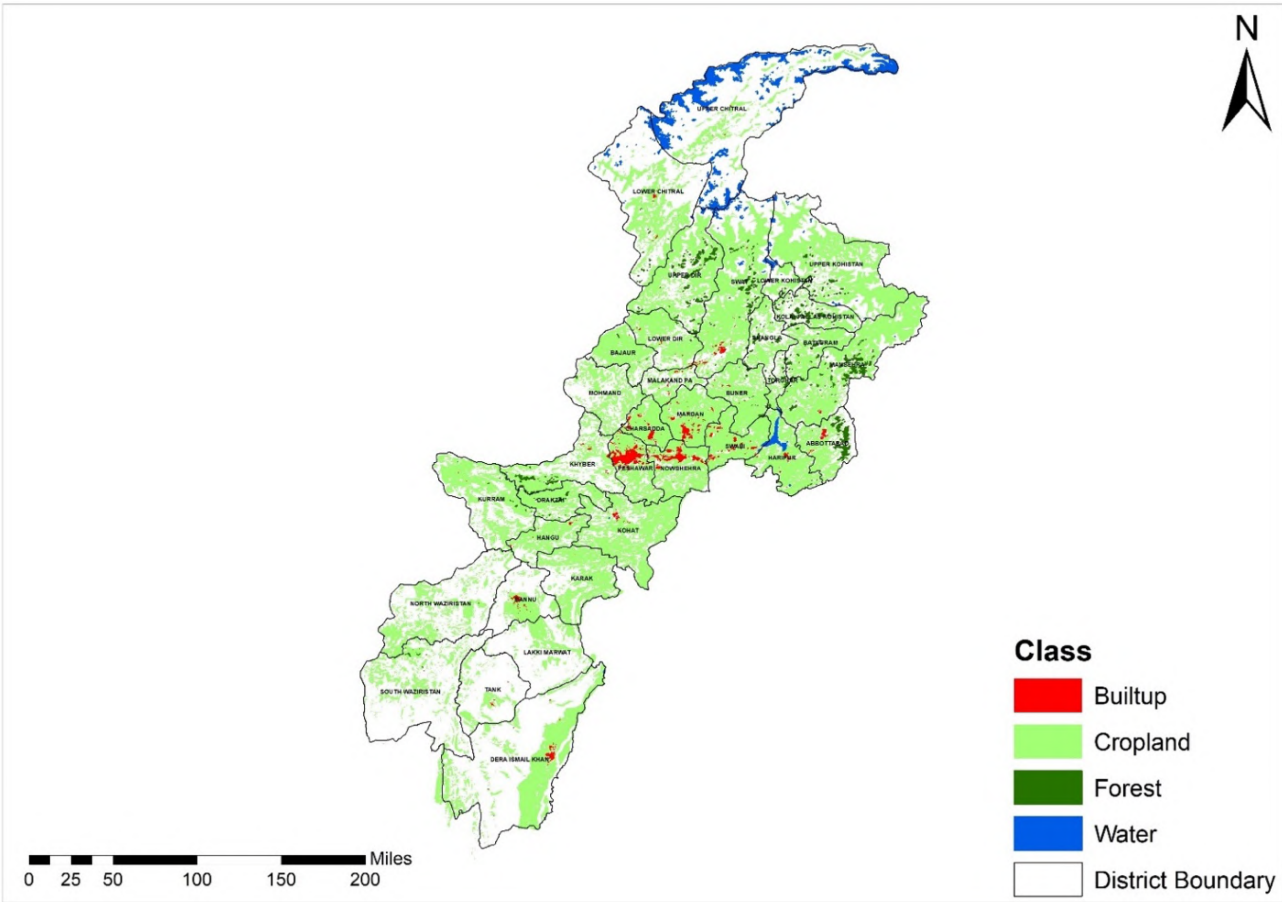


Figure 14: Urban Area Extraction Using NTL-KP Province



Panel A -2001



Panel B -2019

Figure 15: KP land use maps using Landscan (A and B)

We also analyse the use of MODIS data to study urban growth for a particular city. Figure 16 shows change in urban boundary between 2001 and 2019 for Peshawar. The MODIS data indicates some additional urban land use, but it is difficult to have a correct perception of the trend and spatial direction of city growth.

