Bolivar, Osmar

Article

GDP nowcasting: A machine learning and remote sensing data-based approach for Bolivia

Latin American Journal of Central Banking (LAJCB)

Provided in Cooperation with:
Center for Latin American Monetary Studies, Mexico City

Suggested Citation: Bolivar, Osmar (2024) : GDP nowcasting: A machine learning and remote sensing data-based approach for Bolivia, Latin American Journal of Central Banking (LAJCB), ISSN 2666-1438, Elsevier, Amsterdam, Vol. 5, Iss. 3, pp. 1-25, https://doi.org/10.1016/j.latcb.2024.100126

This Version is available at:
https://hdl.handle.net/10419/336403

Latin American Journal of Central Banking 5 (2024) 100126

Contents lists available at ScienceDirect

Latin American Journal of Central Banking

journal homepage: www.elsevier.com/locate/latcb

GDP nowcasting: A machine learning and remote sensing data-based approach for Bolivia

Osmar Bolivar[1]

Ministerio de Economía y Finanzas Públicas, Bolivia

ARTICLE INFO

JEL classification:
C10
C14
C22
C53
C80
E17

Keywords:
Nowcasting
Machine learning
Remote sensing
Economic growth forecast

ABSTRACT

This research introduces an innovative GDP nowcasting strategy tailored for developing countries, specifically addressing challenges related to limited data timeliness. The study centers on Bolivia, where the official monthly indicator of economic growth is released with a substantial delay of up to six months. The proposed nowcast estimates effectively narrow this gap from six to two months. This advancement is achieved through the integration of machine learning techniques with data comprising indicators from traditional sources and statistics derived from satellite imagery. The robustness of this approach is rigorously validated using various criteria, including performance comparisons with conventional econometric methods and sensitivity assessments to different feature sets. Beyond enhancing the understanding of Bolivia's economic dynamics, this research establishes a framework for analogous analyses in regions grappling with information availability challenges.

1. Introduction

In developing countries, the challenge of accessing timely and accurate economic data significantly hampers informed policymaking. A critical aspect of this challenge lies in the absence of gross domestic product (GDP) indicators with high temporal and spatial frequency. The scarcity and infrequency of GDP data impede the timely assessment of economic performance, limiting policymakers' ability to formulate effective strategies and respond promptly to emerging trends. This deficiency becomes particularly pronounced in an era marked by rapid technological advancements.

In the Latin American region, crucial data related to GDP or economic activity indicators is often released with substantial delays and subsequent revisions. For example, countries like Bolivia publish official monthly indicators of economic growth with a delay of up to six months from contemporary time.

To address this predicament, literature provides methodological strategies for nowcasting high-frequency GDP data to understand how economic activity evolves closely to contemporary time. Nowcasting, formally defined as the prediction of the present, the very near future, and the very recent past, involves exploiting information published early and possibly at higher frequencies than the target variable of interest, such as GDP (Giannone et al., 2008).

Giannone et al. (2008) laid the foundation for GDP nowcasting by developing a formal and internally consistent statistical framework. Their method evaluates the marginal impact of intra-monthly data releases on current-quarter forecasts (nowcasts) of real GDP growth. Subsequent studies adopted similar approaches (Bańbura et al., 2013; Bok et al., 2018).

While many studies focus on nowcasting quarterly GDP using econometric and statistical methods, there is scarce literature addressing nowcasting through artificial intelligence – machine learning – and big data. Machine learning algorithms have proven

E-mail address: osmar.economics@gmail.com.

[1] The author of this paper expresses his own opinions and perspectives, which may not necessarily coincide with those of the institutions to which he is associated.

https://doi.org/10.1016/j.latcb.2024.100126
Received 14 November 2023; Received in revised form 5 March 2024; Accepted 11 March 2024
Available online 30 March 2024

efficient in processing voluminous economic data, enabling prompt trend and pattern discernment (Mullainathan and Spiess, 2017). Remarkably, they have been employed to forecast financial crises, GDP growth volatility, unemployment, inflation, demand, sovereign risk, exchange rates, agricultural commodity prices, house prices, entrepreneurial firm valuation, among others.[2]

Another untapped source of big data for nowcasting is remote-sensing information, a process of detecting and monitoring physical characteristics using satellite or aircraft measurements. For instance, Henderson et al. (2012) pioneered using satellite data on night lights to estimate GDP growth, enabling growth measurement at sub and supranational levels. Subsequent studies have proposed and described how remote sensing methods can be used to process satellite imagery data to generate metrics about night lights, climate and weather, topography, agricultural and urban land use, building types, natural resources, and pollution monitoring, as appropriate predictors of economic activity.[3]

Addressing these gaps, this research seeks to introduce a novel methodology for nowcasting monthly economic growth, utilizing machine learning algorithms and incorporating inputs from both conventional and remote-sensing data sources. I aim to assess and validate this approach through a case study of Bolivia.

