Measuring post-war economic activity at the regional level in Ukraine using nighttime lights and machine learning. *

Department of Methodology,
London School of Economics and Political Science
Supervisor: Dr Tom Robinson
Candidate Number: 23498
Word Count: 9779

August 13, 2024

Abstract

This paper investigates how the Ukrainian economic activity at the regional level has changed since the Russian invasion. I use data on nighttime light luminosity and machine learning to arrive at the GDP estimates for each Ukrainian region in 2022. The results suggest that the Ukrainian economy has shrunk by 31% between 2021 and 2022, with eastern Ukraine experiencing a more pronounced decline in economic activity. Estimated change in GDP at the regional level ranges from -68% to -1.1%. I offer several possible explanations for differences in the predicted decline in economic activity. To validate my approach, I investigate whether it accurately predicts regional GDP for Poland, which has experienced a large refugee influx due to the war. Additionally, I explore whether using a CNN to extract features for time-series prediction yields better results. To justify my choice of NTL data empirically, I investigate how the model accuracy and predictions change depending on the composite used. Finally, I demonstrate how my predictions can be used to answer an applied research question about how different types of warfare affect economic activity.

Acknowledgements

I would like to express my gratitude to Dr Tom Robinson, Dr Franscesca Panero, Dr Friedrich Geiecke, and my sister for their insightful feedback and support.

*The code for this paper can be found *here*.

Contents

List of Abbreviations

ACLED Armed Conflict Location & Event Data Project

ARIMA Autoregressive Integrated Moving Average

BRDF Bidirectional Reflectance Distribution Function

CNN Convolutional Neural Network

DMSP Defense Meteorological Satellite Program

GDP Gross Domestic Product

MAE Mean Absolute Error

MSE Mean Squared Error

MODIS Moderate Resolution Imaging Spectroradiometer

NTL Nighttime Lights

PLN Zloty, Polish currency

UAH Hryvnia, Ukrainian currency

VIIRS Visible Infrared Imaging Radiometer Suite

1 Introduction

On 24 February 2023, Russia launched a full-scale invasion of Ukraine, which has been ongoing until the time of writing this paper. The war caused immense suffering to the citizens of Ukraine and led to one of the biggest refugee crises since World War II (UNHCR 2024). Presently, fighting is taking place primarily in the east and south of the country, with the frontline going through Kharkiv, Luhansk, Donetsk, Zaporizhzhia, and Kherson regions. The exact impact of the war on different parts of Ukraine remains unknown since gathering reliable information about the condition of society and infrastructure is challenging during an armed conflict. In particular, Ukrainian authorities have stopped producing regional GDP estimates since the war outbreak. Such estimates could be useful to scholars studying the conflict and international organisations involved in the war on the ground (including those providing humanitarian aid), who could use them to decide where to allocate their resources.

The main goal of this paper is to understand how the GDP at the subnational level in Ukraine has changed since the full-scale Russian invasion began. To achieve this, I train a machine learning model that maps NTL data onto GDP at the regional level for all Ukrainian regions and predicts GDP in 2022 using that mapping. Figure 1 captures the main idea behind this paper. The image on the left-hand side shows how NTL radiance intensity is related to Ukraine's development. By looking at the map, we can identify all major cities (such as Kyiv in the centre-north, Kharkiv in the north-east) and roads (such as the east-west E40 route going through Lviv, Rivne, Zhytomyr, Kyiv, Poltava, Kharkiv, and Luhansk). If we zoomed in more, we could identify smaller towns, regional roads, and infrastructure such as power plants or airports, providing valuable information about economic activity in different parts of Ukraine. I intend to use this association between

NTL and economic activity to produce post-war GDP estimates at the regional level.

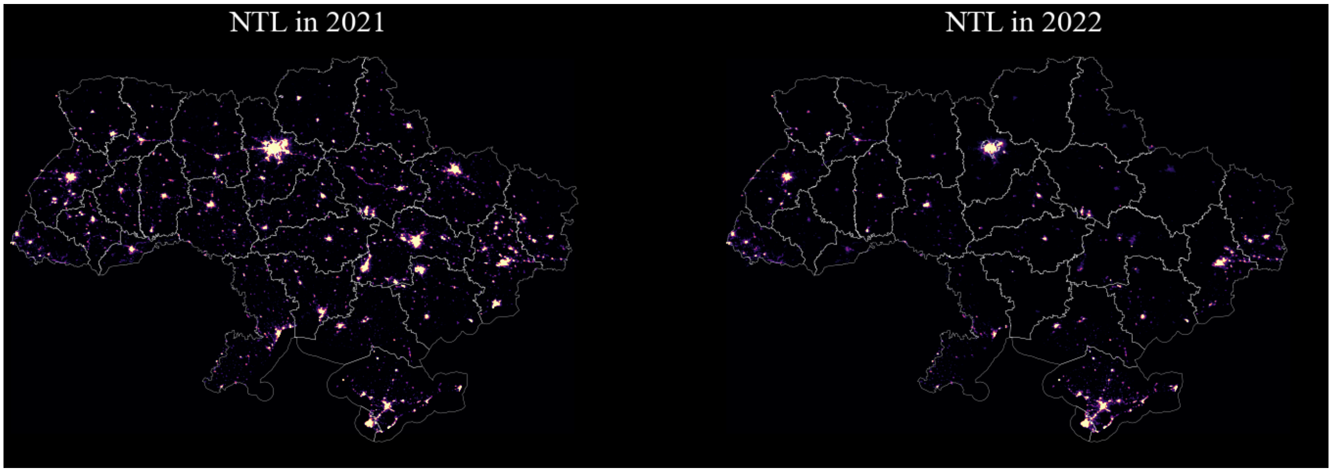

Figure 1: Comparison between Nighttime Lights in Ukraine before and after the Russian invasion.

The difference between the image on the left (pre-war NTL) and the image on the right side (post-war NTL) is staggering. Many regions have experienced a substantial drop in NTL radiance, and some cities, such as Mykolaiv in the south, seem to have disappeared. Nearly all connections are barely visible, suggesting fewer people travelled in 2022 than in 2021. Cities such as Lviv in the east remained visible despite a drop in the NTL radiance, particularly in the suburbs. Judging by the images only, Zakarpattia in the southwest and Crimea in the south, which has been under Russian control since 2014, seem to have been the least affected by the war outbreak. Although the extent of the decrease in NTL radiance varies from region to region, nearly all became darker at night. The consequences of war are visible to the naked eye.

Economists have been using NTL radiance to predict GDP for over a decade. Methodologically, most papers use some aggregate measure of NTL radiance (such as the sum) and a fixed effects model to predict GDP (Chen & Nordhaus 2011, Henderson et al. 2012). One limitation of this approach is that it does not account for the distribution of NTL radiance across the unit of observation and the spatial aspect of the data. When aggregating the data in that way, some information (such as the location and size of the cities and connections between them) is lost, which likely leads to less accurate predictions. Moreover, a simple regression model cannot capture more complex, non-linear relationships and likely underfits when predicting regional GDP using NTL radiance.

Remote sensing scholars have implemented more sophisticated methods, particularly machine learning techniques, and often achieved more accurate results than economists using regression models. Unfortunately, the overlap between these two fields is limited despite numerous commonalities.

In this paper, I propose a more complex methodological approach to this prediction problem to overcome some shortcomings of fixed effects models, hoping to bridge the gap between economics and remote sensing. I use a neural network and a more sophisticated method of aggregating NTL images into features to capture nonlinear relationships and retain more information about NTL radiance. I compare the performance of this approach against the fixed effects model and machine learning models commonly used in the literature. Additionally, I explore whether a novel approach, a CNN, yields better results. CNNs have been successful in pattern recognition since they utilise the image-like data structure better by reducing dimensionality whilst retaining the key features of the data. In this paper, I check whether they can be applied to the time-series prediction of regional GDP using NTL. Lastly, I explicitly address the multitude of available NTL radiance estimates by checking how the predictions change depending on the composite used to improve our understanding of NTL data in the economic context.

The primary goal of this paper is descriptive inference. Its main contribution lies in describing a previously unknown feature of our social reality, namely how GDP at the regional level has changed in Ukraine since the war outbreak. Additionally, it has a methodological contribution by exploring whether more complex methods produce more accurate time-series predictions of regional GDP than linear models, which remain a commonly used technique. Lastly, it improves our understanding of NTL data by exploring how the model's performance and predictions change depending on the composite used.

The paper is structured as follows. The next section reviews the literature on using NTL radiance data to predict GDP. Section 3 describes the data used, the cleaning and processing steps, and the methods used to arrive at the predictions. Section 4 presents the main results, and section 5 includes a validation exercise and checks how the models' performance and predicted GDP change depending on the NTL data used. Section 6 demonstrates how my predictions could be used to answer an applied research question. Section 7 discusses the main contributions, limitations, and possible extensions. Section 8 concludes the paper.

2 Literature Review

Subnational GDP estimates provide useful information to scholars, policymakers, and the public. Economists and social scientists use them to study the reasons behind cross-sectional variation in economic variables within a country. Government and international organisations can use them to evaluate the current development status and make more informed decisions about where to direct public resources to facilitate future growth. Firms might find subnational GDP estimates useful when making investment decisions.