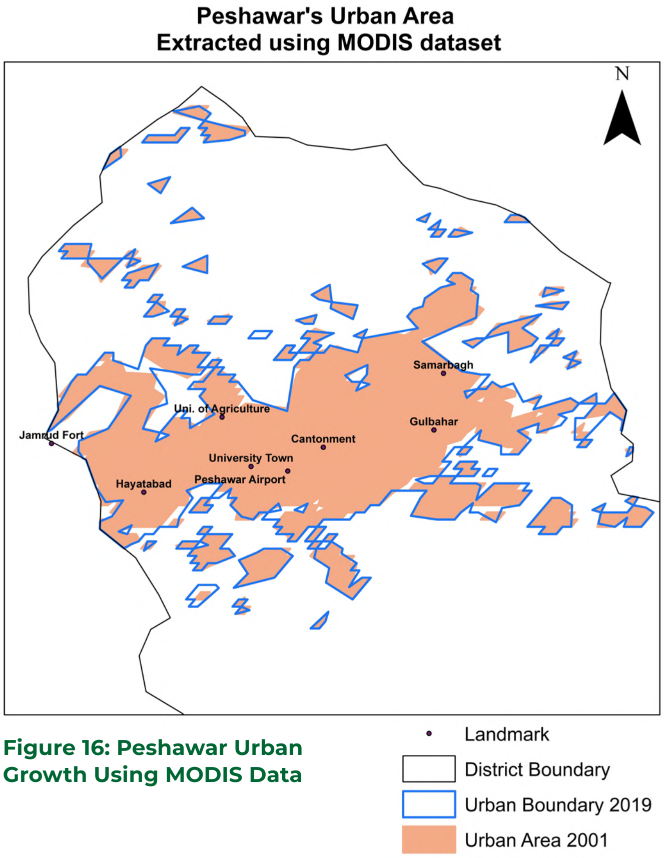


Figure 16: Peshawar Urban Growth Using MODIS Data

In comparison use of NTL with the Brightness gradient methodology yields very clear results in terms of direction and quantum of urban growth as shown in figure 17.

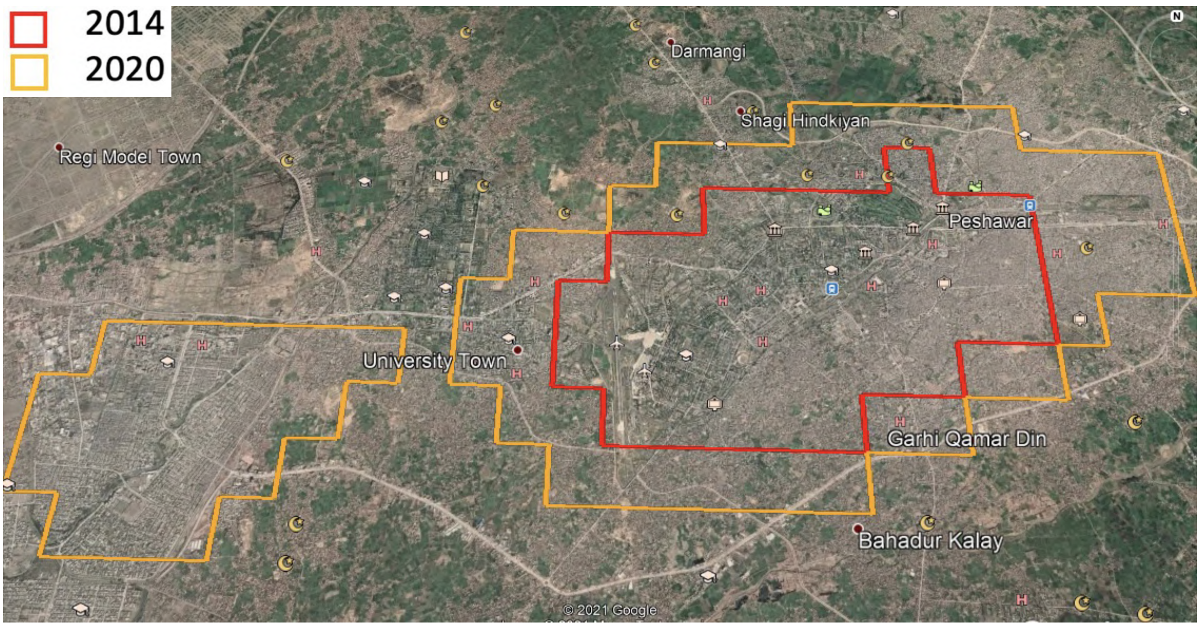


Figure 17: Peshawar Urban Boundary - Using NTL and BG

The NTL data shows a common trend where population growth in large cities of KP such as Peshawar, Mardan, Nowshera, and Abbottabad has resulted in urban sprawl. Apart from Peshawar, there is dimming of nightlights at the core of the city and more growth on the periphery which once again highlights the concerns related to horizontal urban growth and issues related to efficient provision of municipal services and public transportation. Urban growth on the periphery of cities is sometimes not truly captured by population census. Hence, this form of urban growth could lead to "messy" and "hidden" urbanization as also highlighted by the 2015 World Bank report on the South Asian cities.