The choice of Bolivia as the focal point of this study arises from its transformative journey since 2005, characterized by significant growth fluctuations and economic challenges. While experiencing commendable growth between 2006 and 2013, the economy encountered obstacles post-2014, driven by low commodity prices, the impact of climate change, and social conflicts. The COVID-19 pandemic further exacerbated the economic downturn. Throughout this period, a persistent challenge has been the acquisition of timely and reliable economic data.

This historical context lays the foundation for the research imperative, underscoring the necessity for innovative methodologies to nowcast GDP and refine economic forecasting in Bolivia. Notably, Bolivia's official monthly indicator of economic growth, the Global Index of Economic Activity, is released with a delay of several months from the contemporary time. Consequently, this study endeavors to formulate a nowcast metric for monthly economic growth, narrowing the gap between real-time estimates and official reports from six months to just two months. This advancement aims to furnish more timely metrics, thereby empowering policy-making decisions, aiding researchers, and supporting economic analysts in their analyses.

The paper comprises four major sections: the current introduction, the methodology detailing the target variable, predictors from conventional and remote-sensing data, and the implementation of machine learning algorithms; results covering hyperparameter tuning, feature importance, and findings on Bolivian monthly economic activity nowcasting, along with sensitivity analysis; and concluding remarks.

2. Methodology

2.1. Target variable

This study aims to devise a methodology for nowcasting monthly economic growth, employing machine learning algorithms and incorporating inputs from both conventional and remote-sensing data sources, with a focus on testing this approach for Bolivia's monthly economic growth.

In Bolivia, the absence of a monthly GDP indicator is mitigated by the availability of the Global Index of Economic Activity (IGAE), officially published by the National Institute of Statistics. The IGAE, a monthly synthetic index derived from the aggregation of sectoral production indices, serves as the primary metric for gauging the country's economic activity.[4]

Despite its original conceptualization as a prompt metric of economic activity, it is pertinent to note that the release of IGAE data is typically delayed by up to six months from the present period. For instance, as of this research completion date on April 20, 2023, the available IGAE data only extended up to September 2022. This temporal lag creates a six-month gap without a metric for Bolivia's aggregated economic activity.

Therefore, the target variable under consideration for nowcasting in this study is the monthly growth rate of the IGAE. This choice is preferred due to the superior prediction performance observed compared to alternative transformations such as year-over-year (y-o-y) and cumulative growth rates or the log of the IGAE.[5]

2.2. Features

The initial phase of the methodology for nowcasting the monthly growth rate of the IGAE involves assembling a collection of variables exhibiting robust correlation with the target variable, serving as viable predictors (features).

A contribution of this research is the construction of the set of features for the target variable, combining indicators from conventional sources of information with remote-sensing data derived from satellite imagery.

---

[2] See Liu et al. (2022); Balcilar et al. (2022); Gogas et al. (2022); Medeiros et al. (2021); Bajari et al. (2015); Belly et al. (2023); Plakandaras et al. (2015); Bonato et al. (2022); Milunovich (2020); and Zhang et al. (2023).

[3] See Keola et al. (2015); Donaldson and Storeygard (2016), Pinkovskiy and Sala-i Martin (2016); and Storeygard (2016).

[4] Further information about IGAE: https://www.ine.gob.bo/index.php/estadisticas-economicas/indice-global-de-actividad-economica-igae/.

[5] While training several machine learning algorithms using the same features described in Section 2.2, attempts were made to nowcast y-o-y and cumulative growth rates, as well as the log of the IGAE. However, the results indicated higher mean squared errors and lower R^2 scores, leading to the exclusion of these outcomes from the paper for conciseness.

2.2.1. Features from conventional sources

Producing a nowcast of Bolivia's monthly economic growth requires predictors with a monthly frequency. The chosen predictors or features (X) must encapsulate the fundamental behaviors driving the Bolivian economy, incorporating metrics from key sectors (hydrocarbons, mining, etc.) and variables from fiscal, monetary, financial systems, and trade indicators. Additionally, consideration is given to other indicators exhibiting a statistically significant correlation with the IGAE.

A comprehensive process was initiated to gather features meeting the specified attributes. Initially, 261 variables were collected as potential predictors; see Appendix A for a broad description of these variables.

However, literature on machine learning prediction suggests that feature reduction by filtering variables proving a positive and relatively strong correlation with the target variable enhance forecast performance (Kotsiantis, 2011). Therefore, a correlation analysis was conducted to examine the relationship between the IGAE and the 261 variables collected, both for their contemporaneous values and lags ranging from the first to the twelfth, on a monthly growth rate basis.

The analysis revealed that contemporaneously, electricity generation and its specific consumption by small and large industries, the production of liquefied petroleum gas and tin, rail transport, and construction permits displayed a higher correlation with the IGAE. From an economic perspective, the selection of these variables aligns with the significance of sectors such as hydrocarbons, mining, industry, transport, and construction in the Bolivian economy.