Despite the importance of reliable GDP data at the regional level, its availability varies across countries. Although publishing reliable GDP estimates at the subnational level annually is a norm in most developed countries, many countries do not produce them regularly (Jerven 2014), leading scholars to explore methods of predicting regional GDP from available information. One of the approaches commonly used in the literature is to estimate subnational GDP using NTL. Satellites able to measure anthropogenic lighting have been orbiting the Earth for roughly 50 years. Initially, the primary reason for installing these satellites was to observe clouds for short-term weather forecasting - they captured the NTL radiance incidentally. Data extracted from these satellites has been widely used for research since 1992, when an archive of nighttime light images from the DMSP was made available (Gibson et al. 2020). However, since obtaining this data was not the primary objective of the satellites that captured NTL radiance, the estimates are noisy and often require extensive preprocessing before scholars can use them for scientific endeavours. Despite these difficulties, researchers within and outside the social sciences have used the DMSP data extensively for the past 30 years (Gibson et al. 2021, Zhao et al. 2017).

The two most influential papers that use the DMSP data and explore the possibility of using NTL data for measuring economic activity are Henderson et al. (2012) and Chen & Nordhaus (2011). Both studies recognise the difficulties associated with using the DMSP data for predicting economic activity, but they disagree on the extent of usefulness of this approach. Henderson et al. (2012) claim that predicting economic activity using the DMSP data can be insightful for a wide range of economies with varying levels of economic development and quality of available data, whereas Chen & Nordhaus (2011) claim the use cases are limited only to economies where the available economic data is of extremely poor quality (they suggest less than 40 countries would benefit from it). Subsequent studies in economics shed more light on the shortcomings of using

DMSP data (Addison & Stewart 2015, Bickenbach et al. 2016, Chen & Nordhaus 2015, Bennett & Smith 2017).

As the number of potential use cases of NTL data has increased, so has the demand for high-quality estimates of NTL radiance that would address the problems researchers faced when using the DMSP data. From 2012 onwards, data from the VIIRS has become available, and scholars across multiple disciplines have been rapidly switching to using this newer and better data (Gibson et al. 2020). Studies have shown that VIIRS data is better at predicting regional economic outcomes than DMSP data. Nonetheless, the accuracy of estimates remains relatively low for rural areas (Elvidge et al. 2013, Chen & Nordhaus 2015, Gibson et al. 2021) and for time-series GDP prediction relative to cross-sectional GDP prediction (Chen & Nordhaus 2019). A further improvement was introduced in 2020 when NASA published the Black Marble algorithm (Román et al. 2018), which produced data adjusted for atmospheric and BRDF effects. The result is the VNP46A2 dataset (updated daily), which is the most accurate NTL dataset to date (Alahmadi et al. 2023).

Perhaps the most significant advantage of the VNP46A2 dataset is its availability and granularity. Satellites orbiting around the Earth take images of the entire planet daily, which means NTL data is available for all countries, regardless of their development level and regime type. This data is gathered at a pixel level, with over 3 billion pixels covering the entire planet. Each pixel corresponds to an area smaller than one square kilometre, which allows us to predict GDP even for small geographical units.

A crucial feature of the VNP46A2 dataset, which is often omitted in the literature, is that there are different types of estimates. Firstly, the position of the satellite relative to the surface of the Earth changes. In general, the distinction is drawn between near-nadir, off-nadir, and all-angle estimates (nadir is the point directly below the satellite), with near-nadir estimates usually taking higher values (Román et al. 2018). Secondly, estimates can be either snow-free or snow-covered, with the latter being higher since snow reflects light. Lastly, other atmospheric conditions (such as clouds and rainfall) can also impact the estimates, which led scholars to add a binary quality flag (high-quality or all-quality) to all NTL radiance estimates. Scholars tend to use all-angle, snow-free, high-quality estimates due to their consistency, although the reasons behind choosing this particular satellite angle category are unclear. Exploring the differences between the NTL composites and how they affect predictions would improve our understanding of the VNP46A2 dataset and help us make more informed methodological choices when using it in the social sciences.

Scholars have used NTL data to predict economic activity for different countries with varying levels of economic development and regime types that do not produce reliable subnational GDP estimates for various reasons. Most commonly, it is due to a lack of resources (material or human) to conduct research necessary to obtain this information. Many countries in Sub-Saharan Africa and Latin America report annual GDP only at the national level. Scholars have used NTL data to predict subnational GDP in Uganda (Wang et al. 2019), Paraguay (McCord & Rodriguez-Heredia 2022), Kenya and Rwanda (Bundervoet et al. 2015), among other countries. Another reason for estimating regional GDP using NTL data is the unreliability of the existing estimates, which can be a result of a large informal sector (Tahsin 2022) or an authoritarian regime's attempt to inflate the perceived development of their country (Kim 2022, Martínez 2022). Finally, researchers have used NTL data when a government failed to produce GDP estimates due to an external shock such as the COVID-19 pandemic (Dasgupta 2022), natural disasters (Pagaduan 2022), and wars

(Dodd 2021, Wijesekera 2023). In particular, scholars have described the changes in NTL radiance in Ukraine since the war outbreak (Wang et al. 2024, Ialongo et al. 2023), and applied it to study the changes in refugee population (Huang et al. 2023). However, to my knowledge, no studies (to date) have aimed to predict post-war GDP in Ukraine at the regional level.

Methodologically, most economists use a panel regression framework, often utilising auxiliary data to predict GDP (Bennett & Smith 2017, Chen & Nordhaus 2019). That is likely driven by higher interpretability and the possibility of estimating the elasticity between GDP and NTL radiance. However, more recent studies within remote sensing have shown that the predictive accuracy of machine learning models is superior (Liu et al. 2021, Musthyala et al. 2022, Bansal et al. 2020). In particular, neural networks have been shown to outperform regression methods such as ARIMA in predicting GDP for years when a disruptive event (they use COVID as a case study) happened (Musthyala et al. 2022). Other scholars have achieved promising results when using CNNs to predict regional-level GDP using NTL and national GDP estimates (Liu et al. 2021). However, this method has not been used for time-series prediction during a disruptive event (such as a global pandemic or war). Some scholars have pointed out a surprisingly small citation overlap between economists and remote sensing scholars using NTL data (Bennett & Smith 2017). Applying methods from remote sensing in an explicitly economic context could help bridge the gap between the two fields, which will hopefully help both communities to collaborate more closely.

3 Data and Methodology

3.1 Data sources

In this paper, I use data coming from five secondary sources. The data collection process adheres to the Research Ethics Policy of the London School of Economics and was approved by the university's Research Ethics Committee. The first source is the VNP46A4 product, which I accessed via the NASA Earth Data portal (Román et al. 2018). The VNP46A4 product contains 460 files for each year, each corresponding to one non-fill tile from the MODIS sinusoidal tile grid. Each tile covers approximately $10° \times 10°$, and all 460 tiles cover the entire Earth. The geographical scope of my analysis is limited to Ukraine and Poland (I use the latter for validation), and the surface area of these two countries lies within seven tiles. I use the data spanning from 2012 to 2022 (2012 is the first year for which the annual product is available), bringing the total number of files I accessed from the VNP46A4 product to seventy-seven.

Each file contains 28 layers that describe the various aspects of NTL radiance for a given year in the area corresponding to a particular tile from the grid. Each layer is a 2400 x 2400 matrix with entries corresponding to a geographical area of 15 arcsec x 15 arcsec (approximately 500m x 500m at the equator). The layers of interest contain information about NTL radiance depending on zenith angle categories of the satellite (off-nadir, near-nadir and all-angle), snow status (snow-covered or snow-free), and the quality of the annual estimates (all-quality estimates or high-quality estimates only). There are twelve NTL composites in total. At its core, this dataset is a snapshot of nighttime lights for a particular region in a year, with several channels capturing different luminosity types. The remaining 16 datasets include the number of daily observations used to arrive at the annual estimate, standard deviation, land water mask, platform, latitude and longitude.

The second source is the State Statistics Service of Ukraine, which provides real GDP estimates (in hryvnia, the Ukrainian currency) for the years 2004 – 2021 for each of the 27 first-order administrative districts of Ukraine, 24 of which are regions (ukr. oblast) and 3 are special areas (cities of Kyiv, Sevastopol and Autonomous Republic of Crimea). Due to the annexation of Crimea, data for the Autonomous Republic of Crimea and the city of Sevastopol is missing from 2014, so I exclude these two regions from the analysis.

The third source is Statistics Poland (also known as the Central Statistical Office of Poland), which provides information about real GDP estimates (in zloty, the Polish currency) for the years 2012 – 2022 for each of the 16 first-order administrative districts of Poland (pol. województwo). Additionally, I extracted the share of agriculture in each region's GDP for 2021. This dataset is used in Section 6 for one of the validation exercises.

The fourth source is ACLED, whose Ukraine Conflict Monitor describes the ongoing war. Its datasets include information on how often different types of military events (bombings, battles, violence against civilians, etc.) occurred in each region and the number of fatalities associated with them. I use the ACLED dataset in Section 7 to illustrate how my regional-level GDP estimates can be used to answer an applied research question.

The final source is Cartography Vectors, where I obtained polygons for all regions in Poland and Ukraine, together with their coordinates. I use these polygons to build region-level NTL radiance images and produce maps.