Figures 16 and 17 focus on the Peshawar district and the city specifically. To examine intracity patterns of urbanization in our analysis, we divide the city into five distinct zones. The High and Medium-high refer to the city core and areas surrounding the core. The medium region is the relatively less urbanized part of the district. The Medium-low region is the low-density boundary between the urban and rural regions and is indicative of urban sprawl and future direction of development. Finally, the Low region is the rural area of the district. Figure 4 clearly shows that in the case of Peshawar, the high-density city core and low-density city periphery have significantly increased over time. In figure 5, we provide more details by placing the core area boundary for the years 2014 and 2020 on Google map to add a specific place perspective and indicate the pattern of growth more explicitly.

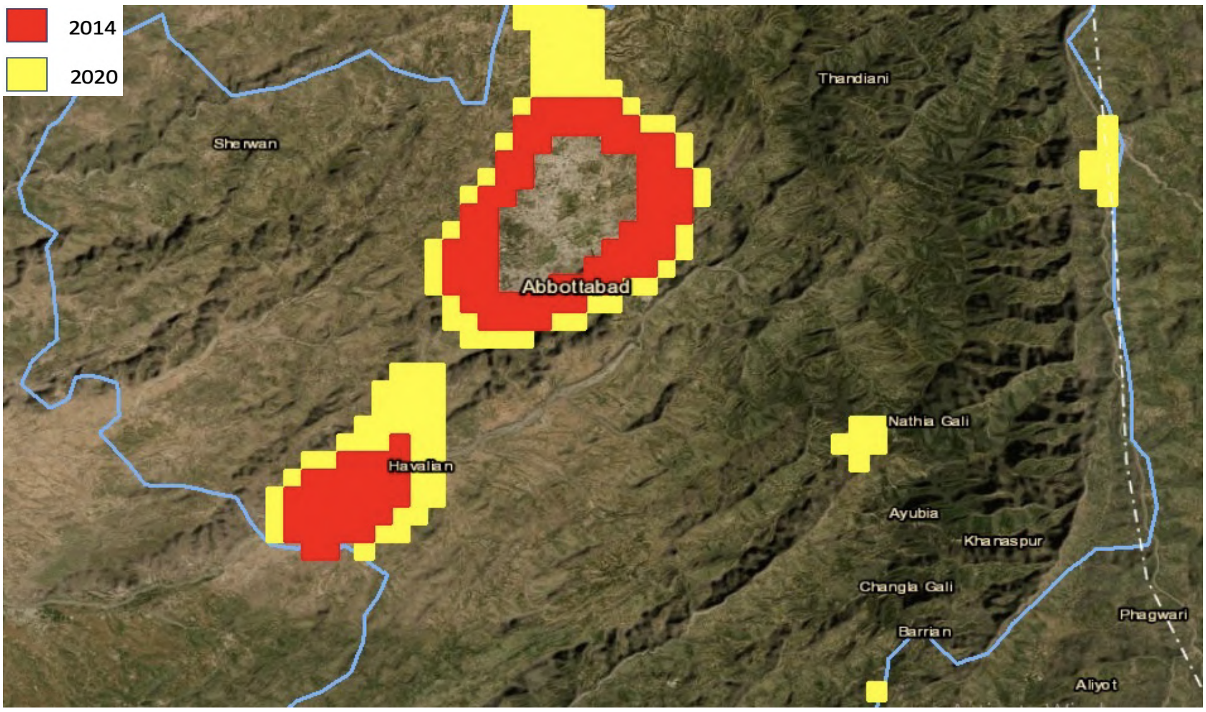


Figure 18: Periphery Growth in Abbottabad (2014-2020)

In the case of Abbottabad, the analysis shows that there is little growth in the city core. However, most of the recent growth has taken place on the city periphery which is indicative of urban sprawl. Figure 18 highlights the change in the low-density development on the city periphery and shows that both tehsils of Abbottabad district – Abbottabad and Havelian as well as locations in the Galiyat region are experiencing new low density urban development which is often considered an inefficient use of land.

6.4 GDP of FATA/ Newly Merged Districts of KP Province

The GDP of the KP province estimated by the provincial BOS (2005-06 to 2019-2020) does not include the economic output of the FATA region / newly merged districts. The survey-based data collection for the FATA region has also remained sketchy due to the special status of the area and security issues involved in data collection and hence factual assessment of economic activity beyond agriculture is difficult. The PBS documentation discussing National Accounts does not discuss this specifically as well.