Furthermore, several variables in foreign trade, such as mineral exports (total, zinc, and borates) and agricultural exports (total and chia), unveiled a strong and positive correlation with the target variable. Increases in the imports of durable consumer goods, raw materials, and intermediate products for agriculture and industry were also associated with monthly and contemporaneous increases in the IGAE. Imports of clothing, chemicals and chemical products, manufactured goods, rubber and plastic products, non-metallic mineral products, and food and beverages – classified according to the International Standard Industrial Classification –, showed a high and positive correlation with the monthly variation of the IGAE.

Moreover, certain lags of international commodity prices were found to have a positive relationship with the IGAE. For instance, the first lag of oil prices revealed a positive correlation, consistent with the importance of natural gas – its price is indexed to the price of oil – in Bolivia's exports. Gold prices showed a positive degree of association with the IGAE when considered with two lags.

To capture seasonality, lags 1, 2, 3, 6, 9, and 12 of the IGAE monthly growth, along with temporal signaling variables using sine and cosine functions, are included among predictor variables.[6]

Appendix B presents the final selection of 35 features from conventional sources, accompanied by their correlation with the target variable.

2.2.2. Features from remote sensing

Remote sensing, which utilizes sensors to acquire data about the Earth's surface and atmosphere from a distance, has proven indispensable across diverse fields, encompassing agriculture, forestry, geology, and environmental studies. Its application in generating socioeconomic indicators by analyzing satellite information yields detailed insights into the distribution and characteristics of human settlements, infrastructure, and economic activities. This approach, validated in monitoring urbanization, population growth, and economic development, furnishes data for well-informed decision-making and the formulation of sustainable strategies.

For this research, statistics derived from remote-sensing data for Bolivia's extension, including nighttime lights, built-up areas, and vegetation, as well as precipitation for major crops, are explored as predictors for nowcasting the monthly growth rate of the IGAE.

As an illustration, the correlation of certain remote-sensing-based features with the target variable is presented in Appendix B.

Nighttime Lights

Considerable research has established a robust correlation between nighttime lights and Gross Domestic Product (GDP) measures at national, state, and regional levels (Ghosh et al., 2010; Chen and Nordhaus, 2011; Henderson et al., 2012). Nighttime light observations serve as a reliable proxy for economic activity, especially in data-scarce areas or where statistical systems are of low quality. Changes in luminosity offer an additional measure of income growth when other indicators are unavailable.[7]

To this end, one of the remote-sensing predictors is constructed based on nighttime lights captured by satellite images for Bolivia. Daily metrics of visible and near-infrared nocturnal lights are utilized, produced by NASA's Land Processes Distributed Active Archive Center (LP DAAC).[8]

To execute the nowcast of the target variable, inputs with a monthly frequency are required. Therefore, a time series is assembled, comprising the monthly median values of brightness in pixels located within the territory of Bolivia. This variable records its initial observation in January 2012 when nighttime VIIRS images became available.

Furthermore, to assess the suitability of regional time series of nighttime lights as features and as part of a sensitivity analysis, forecast algorithms are tested using variables of the monthly median luminosity within the three major metropolitan regions of

[6] Feature importance and forecasts demonstrate the effective capture of seasonal patterns.

[7] The terms "nighttime lights", "light brightness", and "luminosity" are used interchangeably.

[8] Specifically, the image collection is the VIIRS Lunar Gap-Filled BRDF Nighttime Lights Daily L3 Global 500 m. This collection encompasses the band *"DNB_BRDF_Corrected_NTL"*, which is an enhanced version of the average radiance band from the conventionally used Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) (Román et al., 2018). In contrast to the US Defense Meteorological Satellite Program's (DMSP) night lights, this product does not exhibit blurring or overglow effects, nor does it suffer from oversaturation.

Bolivia: La Paz, Kanata, and Santa Cruz. These regions are recognized as the primary population and economic centers of Bolivia. Refer to Fig. C.1 in Appendix C for an illustration of the nighttime lights in these regions.[9]

Vegetation Indices

Vegetation indices derived from satellite imagery are intricately linked to agricultural production, thereby implying a useful proxy of aggregated economic activity (Wiegand et al., 1991; Hu and Xia, 2019).[10] The pivotal role of agricultural production in the Bolivian economy, contributing both to food security and income, is evident; any increase in agricultural output typically correlates with a rise in overall economic activity.[11]

Machine learning models exhibit efficacy in utilizing information derived from vegetation indices for accurate predictions of agricultural output and, consequently, economic activity (Jung et al., 2021; Sharifi, 2021; Tang et al., 2022). In this study, the Normalized Difference Vegetation Index (NDVI) imagery serves as input for generating predictors of economic growth in Bolivia. The NDVI images are sourced from the Suomi National Polar-Orbiting Partnership (Suomi NPP) NASA Visible Infrared Imaging Radiometer Suite (VIIRS) Vegetation Indices (VNP13A1) Version 1 data.