3.2 Data preparation

There are two main aspects of the data preparation process. The first one is aligning the unit of analysis. Since my goal is to predict regional GDP, variables across all datasets need to relate to a specific region and year combination. To achieve this, I merge all seven NTL radiance datasets for each year and filter the tiles through administrative boundaries, which gives me 40 regional NTL radiance images per year (24 for Ukraine and 16 for Poland). Each image contains all twelve NTL composites as separate channels. Secondly, I sum up the number of peaceful protests, fatalities resulting from battles, and fatalities resulting from violence against civilians in every Ukrainian region in 2022. The real GDP is already provided at the regional level and does not require further preparation.

The second aspect is preparing the data for modelling. I begin by computing real GDP at the regional level for Poland and Ukraine in 2012 prices in the local currency. To prepare the regional NTL images for the CNN, I centre each image and add padding around (entries with values of 0) so that all images have the same shape. The resulting images are 765 x 1076 matrices for Ukraine and 609 x 911 matrices for Poland. I repeat the above process for all 12 composites defined by different satellite angles, snow status, and data quality combinations.

Afterwards, I calculate the sum of NTL radiance for every regional image. To provide additional information about the light luminosity, I define ten intervals of equal length beginning at 0 and ending at $M = \max(\ln(x_i) : x \in \text{NTL}_{L,c})$, where $\text{NTL}_{L,c}$ are all NTL radiance values for layer L (12 possible composites) and country c (either Ukraine or Poland). All 10 intervals cover the entire $(0, M)$ range, with the k-th interval ranging from $\frac{(k-1)M}{10}$ to $\frac{kM}{10}$. I took the natural logarithm before defining the intervals because the distribution of NTL radiance has a substantial right skew;

the median for high-quality all-angle NTL radiance values for Ukraine equals 6, and the mean is 18.8, with some pixels taking values over 20,000.

I computed the above measures after adding padding to the images to contextualise the information by accounting for the relative size of regions. A high number of pixels with an NTL radiance of 0 can indicate that the region is relatively small (and zeros come from the padding), which could be valuable information for predicting the change in real GDP. The approach of calculating the number of low- and high-radiance pixels for each region for a range of radiance intensity levels is similar to the one proposed by Jasiński (2019). I use this data format to train the linear, XGBoost and Random Forest models and the Neural Network.

3.3 Methodology

I train six different models to predict regional-level GDP in Ukraine in 2022. I use all-angle snow-free high-quality NTL radiance estimates for 2012 – 2021 and the corresponding GDP values in Ukraine as the training set. Using snow-free estimates makes the predictions robust to the snowfall difference between the test year and training years. Since snow reflects light, a higher snowfall in a test year relative to training years would artificially overstate the GDP. Similarly, using high-quality data prevents the estimates from being influenced by atmospheric conditions such as cloud coverage and rainfall. To embed my work in the existing research, I decided to use all-angle estimates since this is the predominant approach in the literature. I discuss how using different NTL radiance composites affects the results in more detail in Section 5.

I estimate every model for all years in the 2013-2021 range[1] (this includes hyperparameter

[1]For models using first differences as predictors and the training objective, 2012 cannot be used for testing or training since the difference between 2011 and 2012 is unknown.

tuning for some models), each time using all observations from a given year as the test set and observations for all remaining years as the training set. Afterwards, I calculate the mean absolute error for the best-performing model for each year and use this metric to select the model I later employ for prediction.

The first attempts to predict economic activity with data on lights at night used a linear model with a single independent variable capturing the intensity of NTL radiance with region and year fixed effects (see Henderson et al. (2012) for an example). Although presently, many scholars use more advanced modelling techniques, linear models remain an important reference point (Dasgupta 2022). I estimate two linear models, which take the following forms:

$$\begin{aligned} \ln(GDP_{i,t}) &= \beta_0 + \beta_1 \ln(sumoflights_{i,t}) + \gamma_i + \epsilon_{i,t} \\ \Delta GDP_{i,t} &= \beta_0 + \beta_1 \Delta sumoflights_{i,t} + \sum_{k=1}^{10} \delta_k \Delta \log(count_{i,t})^k + \gamma_i + \epsilon_{i,t} \end{aligned} \tag{1}$$

Where $GDP_{i,t}$ and $sumoflights_{i,t}$ are the real GDP and sum of NTL radiance for region i in year t, γ_i are region fixed effects, $\log(count^k_{it})$ is the number of pixels in the k^{th} category for region i in year t, and $\epsilon_{i,t}$ is the error term. Since the objective is time-series prediction, I excluded time fixed effects from the models.[2] The first model is a log-log model, which means β_1 can be interpreted as the elasticity of the real GDP with respect to the sum of NTL radiance. The second model is a first differences model, which means all variables are calculated according to the following formula:

$$\Delta X_{i,t} = X_{i,t} - X_{i,t-1} \tag{2}$$

Both the log-log model and the first differences model have been widely used across various fields

[2] I also tested an alternative approach of including time fixed effects and setting the coefficient to 0 for the test year, which achieved a worse prediction accuracy.

within economics (including predicting GDP with NTL data), so they are a natural starting point. In this paper, I use them as a benchmark against which I judge the performance of other models.

Model	**Parameter**	**Values**
XGBoost	Eta	0.1, 0.2, 0.3, 0.4
	Gamma	0, 5, 10, 20
	Max. depth	4, 6, 8, 10
	Min. child weight	3, 4, 5, 6
Random Forest	Number of trees	100, 200, 300, 400
	Max. depth	5, 10, 15, 20
	Min. samples split	2, 4, 6, 8
	Min. samples leaf	2, 4, 6, 8

Table 1: Hyperparameter grids used to tune the XGBoost and Random Forest models.

Another two sets of models are XGBoost and Random Forest, which have been used in more recent literature (Bansal et al. 2020, Otchia & Asongu 2019). They are ensemble learning methods based on decision trees, which means they combine multiple models (trees) to improve performance relative to individual models and prevent overfitting. I train the models with the difference in the sum of NTL radiance, differences in the ten log-transformed counts, and regional dummies as predictors (the predictors are the same as the independent variables in the first differences model) with the MSE as the loss function and the difference in real GDP as the training objective. I use an exhaustive grid search across a range of hyperparameter values and 5-fold cross-validation to select the best-performing hyperparameter combination for each year (these might change over time). Table 1 summarises the hyperparameter values used to tune the models.

The fifth set consists of neural networks, which have been successfully implemented in predicting GDP using NTL in developed countries (Musthyala et al. 2022). Similarly to the ensemble methods, I train the networks using independent variables from the FD model, the difference in real GDP as the training objective, and MSE as the loss function. I use a batch size of 64 for a maximum of 500 epochs, allowing for early stopping if the performance on the validation set starts to degrade, which helps to avoid overfitting. Since neural networks are sensitive to the scale of inputs, I standardised the parameters before training (XGBoost and Random Forest models were unaffected by standardisation). To tune the model, I experimented with several different architectures and selected the one that led to the most accurate predictions as measured by the MSE. Figure 2 describes the final network architecture. Adding additional dense layers, dropout layers or reducing the depth of the network didn't yield better results.

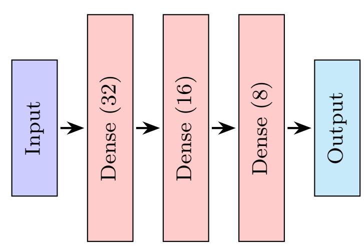

Figure 2: Neural Network Architecture.

The final model consists of two stages. In the first stage, I use NTL radiance images and the corresponding real GDP values to train a CNN that maps those images onto 10-dimensional vectors. I rescale the images and the dependent variable so that all values fall in the (0, 1) range by dividing the NTL radiance estimates and real GDP by the corresponding maximum values. I train the CNN for 100 epochs using a batch size of 64, again allowing for early stopping. In the second stage, I take the first differences of the 10-dimensional vectors, the difference in the sum of NTL radiance, and regional dummies to predict the change in real GDP using a second neural network. The architecture and the details of the training process for the second network are the same as for the fifth model (the only difference is the features I use for training). This approach allows me to assess whether CNN-extracted features perform better than simple NTL radiance measures (the number of pixels in each radiance intensity category). The rationale behind the final model is to utilise the image-like structure of the NTL data in the first stage to extract the most relevant information about the NTL radiance given the training objective and combine these characteristics with regional dummies and the NTL radiance sum to account for the cross-sectional nature of the data. Although CNNs have been used to produce subnational GDP estimates from national-level data, their applicability to time-series prediction remains unknown (Liu et al. 2021).

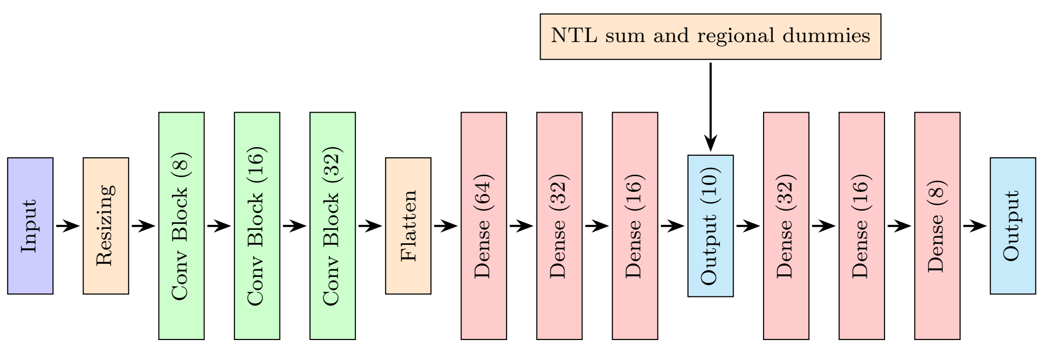

Figure 3: CNN architecture. Each convolutional block consists of a convolutional layer with a 3x3 kernel, a max pooling layer with a 2x2 filter, and a batch normalisation layer.