In this study, we aim to quantify the GDP of the FATA region using the Beyer et al. (2018) methodology. As can be seen from table 11, all the districts/ agencies belonging to the FATA region except Mohmand and North Waziristan have nearly null share in the manufacturing/services sector as assessed by the NTL. As such we can safely assume that most economic activity in the region relates to the farming sector.

Although the GDP contribution of FATA / newly merged districts is not reflected in the GDP figures of the KP province, this must be included in the national figures. We therefore aim, to spatially segregate this component from the national GDP figures by applying Beyer et al. (2018)

approach on the most recent 2019-2020 national GDP figures. The rural population of the region is estimated using the Landscan data as discussed earlier, and this figure is 5,097,461 which is close to the 2017 census number. Using the NTL share and the rural population share in (4), our calculations show that the region has a GDP that is around 1.03% of the national GDP.

6.5 Limitations

The results presented in this study are subject to certain limitations. Data based on satellite imagery such as NTL is not appropriate to estimate the agriculture sector's contribution to the economy or for regions with low output density. Similarly, daytime landcover data such as MODIS provides one-dimensional information and cannot account for yield changes due to other technological improvements such as mechanization etc. As such we relied on the survey-based estimates of agriculture output as it transpired that the systems and procedures for agricultural data collection are in place for a long while and the department conducting crop surveys has extensive experience in this exercise. However, we want to highlight that our attempt to validate satellite imagery data (MODIS or Landsat data) with crop survey data indicated some discrepancies between the two and hence warrants more scrutiny which was beyond the scope of the current study.

Also, for the provincial GDP estimations, we had to use data for Indian states as the number of observations and data for Pakistani provinces since 1992 were not enough for good model estimation. As such provincial estimations are done for the period 2005-2014 as shown in the text.

Another limitation we faced related to the application of the brightness gradient approach in KP districts. This was mainly due to the limited level of urbanization in KP districts. As such the quadratic model fitted few districts and we had to employ the thresholding method for other districts.

Finally, the provincial GDP estimates by the BOS KP still do not include the economic output of the newly merged districts from the erstwhile FATA belt. As such our calculation of district-level GDP is limited to the districts in the pre-merger KP province though we do provide an estimated GDP for the FATA region.

USING THE NTL DATA

07

7.1 Possible Uses Learnt from the Study

This study provides several insights based on analysis using NTL data that can be utilized for better planning.

Firstly, it gives a clear understanding of the economic growth and development of districts in the KP province. The availability of this information is a prerequisite for planning infrastructure (for example to connect vibrant economic clusters and industrial parks) and is vital for designing policies to address regional disparities.

Secondly, our intra-city analysis gives a clear picture of the growth of cities, in particular the pace and direction of new development. This information can feed into urban planning, for example addressing issues of urban sprawl, planning for new infrastructure, and provision of urban services and local amenities (roads, transportation, utilities, crime control, waste disposal, mitigation of environmental degradation) to accommodate growing cities, realize their growth potential and help make them sustainable and efficient. It can also be used to develop a system of cities with plans for secondary cities, that so far appear to be less visible in the case of KP. The report also documents a low and stagnating level of urbanization in KP overall, which means that the agglomeration and growth potential of cities is currently highly underutilized and is an area where policymakers in KP can devote more attention.

In addition to these immediate insights, once the capacity to understand and leverage nightlights data is developed in KP, it opens the doors to a reliable, live, and inexpensive source of information on future topics where official data is usually silent.

7.2 Future Extensions

The spatial information embedded in the nightlights data gives it the unique advantage of providing granularity to the level of 1 sq. km. Estimation of sub-national GDP to the level of a district or a tehsil/town yields an economic indicator that can be extensively used for policy planning and evaluation. Nightlights have been used to (a) obtain reliable estimates of poverty; (b) study the impact of cash transfers; (c) create wealth index; (d) create development indices; (e) validate district per capita income values obtained from Labor Force Surveys; (f) calculate regional inequality measures (g) understand urban crime patterns, etc. Most of these studies have been carried out by combining gridded nightlights data, gridded population data, and socioeconomic variables from other available datasets.

Recently, nightlight data has been used to study the economic impact of various events, for example, restrictive government measures such as the imposition of lockdowns following the COVID-19 outbreak. The argument is that non-pharmaceutical measures to contain disease transmission while mitigating their economic impact require an assessment of the economic situation in near real-time and at high spatial granularity. A comparison of changes in light intensity before and after lockdowns indicates the impact on local economic activity and hence provides useful information for policymakers. This approach could be useful not just to measure COVID impacts but also to assess the impact of natural disasters or other policy measures, in a sense allowing the policymaker a way to continually monitor the pulse of economic activity in the province.