For each of the primary geographical zones encompassing Bolivia's key crop lands, including rice, sugar cane, soybean, and wheat, time series are computed with monthly average values of the NDVI.[12] While not all crop areas in Bolivia are considered in constructing variables with the average value of the NDVI, the inclusion focuses on the most representative ones in terms of large-scale production with export destinations.[13]

For a visual representation of how NDVI imagery maps soybean cultivation areas in the department of Santa Cruz, Bolivia, please refer to Fig. C.2 in Appendix C.

Built-up Areas

The identification and mapping of built-up areas involve assessing regions with high-density human development, including buildings and infrastructure. These areas serve as proxies for economic activity, given their correlation with high population density and economic productivity (Zhou et al., 2021; Liu et al., 2021).

The Global Human Settlement Layer (GHSL), a widely recognized mapping tool developed by the European Commission, provides global coverage of built-up areas. This high-resolution dataset, generated using satellite imagery, census data, and other sources, offers detailed information on the spatial distribution of human settlements worldwide (Melchiorri et al., 2018). Unfortunately, the GHSL's last update is available for 2015, posing a challenge for researchers and policymakers in need of more frequent updates on urban development. To address this limitation, machine learning algorithms, such as those proposed by Kussul et al. (2017) and Corbane et al. (2021), can be employed to train models for accurate monthly predictions of built-up areas.

To introduce new inputs for nowcasting Bolivia's monthly overall economic activity, a Random Forest (RF) algorithm was trained using GHSL images to classify 1 km pixels into built-up or non-built-up areas. The algorithm utilizes data from the Landsat −8 satellite's blue, green, red, near-infrared, and shortwave infrared 1 and 2 bands, along with light brightness levels.[14] The RF model is chosen for its effectiveness and widespread application across various domains and problems (Duro et al., 2012); for further details about the trained RF model, please allude to Appendix D.[15]

Subsequently, the outputs of the built-up image classification are employed to generate monthly time series depicting changes in the proportion of built-up pixels over total pixels within the following regions of interest: Bolivia, Metropolitan Area of La Paz, Metropolitan Area of Kanata, and Metropolitan Area of Santa Cruz.

Precipitation

Precipitation, the amount of water that falls from the atmosphere to the ground in the form of rain, snow, or hail, plays a relevant role in the economic growth of regions worldwide. It is indispensable for agricultural production, a significant sector particularly in developing countries like Bolivia. Rainfall also impacts industrial processes, hydroelectric power generation, and transportation infrastructure.

Using sensors and instruments mounted on satellites, remote sensing measures various physical properties of the Earth's surface, including temperature, reflectance, and moisture content. These measurements are utilized to calculate rainfall, primary beneficial in regions with limited or nonexistent traditional meteorological data.

Hence, an additional metric involving processed satellite imagery is a time series featuring the median precipitation within Bolivia's rice, sugar cane, wheat, and soybean cultivation regions — the same regions used in NDVI computation. These variables are obtained from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), providing precipitation raster files with 0.05° resolution pixels.

[9] The Metropolitan Area of the City of La Paz refers to the regional unit composed of the municipalities of Nuestra Señora de La Paz, El Alto, Viacha, Achocalla, Palca, Laja, Pucarani, and Mecapaca in the highlands of Bolivia; these eight municipalities have a population of nearly 2 million inhabitants. The Kanata Metropolitan Region is composed of the municipalities of Cochabamba, Colcapirhua, Quillacollo, Sacaba, Sipe Sipe, Tiquipaya, and Vinto, totaling seven con-urbanized municipalities where approximately 1.5 million inhabitants coexist. The Metropolitan Area of Santa Cruz de la Sierra, currently the most populated urban center in the country (2.4 million inhabitants), is a con-urbanization of six municipalities in the department of Santa Cruz: Santa Cruz de la Sierra, La Guardia, Warnes, Cotoca, El Torno, and Porongo.

[10] Vegetation indices are mathematical formulations quantifying the status and health of vegetation using satellite data.

[11] Between 1990 and 2021, the Agriculture sector represented an average of 13% of Bolivia's nominal GDP, according to data from the National Institute of Statistics.

[12] Results are comparable to those obtained using the median NDVI, but average values display a higher correlation with the IGAE.

[13] Besides, vector files providing georeferencing for these crop zones are publicly available.

[14] For instance, Jiang et al. (2020) employed nighttime lights to detect urban growth dynamics in Africa.

[15] For example, visit: https://osmarbolivar.github.io/R_built_up.html to download Bolivia's 2022 Built-Up Classification.

2.3. Training, validation and test sets

The feature matrix X comprises a total of 44 variables, integrating time series data from both conventional sources and those derived from remote sensing techniques. This matrix serves as input for training and validating machine learning algorithms with the objective of nowcasting the monthly growth rate of the IGAE (see Appendix B for the list of these variables).

While most variables from conventional sources have data available since 2010 or earlier, information for variables generated based on satellite imagery is accessible from April 2013 onward, including lag transformations. Thus, the complete set of features spans from April 2013 to January 2023.