I repeated the above steps for several different CNN architectures and selected the one which resulted in the lowest MSE. Figure 3 summarises the network architecture used. Experimenting with different kernel sizes, a different number of convolutional and dense layers, the number of neurons in a layer, and the size of the feature vector didn't yield better results.

4 Results

4.1 Model performance comparison

Figure 4 presents the predictive accuracy of the six models as measured by MAE for 2013 – 2021. The benchmark log-log model achieved an average MAE of 3,879 UAH between 2013 and 2021. For reference, the average real GDP at the regional level equalled 63,567 UAH over the same period. Perhaps unsurprisingly, the simplest model has the lowest prediction accuracy. To estimate the elasticity of GDP with respect to the sum of NTL radiance, I retrain the model on the entire dataset with time fixed effects. The R-squared of the model equals 0.97 with β_1 coefficient equal to 0.64 and statistically significant at the 0.05 level (the t-value was more than 9). The result indicates that, on average, a 1% higher sum of NTL radiance is associated with a 0.64% higher real GDP after controlling for time and regional fixed effects. Providing more detailed information about the distribution of NTL radiance by adding log-transformed pixel counts and predicting the difference rather than the level of real GDP improves the performance of the linear model substantially. The FD model has achieved an average MAE of 2,832 UAH, which is approximately 27% lower than the log-log model. Additionally, it is the most precise[3] out of all models, with the standard deviation of MAEs across 2013-2021 equal to 490 UAH.

Employing ensemble methods yields mixed results. On the one hand, the XGBoost model

[3]In this section, I use the word *precise* to describe models' MAE variability across years. The more precise a model is, the closer its MAEs for different years are. Equivalently, more precise models have a lower standard deviation of MAEs.

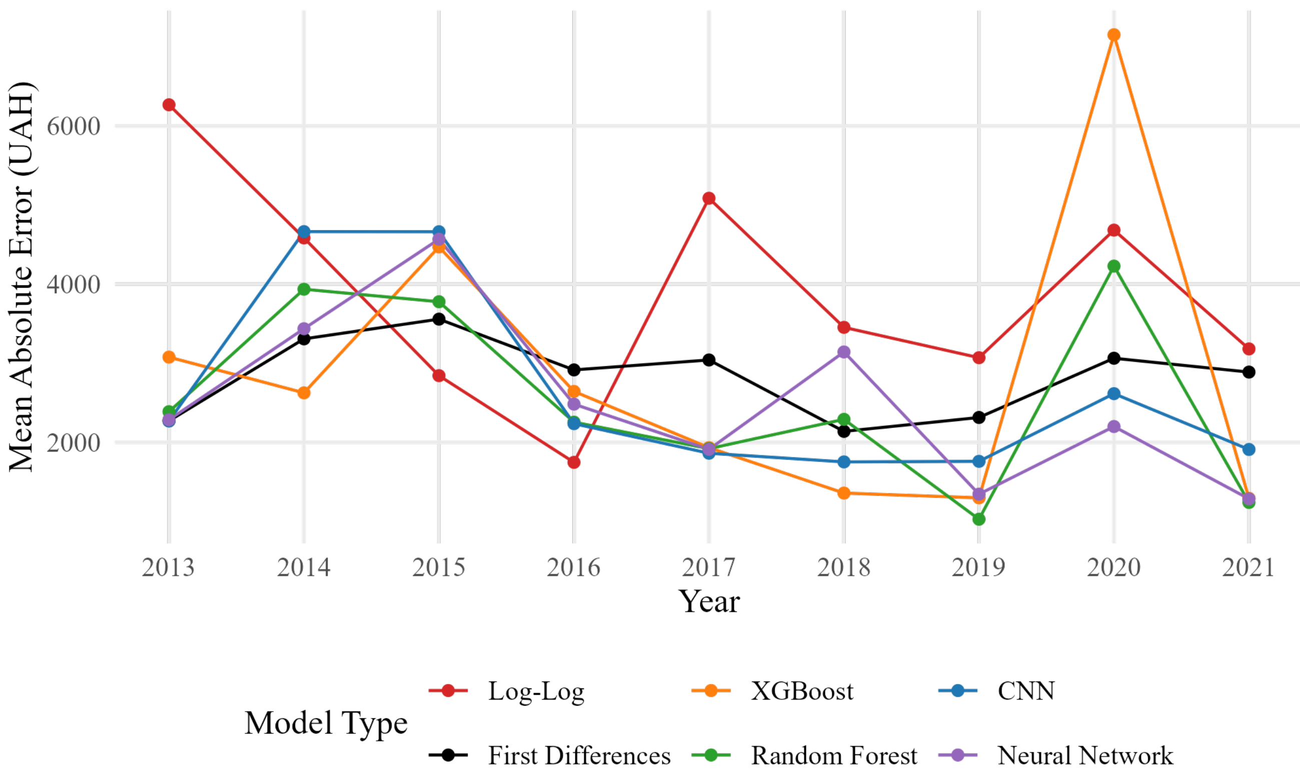

Figure 4: Accuracy comparison between the models used.

achieved an average MAE of 2,871 UAH, which is worse than the FD model. Its precision is the worst of all models, with the standard deviation of MAEs equal to 1,911 UAH. On the other hand, the RF model achieved an average MAE of 2,561 UAH, which is 10% lower than the FD model, and its precision is comparable to the CNN. The XGboost's low accuracy is primarily due to the model's poor performance for 2020, which was the year when Ukraine introduced lockdown to prevent the spread of SARS-CoV-2. RF's accuracy for that year is higher than XGboost's, but its performance in 2020 is inferior to both neural networks and the FD model.

A notable increase in prediction accuracy in 2020 is a common feature across most models. This is likely due to the pandemic and the policies implemented in response, particularly lockdowns. During the pandemic, the relationship between NTL radiance and GDP differed from non-pandemic years because of restricted travel and outdoor activities. Some models may struggle to adapt to these changes, leading to poorer performance when confronted with shocks like COVID-19. Out of the models employed, the FD model appears to be the least affected by the pandemic's impact on accuracy.

War outbreaks might have a similar effect on the relationship between NTL radiance and GDP. For instance, turning lights off and living underground might be a way to avoid air strikes and artillery barrages. Such behaviour would likely lead to a substantial decrease in NTL radiance but may not have such a large negative effect on GDP, which could lead some models to overestimate the fall in GDP.

The CNN achieved an average MAE of 2,636 UAH, slightly underperforming compared to the RF model. This lower accuracy is primarily due to its performance in the years 2013–2015. However, from 2016 onward, CNN's performance has been relatively good and the most consistent of all models. Despite this, the neural network trained with log-transformed pixel counts emerges as the best-performing model overall, with an average MAE of 2,516 UAH, making it the most accurate among the six models across the entire 2013–2021 period. It is also the second most precise model for the whole period, surpassed only by the FD model. One limitation of this neural network is its relatively poor performance in 2018, which cannot be attributed to an external shock like the pandemic. Other machine learning models, particularly XGBoost and CNN, maintained their accuracy for that year. Nevertheless, I chose to use the neural network for predicting post-war real GDP in Ukraine at the subnational level due to its superior accuracy and precision over the 2013–2021 period compared to other models.

4.2 Regional GDP predictions

I retrain the neural network on the entire training set (with a 64 batch size and allowing for early stopping) using the best-performing architecture and predicted regional GDP in 2022. Figure 5 shows the predicted GDP change for Ukrainian regions.

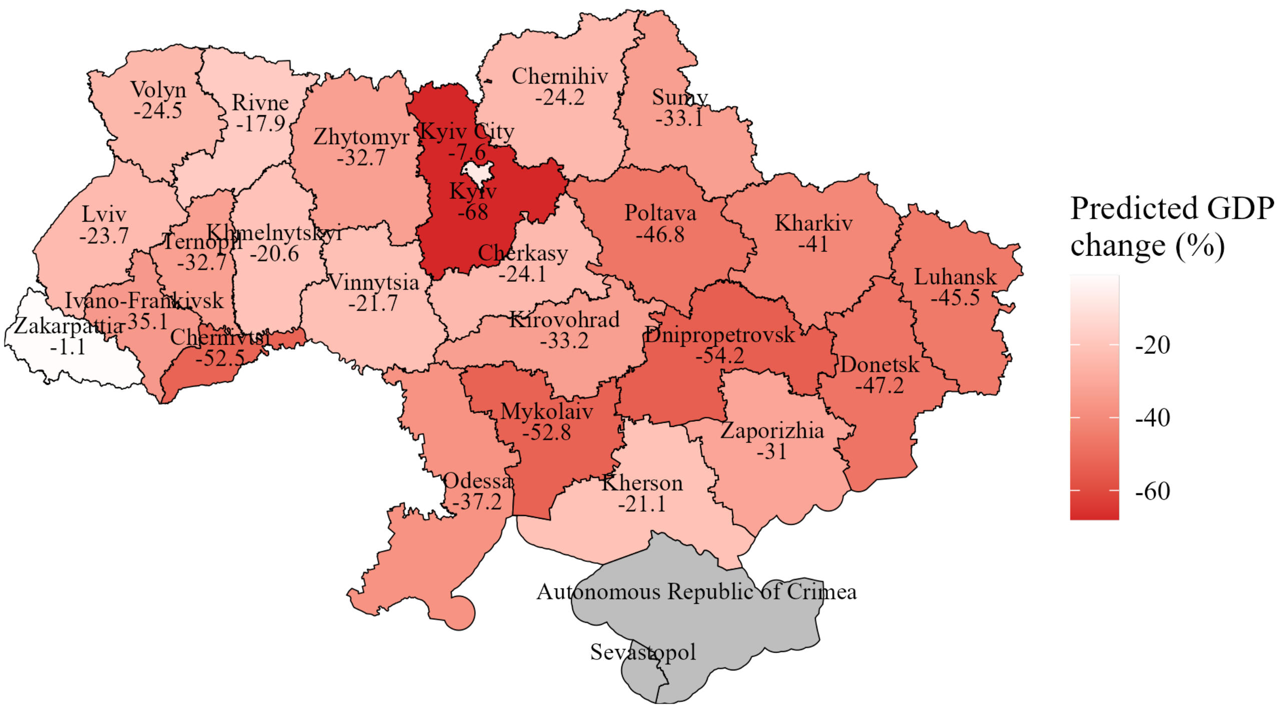

Figure 5: Predicted GDP change between 2021 and 2022 in Ukraine.