The findings of the study can be used for policy design and analysis in the following areas:

a) Understanding inter-district disparities in the scale of economic activity

b) Adopting place-based interventions to uplift lagging districts

c) Linking district level economic activity from industry and service sectors with tax collection figures to determine potential revenue generation

d) Use urban growth trends -both inter-district and intra-district for planning and management of existing and future cities in the province.

The utilization of satellite imagery, in particular nightlights, is a relatively new concept and is still evolving as researchers find novel ways to use it to understand development. It is an opportune time for KP to develop the expertise to analyze and utilize this rich source of information to help achieve its development targets.

08 REFERENCES

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. Journal of public economics, 95(7-8), 476-487.

Arby, M. Farooq (2008), “Some Issues in the National Income Accounts of Pakistan (Rebasing, Quarterly and Provincial Accounts and Growth Accounting),” PhD Dissertation, Pakistan Institute of Development Economics, Islamabad

Alesina, A., Michalopoulos, S., & Papaioannou, E. (2016). Ethnic inequality. Journal of Political Economy, 124(2), 428-488.

Ali, A. (2016). The Impact of Electricity Outages on Households (Doctoral dissertation)

Ali, R., Barra, A. F., Berg, C. N., Damania, R., Nash, J., & Russ, J. (2015). Transport infrastructure and welfare: an application to Nigeria. The World Bank.

Alder, S., Roberts, M., & Tewari, M. (2017). The effect of transport infrastructure on India’s urban and rural development. Working Paper, University of North Carolina, Chapel Hill.

Amaral, S., Câmara, G., Monteiro, A. M. V., Quintanilha, J. A., & Elvidge, C. D. (2005). Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data. Computers, Environment and Urban Systems, 29(2), 179-195.

Azar, D., Engstrom, R., Graesser, J., & Comenetz, J. (2013). Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data. Remote Sensing of Environment, 130, 219-232

.

Baum-Snow, N., Henderson, J. V., Turner, M., Brandt, L., & Zhang, Q. (2015). Transport infrastructure, urban growth, and market access in China.

Baragwanath, K., Goldblatt, R., Hanson, G., & Khandelwal, A. K. (2019). Detecting urban markets with satellite imagery: An application to India. Journal of Urban Economics, 103173.

Bengali, Kaiser, and Mahpara Sadaqat. "Provincial Accounts of Pakistan: Methodology and Estimates 1973-2000." SPDC WP 5 (2005): 2005.

Berg, C. N., Blankespoor, B., & Selod, H. (2018). Roads and rural development in Sub-Saharan Africa. The Journal of Development Studies, 54(5), 856-874.

Beyer, R. C., Chhabra, E., Galdo, V., & Rama, M. (2018). Measuring districts' monthly economic activity from outer space. The World Bank.

Belleflamme, P., Picard, P., & Thisse, J. F. (2000). An economic theory of regional clusters. Journal of urban economics, 48(1), 158-184.

Bickenbach, F., Bode, E., Nunnenkamp, P., & Söder, M. (2016). Night lights and regional GDP. Review of World Economics, 152(2), 425-447.

Bird, J. H., Lebrand, M. S. M., & Venables, A. J. (2020). The Belt and Road Initiative: Reshaping economic geography in Central Asia? Journal of Development Economics, 144, 102441.

Burki, S. J., and A. G. Pasha. "State of the Economy." The Punjab Story. Lahore: Institute of Public Policy, Beaconhouse National University (2012).

Bhandari, L., & Roychowdhury, K. (2011). Night lights and economic activity in India: A study using DMSP-OLS nighttime images. Proceedings of the Asia-Pacific advanced network, 32(0), 218.

Callen, T. (2012). Gross domestic product: An economy’s all. International Monetary Fund, 28.

Center for International Earth Science Information Network-CIESIN-Columbia University. (2018). Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 11 Data Sets.

Doll, C. N., Muller, J. P., & Morley, J. G. (2006). Mapping regional economic activity from night-time light satellite imagery. Ecological Economics, 57(1), 75-92.

Ellis, P., & Roberts, M. (2015). Leveraging urbanization in South Asia: Managing spatial transformation for prosperity and livability. The World Bank.

Elvidge, C. D., Zhizhin, M., Baugh, K., Hsu, F. C., & Ghosh, T. (2016). Methods for global survey of natural gas flaring from visible infrared imaging radiometer suite data. Energies, 9(1), 14.

Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., & Ghosh, T. (2012). The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data. Social Geography, 7(1), 23-35.

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization, and localized competition. European economic review, 43(2), 409-429.

Friedl, M., Sulla-Menashe, D. (2019). MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2021-05-15 from https://doi.org/10.5067/MODIS/MCD12Q1.006

Gibson, J., Olivia, S., & Boe-Gibson, G. (2020). NIGHT LIGHTS IN ECONOMICS: SOURCES AND USES 1. Journal of Economic Surveys, 34(5), 955-980.

Gibson, J., Boe-Gibson, G., & Stichbury, G. (2015). Urban land expansion in India 1992–2012. Food Policy, 56, 100-113.

Gramlich, E. M. (1994). Infrastructure investment: A review essay. Journal of economic literature, 32(3), 1176-1196.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. Journal of political economy, 100(6), 1126-1152.

Glaeser, E. L., Scheinkman, J., & Shleifer, A. (1995). Economic growth in a cross-section of cities. Journal of monetary economics, 36(1), 117-143.

Ghosh, T., L Powell, R., D Elvidge, C., E Baugh, K., C Sutton, P., & Anderson, S. (2010). Shedding light on the global distribution of economic activity. The Open Geography Journal, 3(1).

Harari, M. (2020). Cities in bad shape: Urban geometry in India. American Economic Review, 110(8), 2377-2421.

Hashim, S. R. (2014). Urban Development and the Poor. Social Change, 44(4), 505-518.

Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. Journal of political economy, 103(5), 1067-1090.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. "Measuring economic growth from outer space." The American Economic Review 102, no. 2 (2012): 994-1028.

Holmgren, J., & Merkel, A. (2017). Much ado about nothing?–A meta-analysis of the relationship between infrastructure and economic growth. Research in Transportation Economics, 63, 13-26.

Imhoff, M. L., Lawrence, W. T., Stutzer, D. C., & Elvidge, C. D. (1997). A technique for using composite DMSP/OLS "city lights" satellite data to map urban areas. Remote Sensing of Environment, 61(3), 361-370.

Lay, M. G. (1992). Ways of the World: A History of the World's Roads and of the Vehicles that Used Them. Rutgers university press.

Li, X., Zhou, Y., Zhao, M., & Zhao, X. (2020). A harmonized global nighttime light dataset 1992–2018. Scientific data, 7(1), 1-9.

Luoto, J. (2011). Aggregate infrastructure capital stock and long-run growth: Evidence from Finnish data. Journal of Development Economics, 94(2), 181-191.

Ma, T., Zhou, Y., Zhou, C., Haynie, S., Pei, T., & Xu, T. (2015). Night-time light derived estimation of Spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. Remote Sensing of Environment, 158, 453-464.

Mahmood, K.H., Majid, H., & Chaudhry, M.A. (2017). Quantifying economic and urban growth of Pakistan: sub-national analysis using nighttime lights data. Punjab Economic Research Institute Discussion Paper.

Mohanty, R. K., & Bhanumurthy, N. R. (2019). Analyzing the Dynamic Relationships between Physical Infrastructure, Financial Development, and Economic Growth in India. Asian Economic Journal, 33(4), 381-403.

Pasha, H. A. (2015). Growth of the provincial economies. Institute for Policy Reforms (IPR). http://ipr. org. pk/wp-content/uploads/2016/04/GROWTH-OF-PROVINCIAL-ECONOMICS-. pdf

Prakash, A., Shukla, A. K., Bhowmick, C., & Beyer, R. C. M. (2019). Night-time luminosity: Does it brighten understanding of economic activity in India. Reserve Bank of India Occasional Papers, 40(1).

Reed, T., & Trubetskoy, A. (2019). Assessing the value of market access from Belt and Road projects. World Bank Policy Research Working Paper, (8815).

Roberts, M. (2018). Urban growth in South Asia: A view from outer space. In Alternative Approaches in Macroeconomics (pp. 269-302). Palgrave Macmillan, Cham.

Small, C., & Elvidge, C. D. (2013). Night on Earth: Mapping decadal changes of anthropogenic night light in Asia. International Journal of Applied Earth Observation and Geoinformation, 22, 40-52.

Smith, P. J. (1999). Do knowledge spillovers contribute to US state output and growth?. Journal of Urban Economics, 45(2), 331-353.

Sutton, P. C., & Costanza, R. (2002). Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. Ecological Economics, 41(3), 509-527.

Tan, M. (2015). Urban growth and rural transition in China based on DMSP/OLS nighttime light data. Sustainability, 7(7), 8768-8781.