A subset is designated as the training set, covering the period from April 2013 to December 2018. The validation set incorporates data from January 2019 to September 2022. Notably, the choice of this period for the validation set aims to assess the forecast performance during the months of the COVID-19 pandemic, marked by atypical behavior, thereby evaluating the response of trained algorithms under such conditions.

As the IGAE data is publicly available only until September 2022, the testing set corresponds to the period from October 2022 to January 2023.[16] While forecasting the IGAE beyond January 2023 would be ideal, most selected features lack information for those months. It is crucial to emphasize that the nowcast proposed in this research is a conditional forecast; thus, the forecast horizon is contingent upon the availability of information.

Finally, to ensure comparability and mitigate bias, all variables undergo z-score normalization. Specifically, the features are adjusted according to the provided formula:

$$x_j^{(i)} = \frac{x_j^{(i)} - \mu_j}{\sigma_j} \tag{1}$$

Here, j denotes the selected variable or column in the feature matrix (X), while μ_j and σ_j represent the mean and standard deviation, respectively, of the training set values for the selected variable j. The same normalization procedure is applied to y (i.e., IGAE growth rate).

2.4. Machine learning algorithms

After establishing training, validation, and testing sets, the subsequent step involves training the algorithms for predicting the monthly growth rate of the IGAE. This section provides an overview of the selected algorithms used to attain this objective. Detailed information about these algorithms, including their methodological characteristics and execution, can be found in Appendix E.

It should be noted that the selection criteria for these algorithms are based on their common usage in the machine learning literature for time series prediction. Additionally, some specialize in modeling linear relationships, while others excel in identifying non-linear patterns.

Hyperparameter tuning is vital for optimizing machine learning models, involving the selection of values for hyperparameters to control algorithm behavior. Accordingly, the selected algorithms undergo a rigorous tuning process, primarily using a time-aware cross-validation method to prevent information leakage from the future.

The following outlines the main characteristics and hyperparameter tuning process for each algorithm:

1. **Ridge:** A linear regression technique incorporating regularization through an L2 penalty. The regularization hyperparameter, denoted as λ, controls the trade-off between model fit and complexity, effectively preventing overfitting by reducing coefficients' magnitudes towards zero (Hoerl and Kennard, 1970). Tuning involves a 5-fold cross-validation process with 1000 λ values, logarithmically spaced from 10^{-5} to 10^{5}, selecting the optimal value minimizing Mean Square Error (MSE).
2. **Lasso:** A linear regression technique utilizing an L1 penalty to encourage sparsity and feature selection (some coefficients are zero), particularly advantageous when dealing with numerous predictors (Tibshirani, 1996). The hyperparameter tuning method mirrors Ridge regression, involving a 5-fold cross-validation over 1000 λ values from 10^{-5} to 10^{5}, selecting the optimal λ minimizing MSE.
3. **ElasticNet (EN):** It combines Ridge and Lasso, utilizing both L1 and L2 penalties to balance sparsity and coefficient shrinkage (Zou and Hastie, 2005). Hyperparameters λ and α are optimized through a 5-fold cross-validation, considering 1000 values for both λ and α, ranging from 0.05 to 0.95 in 0.01 increments, selecting values minimizing MSE.
4. **Decision Tree Regressor (DT):** A non-parametric algorithm that recursively partitions the predictor space, beneficial for modeling complex, non-linear relationships (Loh, 2011). Hyperparameters, including maximum depth (d), minimum samples to split an internal node (mss), and minimum samples at a leaf node (msl), are optimized through a 5-fold cross-validation, over a range of values based on prior research and machine learning best practices (Mantovani et al., 2018).
5. **AdaBoost Regressor (ADA):** It combines weak learners to form a robust predictor, emphasizing misclassified data points to improve performance in noisy or complex scenarios (Freund and Schapire, 1997). Hyperparameters encompass a variety of factors, including the maximum depth of the tree (d), the maximum number of estimators utilized during boosting (T), the learning rate (α_t), and the loss function employed to update weights after each boosting iteration. The selection of hyperparameter ranges for tuning is contingent upon several considerations, such as model complexity and available training

[16] IGAE time series are available monthly since January 1990.

resources. To determine the optimal value for α_t, a broad range of values is typically assessed to strike a balance between convergence speed and overfitting risk; a starting range of 0.01 to 3 with an increment of 0.01 are used for tuning. For T the range is chosen based on the trade-off between model complexity and performance, with a large number of estimators potentially leading to overfitting and a small number resulting in underfitting; thus, a range of 50 to 200 with an increment of 1 are tested. Additionally, linear, square, and exponential loss functions are evaluated. d is assess in a range from 3 to 10. Optimal values are selected based on minimizing the MSE through 5-fold cross-validation.