Predicted GDP change varies substantially by region, ranging from -68% to -1.1%. Generally, the projected decline in real GDP is more pronounced in the east than in the west, which seems plausible given that the majority of war activities occurred in eastern Ukraine. A notable exception is the Kyiv oblast, which includes a large proportion of the Kyiv metropolitan area. The high predicted fall in real GDP is likely due to the early stages of the war when the Russian forces attempted to take the capital, causing widespread destruction and forcing many residents to flee.

Another region that stands out is Zakarpattia, which has a predicted GDP change of -1.1%, approximately 20 percentage points more than any other Ukrainian oblast. To some extent, this reflects the region's placement - located in the southwest of Ukraine, it is relatively far away from the frontline. It is a mountain region with fairly small cities, which has not been directly affected by the war. Additionally, the core of Zakarpattia's economy is tourism and agriculture. Since NTL data is generally worse at predicting GDP for rural areas (agriculture is associated with less NTL radiance than other sectors), I expect the 1.1% decline in GDP to be the upper bound of the actual change.

To validate the predictions, I aggregated them and compared them against existing national-level estimates of Ukrainian GDP. In 2022, the IMF projected that Ukraine's GDP would contract by 35% due to the war (IMF 2022). Subsequently, the World Bank published its estimate, based on the same data used for my model training, indicating a GDP decline of 28.8% in 2022 (World Bank 2024). My estimates suggest a 31% contraction in the Ukrainian economy, which is 2.2 percentage points higher than the World Bank's estimate.

Multiple explanations could account for the fall in economic activity. Firstly, destruction can be caused directly by warfare. Cities such as Kharkiv, Mariupol, and Odessa experienced intense artillery barrages and bombings in the first months of the war as the Russian forces tried to take over the cities. Deaths and destruction caused by the offensive account for part of the decrease in economic activity.

However, this explanation alone is insufficient, as it does not account for why regions like Dnipropetrovsk or Chernivtsi, which were relatively less affected by direct warfare, still experienced a significant decline in real GDP. A second plausible explanation is the widespread attempts to escape the conflict, particularly through migration to neighbouring countries and internally. The UN estimates that approximately 6.5 million Ukrainians have left the country, with an additional 3.7 million internally displaced since the war began (UNHCR 2024). In Poland, most Ukrainian refugees are from Kharkiv, Dnipropetrovsk, Zaporizhzhia, and Lviv (UNHCR & REACH 2022). Residents of regions close to the frontline, such as Dnipropetrovsk, likely left the country out of concern that the conflict might eventually reach them, which would explain the substantial GDP decline in these areas.

Not everyone can flee a country during wartime. Factors such as socioeconomic status, education, and gender likely influence both the decision to migrate and the destination. For example, Syrian refugees in Germany tend to be better educated than those in Turkey (Pearlman 2020). Similarly, Ukrainian refugees in Poland are generally better educated than the local population, with many having pre-existing ties to Poland (Kowalski et al. 2022).

Many of those who did not flee abroad moved internally. In general, we would expect regions with higher net migration to experience a lower decline in economic activity. Additionally, the closer a safe region or city is, the more compelled people might be to move. Internal migration might explain why Kyiv city experienced a relatively modest decline in GDP, especially in light of the siege that took place in the early months of the war. If people living in the suburbs decided to flee to the city centre, which is further away from the frontline, they would offset part of the decrease in the city's GDP as long as they remained involved in economic activity there. This internal migration would account for both the substantial decline in GDP in the suburbs and the relatively modest fall in the centre.

Another important factor is the presence of national minorities in Ukraine. The large Russian-speaking population in eastern Ukraine and Crimea likely contributed to pro-Russian tensions leading up to the 2014 annexation of Crimea and set the stage for the current invasion (Mitrokhin 2015). Many Russians living in Ukraine have since fled due to the invasion, with Russia receiving the most Ukrainian refugees (UNHCR 2024). Similarly, the significant number of refugees from Lviv in Poland could be partially explained by the relatively large population of Ukrainians of Polish descent living in Lviv before the war, as well as the historical ties between western Ukraine and Poland (Burant 1993).

In every region, a variety of factors contribute to the changes in economic activity observed between 2021 and 2022. While we can hypothesize about which factors were most influential in each region—drawing from our understanding of the situation in Ukraine and the war—verifying verifying these hypotheses empirically is challenging. The difficulty stems from the ongoing conflict, which severely limits the availability of reliable data. As long as the warfare persists and data remains scarce, accurately determining the precise impact of each factor is not just difficult but, in many cases, impossible.

5 Validation and other NTL Composites

5.1 Predicting regional GDP in Poland

Given that war is a significant external shock that disrupts nearly every aspect of the affected country, including the relationship between NTL radiance and economic activity, it is important to scrutinize the predictive accuracy of models trained on data from periods without such shocks.

While the relatively modest increase in MAE during the pandemic year indicates that the model is quite resilient to such disturbances, in this section, I address this issue in more detail and provide further justification for the accuracy of my GDP predictions.

Firstly, I implement the approach I used to arrive at GDP predictions for Ukraine to another country, which has (to a certain degree) been affected by the war but kept producing reliable GDP estimates. Since the outbreak of the war, Poland has provided asylum to over 1.5 million refugees, accounting for 3.7% of the country's population (UNHCR 2024). In major Polish cities, populations have surged by 17% since the war began, with some cities seeing population growths of 30% or even 50% within just a few weeks (Wojdat & Cywiński 2024). This unprecedented influx of people has significantly impacted many aspects of daily life in Poland.

I begin by extracting NTL radiance estimates for each of the 16 Polish regions for 2012 – 2022 and matched them to corresponding real GDP estimates. The cleaning and processing approach was exactly the same as for Ukrainian regions, with the only two differences being the size of regional images and the currency of the GDP. I train a neural network with the same architecture as the one I used to compute the predictions for Ukraine on data for 2012 – 2021 and use it to predict the change in real GDP in 2022. This approach allows me to gain additional information about the performance of the model I used for prediction, but it should not be treated as a standalone attempt to predict regional GDP in Poland. The neural network architecture that yielded the most accurate predictions for Ukraine may not necessarily be the best-performing in the case of Poland.

Figure 6 compares my predictions for each region with the actual change as reported by Statistics Poland. According to the available data, Polish GDP has increased by 5.3% between 2021