Jansen, M. (1985). Mohenjo-Daro, city of the Indus valley. Endeavor, 9(4), 161-169.

Joyeux, R., & Ripple, R. D. (2007). Household energy consumption versus income and relative standard of living: A panel approach. Energy Policy, 35(1), 50-60.

Keola, S., Andersson, M., & Hall, O. (2015). Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. World Development, 66, 322-334.

World Bank. (1994). World Development Report 1994: Infrastructure for Development. The World Bank.

Yue, W., Gao, J., & Yang, X. (2014). Estimation of gross domestic product using multi-sensor remote sensing data: A case study in Zhejiang province, East China. Remote Sensing, 6(8), 7260-7275.

Yu, P. L. (2013). Multiple-criteria decision making: concepts, techniques, and extensions (Vol. 30). Springer Science & Business Media.

Zhou, N., Hubacek, K., & Roberts, M. (2015). Analysis of spatial patterns of urban growth across South Asia using DMSP-OLS nighttime lights data. Applied Geography, 63, 292-303.

APPENDIX A: TABLES

08

Table 1A: Sectoral share (%) in province's GDP

	1971-75	1976-80	1981-85	1986-90	1991-95	1996-00	2000-05
KP							
Agriculture	38.12	32.76	29.32	26.84	24.31	22.25	21.32
Industry	17.85	18.62	20.98	21.74	22.28	22.24	22.46
Services	44.03	48.63	49.7	51.42	53.42	55.51	56.22
Punjab							
Agriculture	43.14	38.65	34.87	32.26	29.75	27.98	25.51
Industry	19.87	21.3	23.05	24.01	24.37	23.59	23.88
Services	36.98	40.04	42.08	43.73	45.88	48.43	50.61
Sindh							
Agriculture	27.88	25.41	22.28	19.40	18.63	20.08	17.83
Industry	27.21	26.55	29.74	31.14	29.86	28.20	28.40
Services	44.91	48.04	47.98	49.46	51.51	51.72	53.77
Balochistan							
Agriculture	32.47	30.06	29.57	27.77	27.24	29.81	25.77
Industry	23.54	25.52	26.31	24.86	26.18	22.35	25.17
Services	43.99	44.42	44.13	47.37	46.57	47.83	49.06

Source: Arby, M. F. (2008). 5-year averages calculated by authors themselves.

Table 2A: Sectoral Share of Industries (%) of KP

	1971-75	1976-80	1981-85	1986-90	1991-95	1996-00	2000-05
Agriculture							
Major crops	27.12	30.8	29.02	26.49	26.93	24.73	17.81
Minor crops	10.2	15.17	13.6	13.48	12.79	11.82	11.29
Livestock	53.52	46.5	46.8	49.16	51.84	57.36	57.3
Fishing	0.03	0.24	0.08	0.08	0.20	0.07	0.11
Forestry	9.13	7.29	10.5	10.79	8.24	6.03	13.49
Industry							
Mining	0.28	0.67	0.66	1.59	1.96	2.26	2.83
LSM	52.69	47.39	54.29	49.62	40.52	35.85	39.36
SSM	4.63	5.29	4.99	6.5	7.51	12.23	15.1
Slaughtering	19.96	20.96	18.88	19.62	24.88	21.73	21.13
Construction	15.04	18.69	14.64	13.71	12.66	11.02	9.06
Electricity	7.41	7.0	6.54	8.98	12.47	16.9	12.52
Services							
Trade	43.09	35.3	36.15	33.84	31.21	29.86	32.61
Transport	13.44	16.63	21.8	22.4	22.7	25.53	24.76
Finance	6.99	6.89	3.5	4.33	3.62	2.75	2.58
Dwelling	3.85	3.99	3.82	3.82	4.07	4.5	5.01
Public Adm.	13.06	17.66	17.05	16.15	14.65	12.93	13.89
Social	19.58	19.52	17.68	19.45	23.74	24.43	21.14

Source: Arby, M. F. (2008). 5-year averages calculated by authors themselves.