6. **Gradient Boosting Regressor (GBR):** It constructs an additive model of weak learners, modifying residual errors in a step-by-step approach, effective in complex or noisy data (Friedman, 2001). The hyperparameters employed are the learning rate (γ_m), maximum depth of the tree (d), the maximum number of estimators at which boosting is terminated (T), and the minimum number of samples required to split an internal node (mss). To determine the optimal values for each of these hyperparameters, 5-fold cross-validation is performed over a range of values. For T hyperparameter, a range of 100 to 500 with a step of 5 is chosen; this range is deemed reasonable as it encompasses a wide range of values and includes some values that have been known to work well in practice. Similarly, for the d hyperparameter, a range of 3 to 9 with a step of 1 is chosen, because it covers a range of depths that have been known to work well in practice, while exceeding a depth of 9 can result in overfitting. In the context of the mss hyperparameter, it is worth noting that the range of admissible values spans from 2 to 20, with a step size of 1; a choice of 2 would result in a greater number of splits and a heightened risk of overfitting, whereas a higher value of 20 would lead to fewer splits and a potential for underfitting. For the learning rate hyperparameter, a range of 0.01 to 0.1 with a step of 0.01 is chosen. Besides, Stochastic Gradient Boosting is tested because it may help reduce overfitting by introducing randomness into the training process; the parameter in sklearn library to assess this option is $subsample$. The optimal values are selected based on the ones that minimize the MSE.
7. **Random Forest Regression (RF):** It amalgamates multiple decision trees for enhanced accuracy and robustness, especially suitable for intricate, non-linear relationships (Breiman, 2001). The hyperparameters for tuning are selected based on prior knowledge and common practice in the literature. The number of estimators (T) is tested in the range of 100 to 300 with increments of 5, as a larger number of trees can improve the model's performance, but at a certain point, the improvement may be negligible. The criterion for splitting nodes ($criterion$) is set to test both the mean squared error and the mean absolute error, as both are common loss functions in regression problems. The minimum number of samples required to split an internal node (mss) is evaluated in the range of 2 to 10 with increments of 1; this range is set to avoid overfitting, as having too few samples required to split a node may lead to over-complex trees, while having too many may lead to underfitting. A 5-fold cross-validation process is conducted to identify the optimal hyperparameters that minimize the MSE.
8. **Extra Trees Regressor (ET):** It is an ensemble technique using random subsets for decision tree construction, introducing additional randomness to mitigate overfitting (Geurts et al., 2006). For tuning hyperparameters a wide range of values were analyzed to identify the combination that yields the best performance. For instance, a range of values from 100 to 500, with a step size of 5 is evaluated for the number of trees in the forest (T). Besides tuning other common parameters such as d, mss, and msl, the algorithm is implemented allowing the option to whether or not bootstrap samples are used when building trees. If bootstrap is enable, there is a parameter to be tuned which is the number of samples to draw from X to train each base estimator ($MaxSample$); a range of values from 0.1 to 1.0, with a step size of 0.01, are tested. The final hyperparameters are chosen based on the ones that minimize the MSE over a 5-fold cross-validation strategy.

2.5. Nowcasting metrics

Following the training and validation of the algorithms, eight distinct z-score normalized predicted values of the target variable can be computed. However, the performance of these algorithms varies, and differences in variability are observed. Thus, it is important to establish criteria that yield aggregated metrics for nowcasting the monthly growth rate of the IGAE. To achieve this objective, the following procedure is proposed:

- Each z-score normalized forecast of the target variable ($\hat{y}_i^z$; $\forall\ i = 1, \dots, 8$) is converted to its original units, representing percentage points of the monthly growth rate of the IGAE ($\hat{y}_i$; $\forall\ i = 1, \dots, 8$).
- Using the predicted monthly growth rates, IGAE levels are calculated for each algorithm ($\hat{IGAE}_i$; $\forall\ i = 1, \dots, 8$). This step is necessary for computing statistics such as the geometric mean, given the caveat that geometric means cannot be estimated with negative values, which is the case when working with growth rates.
- In line with a well-established strategy supported by existing literature (Wolpert, 1992; Breiman, 1996; Dietterich, 2000), aimed at enhancing the quality and reliability of predictions in regression problems by combining individual forecasts through averaging, the following statistics are constructed:
  - The arithmetic mean for all predicted IGAE levels ($\hat{IGAE}_{\text{AM}}$).
  - The weighted arithmetic mean for all predicted IGAE levels, with weights defined as the inverse of the mean squared errors of each algorithm ($\hat{IGAE}_{\text{WAM}}$).
  - The weighted arithmetic mean for predicted IGAE levels from the three best-performing algorithms ($\hat{IGAE}_{\text{WAM-BEST}}$); weights are set as the inverse of the mean squared errors of each algorithm.
  - The weighted geometric mean for all predicted IGAE levels, with weights defined as the inverse of the mean squared errors of each algorithm ($\hat{IGAE}_{\text{WGM}}$).

– The weighted geometric mean for predicted IGAE levels from the three best-performing algorithms ($\widehat{IGAE}_{\text{WGM-BEST}}$); weights are set as the inverse of the mean squared errors of each algorithm.