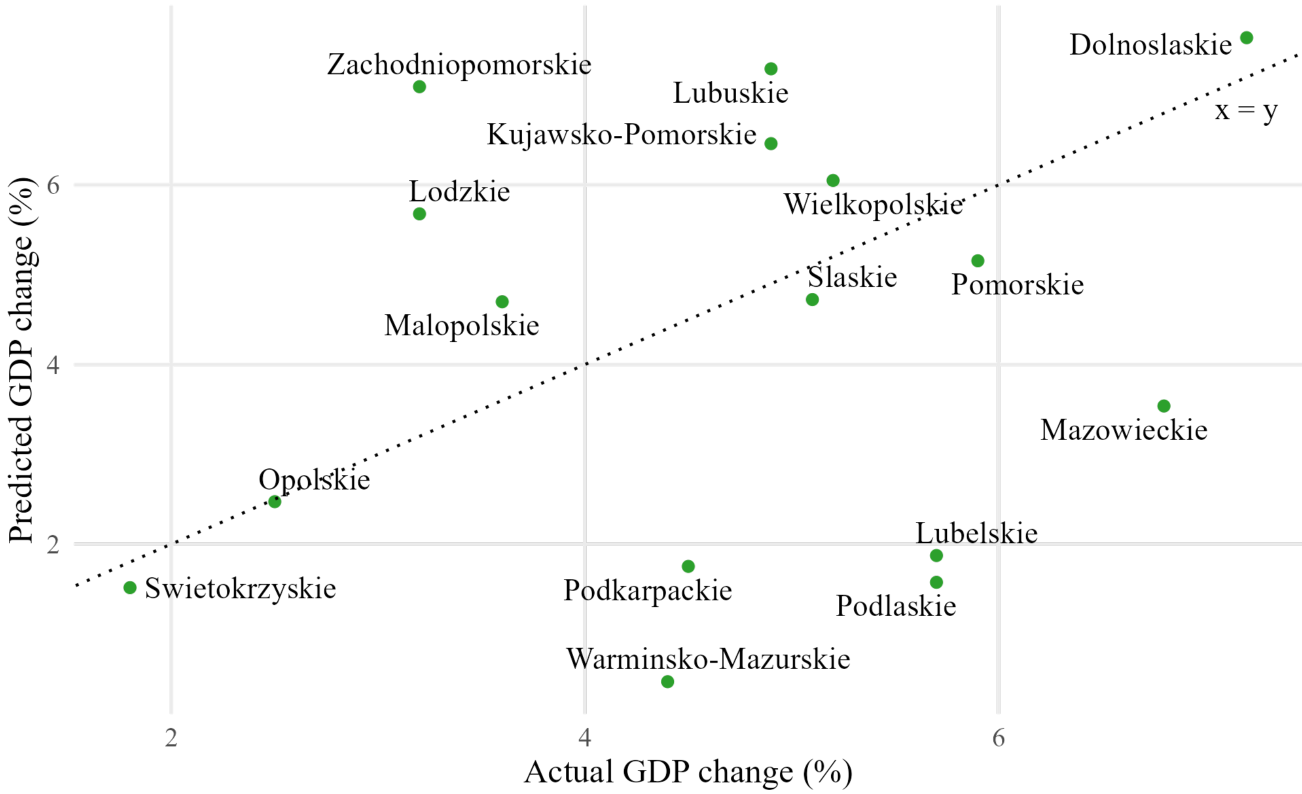

Figure 6: Predicted and actual GDP change between 2021 and 2022 in Poland.

and 2022. For the same period, my model predicted a 4.6% increase. Although the aggregate prediction is relatively accurate, there is a lot of variation in model accuracy between different regions. My predictions for the level of GDP deviate by 2,623 PLN from the actual level (the average level of GDP for Polish regions in 2021 equalled 138,314 PLN) or two percentage points on average. For six regions, the difference between the predicted and actual GDP change is less than one percentage point, and for five of them, it is more than three.

Overall, the model's performance is too poor to provide useful information about regional GDP in Poland, which regularly publishes reliable GDP data. This limitation notwithstanding, the model has correctly identified the GDP trend in many regions and can be helpful in understanding regional GDP dynamics if reliable estimates are unavailable.

Additionally, the variability in the model's performance can help understand where the GDP predictions are likely to be accurate. Previous research indicates that NTL radiance is a better predictor of GDP in urban than rural areas (Chen & Nordhaus 2019, Pérez-Sindín et al. 2021). The results for Poland are consistent with this finding. Podlaskie, Lubelskie, and Warminsko-Mazurskie are the three regions with the highest share of agricultural production in total GDP. On the other side of the spectrum are regions like Slaskie, Dolnoslaskie, and Pomorskie, which have a relatively small agricultural sector and more accurate predictions. To investigate this relationship in more detail, I estimate a linear regression model with the share of agriculture in regions' GDP as the dependent variable and the absolute percentage difference in the predicted and actual GDP as the dependent variable. This simple model explained over 30% of the variation in the prediction accuracy (ie. the dependent variable). The coefficient equalled 0.5 and was statistically significant at the 0.05 level, suggesting that, on average, a one percentage point higher share of agriculture in regions' GDP is associated with a 0.5 percentage point higher absolute difference between the actual and predicted GDP. For most rural regions, GDP predictions derived using NTL data are lower than the actual level of GDP.

I want to highlight two conclusions from this analysis. Firstly, the model most likely correctly identifies the main trends in regional GDP in Ukraine since the war outbreak, although precise figures should be viewed with caution. As shown by the example of Poland, there is substantial variability in the model's performance depending on the region, and the GDP predictions can be far from the actual value in some cases. Secondly, GDP predictions are probably less accurate for rural regions.

5.2 Other NTL composites

Most scholars using NTL data to predict economic variables have used the all-angle, snow-free and high-quality estimates, and this is the approach I have used throughout this paper so far. Various reasons were given to support the use of this composite, but (to the best of my knowledge) there has not been an empirical attempt to check to what extent using different NTL radiance estimates affects the predictions. In this section, I address this gap by training neural networks using other composites and predicting GDP for 2022 (the data preparation process and network architecture remain unchanged).

Table 2 summarises the MAEs of the same neural network trained on nine different composites for 2013 – 2021 and the predicted change in GDP for 2022. Due to data missingness, I could not train the model using snow-covered, high-quality estimates for all satellite angle categories. That should not come as a surprise since snow lowers data quality due to reflection. The differences in average MAE for the remaining nine composites are relatively small – they all fall within the (2516, 2788) range. Despite this, the GDP predictions for 2022 vary substantially between different composites.

In general, off-nadir composites perform the worst. Within each quality and snow coverage category, they achieved the lowest prediction accuracy as measured by the average MAE. On the other side of the spectrum are the all-angle composites, which have consistently outperformed all other satellite angle categories. This result indicates that combining near-nadir and off-nadir observations, which provide more information about NTL radiance than a single composite, yields the best results.

High-quality estimates generally outperform their all-quality counterparts on average, though

Composite	Average MAE	Predicted GDP change
All-angle, snow-free, high-quality	2516	-31.0%
Near-nadir, snow-free, high-quality	2656	-39.3%
Off-nadir, snow-free, high-quality	2711	-12.6%
All-angle, snow-free, all-quality	2639	-36.9%
Near-nadir, snow-free, all-quality	2639	-40.8%
Off-nadir, snow-free, all-quality	2663	-16.1%
All-angle, snow-covered, all-quality	2567	-7.7%
Near-nadir, snow-covered, all-quality	2715	-5.0%
Off-nadir, snow-covered, all-quality	2788	-6.4%

Table 2: Accuracy and prediction comparison between different NTL composites.

the latter exhibit more consistent performance across various satellite angle categories. When it comes to predictions, all-quality composites tend to yield more negative GDP estimates, averaging 3.6 percentage points lower than those from high-quality predictions. Notably, composites with snow cover have the lowest prediction accuracy, significantly understating the decrease in GDP compared to other composites. This inaccuracy is likely due to the high reflectivity of snow. In some regions, these snow-covered composites even predict a substantial GDP increase between 2021 and 2022—a result that appears implausible.

The above analysis suggests that all-angle, high-quality, snow-free composites are the most appropriate for predicting economic activity. They achieve the highest prediction accuracy and tend to produce more plausible predictions than other composites. Even though the differences in average MAE between different NTL composites are relatively small in absolute terms, these slight differences can result in substantially different predictions. To account for this, scholars should carefully select input data, taking into consideration the qualitative differences between different composites and how they affect the models' predictions.

6 Application

In this section, I demonstrate how my regional GDP estimates can be used to answer an applied research question about an ongoing conflict. Although war is as old as, if not precedes, the human civilisation, it has undergone massive changes throughout the centuries. For most of history, humans fought primarily hand-to-hand, with ranged weapons such as the bow and arrows playing a complementary role (Keeley 1997). That changed with the invention of gunpowder, which has led firearms to become the predominant weapon used in combat, including close combat. Later inventions, such as the dynamite, further increased the destructive potential of weapons. The rise of weapons of mass destruction in the 20th century has led civilians to become far more likely to become victims of war. The civilian casualty ratio (the ratio of civilian casualties to combatant casualties) has been increasing since the 19th century. Scholars estimate that over 90% of fatalities in modern wars are civilians, while at the end of the 19th century it was only 5% (Khorram-Manesh et al. 2021). Studying the relationship between different types of warfare and economic activity can help us understand the wider consequences of this historical trend.

Suppose a scholar is interested in learning about how different types of military engagement impact the economic activity of region that is affected by war and would like to use the Russo-Ukrainian war as a case study. One hypothesis she might put forward is that violence towards civilians has a more negative impact on economic activity than violence against soldiers. The reason behind it might be that deaths in battles feel more distant to civilians (who are the ones primarily involved in economic activity while the military is fighting) relative to explosions or acts of terror close to where the civilian population live.

To determine the relationship between economic activity and types of military engagement, I estimate a linear regression model with regional GDP predictions as the dependent variable and the real GDP in 2021, number of peaceful protests, battle casualties, and deaths from civilian violence as the independent variables. Table 3 summarises the results of this model.

	Predicted GDP change (UAH)
Battle fatalities (#)	-1.6 (1.2)
Fatalities from violence against civilians (#)	-43.2*** (10.8)
Peaceful protests (#)	583.9 (353.1)
Real GDP in 2021 (UAH)	0.80*** (0.05)

*p<0.05 **p<0.01 ***p<0.001

Table 3: Regression results. Each observation is a Ukrainian region, heteroskedasticity-robust standard errors are in parentheses.

The signs of the regression coefficients align with what we would generally expect. A higher number of casualties and deaths from civilian violence is associated with a higher decrease in GDP, which means that regions with more warfare directed at civilians (as measured by the number of fatalities) tend to experience a more substantial drop in GDP. Similarly, a higher number of peaceful protests (which happen only when people feel sufficiently safe to go outside) is associated with a smaller decrease in GDP. On average, one peaceful protest is associated with a 584 UAH higher GDP. Cross-sectional variation in the predicted change in GDP aligns with what we would expect based on where most of the fighting took place, which (to some extent) corroborates the proposed GDP predictions.