Table 3A: Percentage distribution of Employment (10 years and above) - KP province

	2009-10	2010-11	2012-13	2014-15	2017-18
Agriculture					
Whole	41.02	37.95	36.79	34.56	32.62
Urban	8.64	5.90	4.90	4.06	7.15
Rural	47.32	44.26	43.30	41.15	38.29
Industry					
Whole	21.46	24.17	22.50	24.62	26.88
Urban	26.27	28.15	24.75	29.00	27.59
Rural	20.52	23.39	22.06	23.68	26.70
Services					
Whole	37.51	37.88	40.71	40.81	40.51
Urban	65.09	65.97	70.35	66.96	65.25
Rural	32.16	32.34	34.64	35.17	35.00

APPENDIX B: KP DISTRICT LEVEL URBAN AREA (HECTARES) USING MODIS

10

District	2001	2005	2010	2015	2019
PESHAWAR	25310	25056	25477	25839	26706
NOWSHEHRA	15815	15979	16102	16213	16561
MARDAN	9741	9664	9570	9659	10033
SWABI	8046	8092	8045	8111	8143
CHARSADDA	6463	6397	6447	6482	6540
SWAT	4862	5042	5331	5617	5836
DERA ISMAIL KHAN	5203	5257	5288	5348	5359
ABBOTTABAD	2978	2959	3018	3186	3384
BANNU	3424	3364	3344	3353	3347
HARIPUR	2144	2229	2251	2263	2355

APPENDIX C: PROVINCIAL GDP ESTIMATES BY KP'S BOS - AVAILABLE AT BOS KP WEBSITES

11

S.No	Description	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19(F)	2019-20(R)
A	GVA of Agriculture Sector	190,622	192,962	197,503	205,670	215,281	214,386	220,539	222,868	230,370	236,035	239,884
1	Crops	36,531	35,305	36,317	39,213	43,214	46,291	46,862	44,424	45,316	44,024	41,084
i	Major Crops	22,258	21,728	23,238	25,041	28,575	30,170	30,428	28,551	29,244	28,075	25,685
ii	Minor Crops	14,273	13,578	13,079	14,172	14,639	16,121	16,434	15,874	16,072	15,949	15,400
2	Livestock & Poultry	143,075	147,299	151,571	155,898	160,403	165,877	170,666	176,044	181,513	187,814	194,240
3	Fisheries	232	1,953	2,579	754	1,860	1,488	770	861	2,081	2,394	2,952
4	Forestry	10,785	8,405	7,036	9,805	9,805	730	2,242	1,538	1,460	1,804	1,608
B	GVA of Industries Sector	186,622	197,660	214,364	219,448	231,067	243,584	265,981	271,893	297,846	297,981	295,273
1	M&Q	37,192	49,521	54,752	63,620	64,555	68,638	73,608	78,768	82,508	84,129	79,086
2	Manufecturing	87,460	91,535	103,642	99,874	103,379	103,812	114,606	110,360	130,579	132,110	127,106
i	LSM	67,722	70,462	81,165	75,893	77,798	76,533	85,479	79,263	97,367	96,603	90,818
ii	SHMI	12,610	13,682	14,825	16,053	17,383	18,810	20,352	22,010	23,807	25,769	26,155
iii	Slaughtering	7,128	7,390	7,651	7,929	8,198	8,470	8,776	9,087	9,405	9,738	10,132
3	Construction	43,428	39,713	40,936	41,377	43,845	47,029	53,461	58,248	64,557	54,548	57,524
4	EGD&GD	18,542	16,892	15,034	14,576	19,288	24,105	24,306	24,517	20,202	27,194	31,557
i	EGD	18,542	16,892	13,752	15,293	18,599	22,651	22,559	22,592	18,082	24,857	28,981
ii	GD	-	-	1,282	- 717	689	1,454	1,746	1,924	2,121	2,337	2,576
C	GVA of Services Sector	463,564	462,149	481,543	518,721	545,197	570,126	610,984	651,440	711,820	711,422	676,043
1	Whole Sale Trade & Retail and Hotel & Resturants	174,414	176,407	196,132	209,503	228,670	239,386	263,415	282,241	322,374	312,853	274,953
2	Transport Storage and Communication	123,912	108,642	96,534	105,703	106,883	109,818	112,443	116,743	120,271	114,895	110,288
3	Finance and Insurance	15,941	16,917	17,998	20,513	20,314	22,671	24,752	27,494	29,995	33,089	34,855
4	Housing	32,019	33,000	34,013	35,058	36,136	37,248	38,396	40,268	40,588	40,836	41,161
5	General Government	44,913	51,227	56,892	63,335	65,144	68,287	74,925	83,791	93,682	98,529	99,538
6	Other Private Services	72,366	75,955	79,975	84,609	88,050	92,716	97,053	100,902	104,911	111,220	115,248
	GDP of Khyber Pakhtunkhwa	840,809	852,770	893,410	943,839	991,545	1,028,097	1,097,504	1,146,201	1,240,036	1,245,438	1,211,200
	GR of GDP	3.16%	1.42%	4.77%	5.64%	5.05%	3.69%	6.75%	4.44%	8.19%	0.44%	-2.75%

SEED

Adam Smith International

UKaid from the British people