- Using the variables generated in the previous step, the following predicted monthly growth rates are calculated: $\hat{y}_{\text{AM}}$, $\hat{y}_{\text{WAM}}$, $\hat{y}_{\text{WAM-BEST}}$, $\hat{y}_{\text{WGM}}$, $\hat{y}_{\text{WGM-BEST}}$. Ultimately, these predictions constitute the aggregated nowcasting metrics.

Additionally, a nowcast interval is established to delineate the range of values for all predicted monthly growth rates of the IGAE. In essence, the upper and lower limits of the interval are determined by the maximum ($\hat{y}_{\text{max}}$) and minimum ($\hat{y}_{\text{min}}$) values of the forecasts generated by the eight algorithms.

3. Results

3.1. Fine-tuning hyperparameters

Table 1 demonstrates the performance of selected machine learning algorithms in predicting the z-score normalized monthly growth rate of the IGAE.

Table 1
Fine-Tuned Hyperparameters by Algorithm.

Algorithm	Without tuning			After tuning (5-fold cross-validation)			
	MSE	RMSE	MAE	MSE	RMSE	MAE	Hyperparameters
Ridge	0.317	0.563	0.350	0.297	0.545	0.357	$\lambda = 14.508$
Lasso	1.687	1.299	1.034	0.421	0.649	0.378	$\lambda = 0.011$
ElasticNet	0.987	0.993	0.733	0.350	0.592	0.377	$\lambda = 0.087$ $\alpha = 0.11$
Decision Tree	0.721	0.849	0.453	0.712	0.844	0.420	$d = 4$
AdaBoost	0.557	0.746	0.371	0.524	0.724	0.332	$d = 5$ $\alpha_t = 2.1$ $Loss = Exponential$ $T = 60$
Gradient Boosting	0.505	0.711	0.304	0.390	0.625	0.292	$d = 7$ $T = 135$ $mss = 10$ $Subsample = 0.4$
Random Forest	0.560	0.749	0.368	0.531	0.729	0.359	$mss = 3$ $T = 215$
Extra Trees	0.501	0.707	0.311	0.496	0.704	0.356	$d = 13$ $MaxSamples = 0.98$

Note: In the process of training and evaluating the algorithms, the scikit-learn library was employed. It should be noted that in comparison to Section 2.4, certain hyperparameters are not included in the table above. However, during the fine-tuning process, it was determined that the default values in the scikit-learn functions were the most suitable; hence, these default values have been omitted from the table.

The first column lists each algorithm, while the second to fourth columns show the mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) without hyperparameter tuning. The fifth to eighth columns exhibit the same metrics after hyperparameter tuning using a 5-fold cross-validation strategy. MSE and RMSE are common metrics for regression problems, with larger errors penalized more severely. MAE treats all errors equally, regardless of their magnitude.

Upon examination, it is evident that hyperparameter tuning improved the performance of all algorithms; thereby, prevented overfitting and increased the models' generalization ability. This indicates that the initial hyperparameters were suboptimal, and fine-tuning enhanced model performance.[17]

Prior to tuning, Lasso and ElasticNet exhibited the highest MSE and RMSE values, while Ridge regression showed superior performance. After fine-tuning, Ridge regression outperformed other algorithms in predicting the target variable in the validation set, displaying reduced MSE and RMSE values. ElasticNet also showed a substantial reduction in MSE and RMSE, primarily due to its L2 regularization similar to Ridge regression.

Of note, Gradient Boosting achieved the lowest MAE and relatively low MSE and RMSE values among all algorithms, indicating its potential suitability for IGAE nowcasting.

Ultimately, Ridge regression emerged as the best-performing algorithm, both before and after hyperparameter tuning. This highlights Ridge's potential as the most suitable algorithm for nowcasting Bolivia's monthly economic growth, given the dataset used in this study. The superior performance of Ridge could be attributed to robust linear associations between the target variable and its predictors. Importantly, Ridge regression, unlike Lasso, retains all predictors when making predictions, a factor contributing to its enhanced performance.

However, when aiming to capture intricate patterns in non-linear relationships, models such as Gradient Boosting deserve careful consideration for nowcasting the IGAE in the Bolivian context.

[17] The hyperparameters in non-tuned models represent default values in the scikit-learn library.

3.2. Feature importance

While having identified algorithms with higher predictive capabilities for IGAE growth, understanding the significance of individual predictors (features) is crucial. Table 2 illustrates the weights assigned to predictors in each algorithm, scaled for visibility by multiplying the original values by 100.

Table 2
Feature importance by algorithm.