Perhaps the most intriguing finding from this simple regression model is the lack of association between battle fatalities and GDP decrease. The coefficient value is approximately zero and not statistically significant at the 0.05 level. When compared with violence against civilians, the difference is staggering. On average, one death from civilian violence is associated with a 43 UAH lower GDP. The Pearson correlation coefficient between these two independent variables is 0.21, and VIF values for all independent variables are below 2, suggesting that multicollinearity is not an issue. Although these results give some evidence for the proposed hypothesis in the case of Ukraine, further analysis that would include determining whether the model meets the assumptions of linear regression and additional robustness checks, which is beyond the scope of this paper, is necessary to establish this relationship.

The presented approach could be a starting point for more complex and robust research examining the relationship between different types of military engagement and economic activity. On their own, post-war regional GDP estimates for Ukraine are likely insufficient to conduct a study with robust findings due to a small sample size (25 observations). Nevertheless, they could be used in a similar fashion as part of a broader analysis that would include multiple conflicts. Alternatively, they could be used as one of the independent variables in a multiple regression model, providing insight into the impact of the GDP dynamics on the dependent variable.

7 Discussion

This paper contributes to the existing literature in three main ways. Firstly, it aims to estimate post-war regional GDP in Ukraine, hopefully allowing policymakers involved in the war to make more informed decisions and help scholars study it. Secondly, this paper explores a novel empirical method for predicting economic variables using NTL radiance data, taking advantage of the image-like structure of the data. Lastly, it explicitly analyses the extent to which the model can be applied to another country and how the predictions change depending on the NTL radiance estimates used.

The proposed GDP predictions are based on the most accurate of several models and seem plausible given where most warfare took place. The prediction at the national level is close to existing estimates, particularly the one put forward by the World Bank. In general, I would expect the proposed GDP predictions to overstate the actual decrease in GDP. That is because, during a war, many people might adapt their way of living, trying to draw as little attention as possible to avoid being targeted by the enemy. Moving a large proportion of economic activity underground would result in a substantial fall in NTL radiance without the corresponding fall in real GDP, leading models to overstate the decrease in real GDP.

A major advantage of the proposed approach lies in its simplicity. Producing reliable GDP estimates is costly, especially during a war. In contrast, my predictions are based on relatively small and publicly available datasets. The entire preparation and training process is simple and not computationally intensive. Given these characteristics, the presented approach is a low-cost alternative to producing regional GDP estimates during crises. Despite this, it is unlikely to provide valuable information for countries that do not experience some disruption.

Methodologically, this study suggests that using machine learning models, particularly a neural network with a relatively simple architecture, leads to more accurate time-series predictions relative to a simple linear regression model. That is likely due to the network's ability to capture more complex, non-linear relationships.

Using features extracted with a CNN did not improve the model's performance relative to log-transformed counts. That might indicate that the relationship between NTL data and real GDP is not that strong, and using simple aggregate features for prediction leads to accuracy close to the best possible accuracy. More complex features of NTL images add little information relevant to predicting regional GDP. Another reason behind CNN's poor performance might be a relatively small training set. With many parameters, CNNs generally need a lot of data to perform well out-of-sample. For small datasets, such as the ones used in this paper, the CNN might overfit the training data, leading to poorer performance on the test set. I would not expect a more complex CNN to yield better results for similar reasons.

When it comes to using different NTL composites, this study suggests the standard approach of using all-angle, snow-free, high-quality estimates is indeed the most appropriate for predicting regional GDP. A possible idea for further research in this area would be to use a combination of near-nadir and off-nadir composites, perhaps as two separate channels, as input for a neural network.

The main weakness of my predictions lies in the fundamental difference between the training and prediction data. The former describes pre-war years, whereas the latter consists of post-war observations. That is problematic since war has most likely disrupted the relationship between NTL radiance and real GDP. Even an extremely well-performing model that can capture the relationship pre-war may not be applicable if the disruption is substantial.

Secondly, the dataset I used to train my model is relatively small, especially for neural networks, which show their true potential when trained on large amounts of data. That is probably why more complex networks performed relatively poorly – they overfitted on the training data and produced less accurate estimates for new observations. One way to overcome this limitation would be to increase the number of data points by disaggregating the regional GDP into lower areas using NTL data. This approach was implemented for other countries to compute sub-national GDP (Wang et al. 2019). Afterwards, subregional GDP estimates and corresponding NTL data could be used to train a model, potentially improving the accuracy.

Another limitation lies in the short timeframe of the study - I produced predictions for only one year ahead. At the time of writing this paper, the war in Ukraine has been ongoing for over two years. Scholars studying the conflict and institutions involved in it might be interested in regional GDP predictions for 2023. However, since reliable data for 2022 is not available, the prediction for 2023 would have to be made using data from 2012 to 2021.

There are several possible ideas for further research stemming from this paper. The example of Poland suggests the proposed method can produce relatively accurate predictions in other contexts (without additional architecture tuning, which would likely lead to even better results). Implementing it for other countries that stopped producing reliable GDP estimates due to war would allow scholars and policymakers to understand the economic consequences of military conflict in those countries better.

Another improvement could come from using auxiliary data, particularly daytime satellite imagery, to improve the quality of predictions. Scholars found daytime satellite imagery to be a good proxy of economic activity for East Germany before the fall of the Iron Curtain (Lehnert et al. 2023). Some evidence suggests that combining daytime satellite imagery with NTL data provides more accurate estimates of economic activity, especially when using CNNs (Liu et al. 2021). Additionally, scholars could use national-level GDP as one of the input variables and model the distribution of GDP (rather than the level).

8 Conclusion

This paper investigated how the regional GDP in Ukraine has changed since the Russian invasion in February 2022 using NTL data and machine learning. I found that all regions experienced a decline in GDP of varying degrees, with those in the east noting a more pronounced decline. According to my estimates, the Ukrainian economy contracted by 31% between 2021 and 2022. Factors such as destruction caused by warfare and internal and external migration might explain the differences in the size of the decline in economic activity between different regions.

The best-performing model was a neural network trained using log-transformed pixel counts. Its performance is relatively consistent throughout the years, including the pandemic year. Linear regression methods, commonly used by economists, achieved a worse prediction accuracy. Ensemble machine learning methods outperformed linear models but did not do better than the neural network. Using a CNN for feature extraction did not improve the prediction accuracy, most likely due to a small training sample and a weaker-than-expected association between NTL data and economic activity.

To validate the approach, I implemented the model for Poland, which experienced a large influx of refugees due to the war in Ukraine but continued producing reliable GDP estimates throughout 2022. The model correctly identified the GDP trajectory for most regions. The difference between the actual and the predicted GDP level equalled two percentage points on average. Additionally, the analysis confirmed the finding from existing literature - NTL data is worse at predicting GDP in rural regions, usually understating it.

Additionally, this paper explored how the model's performance changes depending on the NTL composite used. It provides empirical support for using all-angle, snow-free, high-quality NTL estimates. All other composites achieved a lower prediction accuracy for the training period, and the GDP estimates they produced were less plausible.

Finally, I demonstrated the applicability of my predictions by investigating how different types of warfare affect economic activity. I found some evidence suggesting that violence against civilians has a more negative impact on GDP relative to battles between soldiers.

The main limitation of this paper lies in the fundamental difference in the relationship between NTL and economic activity before and after the war. Further research could explore using auxiliary data such as daytime satellite imagery or national GDP estimates to improve the quality of predictions. Despite its limitations, the proposed method arrived at predictions that seem plausible given where most warfare took place and are close to other estimates at the national level. The overall evidence suggests predicting economic activity using NTL data is a viable, low-cost alternative to conventional methods when collecting reliable information is difficult or impossible due to an armed conflict.

References

Addison, D. M. & Stewart, B. (2015), 'Nighttime lights revisited: the use of nighttime lights data as a proxy for economic variables', *World Bank Policy Research Working Paper* (7496).

URL: *https://ideas.repec.org/p/wbk/wbrwps/7496.html*

Alahmadi, M., Mansour, S., Dasgupta, N. & Martin, D. J. (2023), 'Using Nighttime Lights Data to Assess the Resumption of Religious and Socioeconomic Activities Post-COVID-19', *Remote Sensing* **15**(4), 1064.

URL: *https://www.mdpi.com/2072-4292/15/4/1064*

Bansal, C., Jain, A., Barwaria, P., Choudhary, A., Singh, A., Gupta, A. & Seth, A. (2020), Temporal prediction of socio-economic indicators using satellite imagery, *in* 'Proceedings of the 7th ACM IKDD CoDS and 25th COMAD', CoDS COMAD 2020, Association for Computing Machinery, New York, NY, USA, p. 73–81.

URL: *https://doi.org/10.1145/3371158.3371167*

Bennett, M. M. & Smith, L. C. (2017), 'Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics', *Remote Sensing of Environment* **192**, 176–197.

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

URL: *https://doi.org/10.1007/s10290-016-0246-0*

Bundervoet, T., Maiyo, L. & Sanghi, A. (2015), 'Bright Lights, Big Cities: Measuring National and Subnational Economic Growth in Africa from Outer Space, with an Application to Kenya and Rwanda', *World Bank Policy Research Working Papers* .

URL: *https://hdl.handle.net/10986/22883*

Burant, S. R. (1993), 'International relations in a regional context: Poland and its eastern neighbours—lithuania, belarus, ukraine', *Europe-Asia Studies* **45**(3), 395–418.

Chen, X. & Nordhaus, W. (2015), 'A test of the new VIIRS lights data set: Population and economic output in Africa', *Remote Sensing* **7**(4), 4937–4947.
URL: *https://www.mdpi.com/2072-4292/7/4/4937*

Chen, X. & Nordhaus, W. D. (2011), 'Using luminosity data as a proxy for economic statistics', *Proceedings of the National Academy of Sciences* **108**(21), 8589–8594.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.1017031108*

Chen, X. & Nordhaus, W. D. (2019), 'VIIRS nighttime lights in the estimation of cross-sectional and time-series GDP', *Remote Sensing* **11**(9), 1057.
URL: *https://www.mdpi.com/2072-4292/11/9/1057*

Dasgupta, N. (2022), 'Using satellite images of nighttime lights to predict the economic impact of COVID-19 in India', *Advances in Space Research* **70**(4), 863–879.
URL: *https://www.sciencedirect.com/science/article/pii/S0273117722004185*

Dodd, S. (2021), 'The Impact of Conflict on Economic Activity: Night Lights and the Bosnian Civil War'.

Elvidge, C., Baugh, K., Zhizhin, M. & Hsu, F.-C. (2013), 'Why VIIRS data are superior to DMSP for mapping nighttime lights', *Proceedings of the Asia-Pacific Advanced Network* **35**, 62–69.

Gibson, J., Olivia, S., Boe-Gibson, G. & Li, C. (2021), 'Which night lights data should we use in economics, and where?', *Journal of Development Economics* **149**, 102602.

URL: *https://www.sciencedirect.com/science/article/pii/S0304387820301772*

Gibson, J., Olivia, S. & Boe-Gibson, G. (2020), 'Night lights in economics: sources and uses', *Journal of Economic Surveys* **34**(5), 955–980.

URL: *https://onlinelibrary.wiley.com/doi/10.1111/joes.12387*

Henderson, J. V., Storeygard, A. & Weil, D. N. (2012), 'Measuring economic growth from outer space', *American Economic Review* **102**(2), 994–1028.

URL: *https://www.aeaweb.org/articles?id=10.1257/aer.102.2.994*

Huang, C., Hong, S., Niu, X., Wu, Q., Zhong, Y., Yang, H. & Zhang, H. (2023), 'Mapping of nighttime light trends and refugee population changes in Ukraine during the Russian–Ukrainian War', *Frontiers in Environmental Science* **11**, 1055100.

Ialongo, I., Bun, R., Hakkarainen, J., Virta, H. & Oda, T. (2023), 'Satellites capture socioeconomic disruptions during the 2022 full-scale war in Ukraine', *Scientific Reports* **13**(1), 14954.

IMF (2022), 'Countering the cost-of-living crisis', *World economic outlook* (2022, Oct).

Jasiński, T. (2019), 'Modeling electricity consumption using nighttime light images and artificial neural networks', *Energy* **179**, 831–842.

URL: *https://linkinghub.elsevier.com/retrieve/pii/S0360544219308618*

Jerven, M. (2014), 'Poor numbers and what to do about them', *The Lancet* **383**(9917), 594–595.

URL: *https://linkinghub.elsevier.com/retrieve/pii/S0140673614602099*

Keeley, L. H. (1997), *War before civilization*, Oxford University Press.

Khorram-Manesh, A., Burkle, F. M., Goniewicz, K. & Robinson, Y. (2021), 'Estimating the number of civilian casualties in modern armed conflicts–a systematic review', *Frontiers in Public Health* **9**, 765261.

URL: *https://www.frontiersin.org/articles/10.3389/fpubh.2021.765261/full*

Kim, D. (2022), 'Assessing regional economy in North Korea using nighttime light', *Asia and the Global Economy* **2**(3), 100046.

URL: *https://linkinghub.elsevier.com/retrieve/pii/S2667111522000238*

Kowalski, M., Zymnin, A., Sielewicz, M., Klages, K., Kowcuń, S., Bryzek, S., Lytvynenko, O., Brusnyk, S., Dabrowska, E., Koszykowska, D. & Stelmach, F. (2022), *Refugees from Ukraine – vocational activation in Poland and Germany*, EWL S.A.

Lehnert, P., Niederberger, M., Backes-Gellner, U. & Bettinger, E. (2023), 'Proxying Economic Activity with Daytime Satellite Imagery: Filling Data Gaps Across Time and Space', *PNAS Nexus* **2**(4).

URL: *https://ideas.repec.org/p/iso/educat/0165.html*

Liu, H., He, X., Bai, Y., Liu, X., Wu, Y., Zhao, Y. & Yang, H. (2021), 'Nightlight as a proxy of economic indicators: Fine-grained gdp inference around chinese mainland via attention-augmented cnn from daytime satellite imagery', *Remote Sensing* **13**(11), 2067.

Martínez, L. R. (2022), 'How much should we trust the dictator's GDP growth estimates?', *Journal of Political Economy* **130**(10), 2731–2769.

URL: *https://www.journals.uchicago.edu/doi/10.1086/720458*

McCord, G. C. & Rodriguez-Heredia, M. (2022), 'Nightlights and Subnational Economic Activity: Estimating Departmental GDP in Paraguay', *Remote Sensing* **14**(5), 1150.

URL: *https://www.mdpi.com/2072-4292/14/5/1150*

Mitrokhin, N. (2015), 'Infiltration, Instruction, Invasion: Russia's War in the Donbass'.

Musthyala, R., Kargupta, R., Jain, H. & Chakraborty, D. (2022), 'ReGNL: Rapid Prediction of GDP during Disruptive Events using Nightlights', *arXiv preprint arXiv:2201.07612* .

Otchia, C. S. & Asongu, S. (2019), Industrial growth in sub-saharan africa: Evidence from machine learning with insights from nightlight satellite images, AGDI Working Paper WP/19/046, Yaoundé.

URL: *https://hdl.handle.net/10419/205016*

Pagaduan, J. A. (2022), 'Do higher-quality nighttime lights and net primary productivity predict subnational GDP in developing countries? Evidence from the Philippines', *Asian Economic Journal* **36**(3), 288–317.

URL: *https://onlinelibrary.wiley.com/doi/10.1111/asej.12278*

Pearlman, W. (2020), 'Host state engagement, socioeconomic class, and syrian refugees in turkey and germany', *Comparative Politics* **52**(2), 241–272.

Pérez-Sindín, X. S., Chen, T.-H. K. & Prishchepov, A. V. (2021), 'Are night-time lights a good proxy of economic activity in rural areas in middle and low-income countries? Examining the empirical evidence from Colombia', *Remote Sensing Applications: Society and Environment* **24**, 100647.

URL: *https://linkinghub.elsevier.com/retrieve/pii/S235293852100183X*

Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C., Pandey, B., Seto, K. C., Hall, D., Oda, T., Wolfe, R. E., Lin, G., Golpayegani, N., Devadiga, S., Davidson, C., Sarkar, S., Praderas, C., Schmaltz, J., Boller, R., Stevens, J., Ramos González, O. M., Padilla, E., Alonso, J., Detrés, Y., Armstrong, R., Miranda, I., Conte, Y., Marrero, N., MacManus, K., Esch, T. & Masuoka, E. J. (2018), 'NASA's Black Marble nighttime lights product suite', *Remote Sensing of Environment* **210**, 113–143.
URL: *https://www.sciencedirect.com/science/article/pii/S003442571830110X*

Tahsin, M. (2022), 'Assessing the reliability of macro data using night-time lights models: Bangladesh'.
URL: *https://www.researchsquare.com/article/rs-2390237/latest*

UNHCR (2024), 'Operational Data Portal: Ukraine Refugee Situation'.
URL: *https://data.unhcr.org/en/situations/ukraine*

UNHCR & REACH (2022), 'Uchodźcy z Ukraine w Polsce: Profilowanie Czerwiec 2022'.
URL: *https://data.unhcr.org/en/documents/details/96508*

Wang, L., Lei, H. & Xu, H. (2024), 'Analysis of Nighttime Light Changes and Trends in the One Year Anniversary of the Russia-Ukraine Conflict', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* .

Wang, X., Rafa, M., Moyer, J. D., Li, J., Scheer, J. & Sutton, P. (2019), 'Estimation and Mapping of Sub-National GDP in Uganda Using NPP-VIIRS Imagery', *Remote Sensing* **11**(2), 163.
URL: *https://www.mdpi.com/2072-4292/11/2/163*

Wijesekera, N. K. (2023), 'Illuminating the Economic Costs of Conflict: A Night Light Analysis of the Sri Lankan Civil War'.

Wojdat, M. & Cywiński, P. (2024), 'Urban Hospitality: Unprecedented Growth, Challenges and Opportunities'.

URL: *https://metropolie.pl/fileadmin/news/2022/10/*$UMP_{r}aport_{U}kraina_{A}NG_{2}0220429_{f}inal.pdf$

World Bank (2024), 'Gdp growth', *World Bank national accounts data* .

URL: *https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=UA*

Zhao, N., Liu, Y., Cao, G., Samson, E. L. & Zhang, J. (2017), 'Forecasting china's GDP at the pixel level using nighttime lights time series and population images', *GIScience & Remote Sensing* **54**(3), 407–425.

URL: *https://doi.org/10.1080/15481603.2016.1276705*