Feature	Ridge	Lasso	EN	DT	ADA	GBR	RF	ET
IGAE (lag 12)	26.70	63.50	36.60	65.30	46.30	43.00	60.00	39.90
Bolivia: Median Nighttime Lights	22.70	23.60	26.30	0.00	29.20	25.90	13.90	27.30
IGAE (lag 6)	17.80	5.10	14.40	0.00	0.40	0.60	0.10	3.50
Sine monthly-time signaling variable	10.00	1.90	8.30	0.00	0.50	1.10	0.90	4.10
Small industries: Electricity consumption	8.90	0.90	7.60	0.00	0.70	1.80	1.10	1.80
IGAE (lag 9)	8.10	4.30	6.00	0.00	0.50	1.80	0.40	0.50
Large industries: Electricity consumption	7.30	5.40	6.60	0.00	0.10	0.10	0.10	0.20
Rail transport consumption index	7.10	0.10	5.10	0.00	0.40	1.00	0.00	0.40
Santa Cruz: Average NDVI in sugar cane crops (lag 6)	6.50	2.40	5.80	0.00	0.10	0.30	0.20	0.50
Imports: Manufacture of clothing	5.20	1.30	4.50	0.00	1.70	0.90	4.30	1.20
Imports: Manufacture of rubber and plastic products	5.20	10.10	4.90	0.00	0.10	0.10	0.10	0.30
Imports: Durable Consumer Goods	4.10	0.00	1.10	0.70	0.20	0.20	0.20	0.20
Production: Tin	4.00	3.20	3.00	0.70	0.40	0.70	0.20	0.10
Generation of electrical power	4.00	0.00	2.40	12.60	1.30	5.00	4.70	2.10
Commodity price: Gold (lag 2)	3.40	0.00	1.70	0.00	0.20	0.20	0.10	0.10
Santa Cruz: Median NDVI in wheat crops (lag 9)	3.20	0.00	0.30	0.00	0.20	0.30	0.20	0.40
Bolivia: Share of built-up pixels	2.50	0.00	0.40	5.10	1.00	0.20	1.40	0.30
Exports: Wood and wood products	2.40	2.30	2.80	0.00	0.40	1.90	0.30	0.20
Exports: Agriculture, livestock, forestry and fishing	2.30	2.60	2.00	0.00	0.10	0.80	0.10	0.30
Cosine yearly-time signaling variable	1.90	0.00	0.20	0.00	0.20	0.10	0.10	0.20
Production: Liquefied petroleum gas	1.90	0.00	0.00	0.00	0.30	0.80	0.10	0.50
Imports: Raw and intermediate products for agriculture	1.70	0.00	0.00	0.00	0.00	0.60	0.10	0.40
Export: Chia	1.70	0.00	2.30	0.00	1.50	1.60	1.20	0.50
Imports: Manufacturing industries	1.70	0.00	0.00	0.00	0.10	0.10	0.10	0.20
Imports: Suitcases, handbags, footwear, etc.	1.40	0.00	0.00	0.00	0.80	0.90	0.40	0.40
Export: Mineral extraction	0.80	0.00	0.00	0.00	0.30	0.10	0.10	0.20
Exports: Zinc ore	0.80	0.20	0.00	0.00	1.40	0.00	0.10	0.30
Commodity price: Oil (lag 1)	0.50	0.20	0.00	0.00	0.10	0.00	0.00	0.20
Cosine monthly-time signaling variable	0.50	0.00	0.00	0.70	1.00	0.10	1.00	0.40
Santa Cruz: Median NDVI in soybean crops (lag 9)	0.50	0.00	0.00	0.00	0.30	0.30	0.10	0.20
Imports: Non-metallic mineral manufactured products	0.50	3.70	0.00	0.00	0.10	0.90	0.20	0.20
Construction permits	0.40	0.00	0.00	0.00	0.20	0.10	0.10	0.10
Sine yearly-time signaling variable	-1.10	0.00	0.00	0.00	1.20	3.20	1.70	0.60
IGAE (lag 3)	-1.50	0.00	0.00	0.00	0.20	0.20	0.10	0.50
Rural Cochabamba: Median precipitation (lag 2)	-1.60	0.00	-0.70	0.00	0.50	0.30	0.40	0.70
IGAE (lag 1)	-1.80	0.00	0.00	0.60	2.40	2.00	1.80	0.90
Imports: Manufacture of food and beverage products	-2.30	-2.90	0.00	0.00	0.10	0.20	0.00	0.10
Imports; Chemical substances and products	-2.40	0.00	0.00	0.00	0.00	0.50	0.00	0.10
Rural Cochabamba: Median precipitation (lag 1)	-2.70	-2.80	-2.60	0.00	0.40	0.10	0.00	0.30
Generation of electrical power (lag 12)	-3.00	-6.50	-2.00	0.00	0.40	0.50	0.30	1.90
Imports: Raw and intermediate products for industry	-3.30	0.00	0.00	0.00	0.10	0.50	0.00	0.30
SantaCruz soybean crops: Median precipitation (lag 2)	-3.50	-1.60	-3.20	0.00	0.30	0.10	0.20	0.10
Tarija sugar cane crops: Median precipitation (lag 3)	-4.40	0.00	-1.00	0.00	1.20	0.20	0.60	1.70
IGAE (lag 2)	-6.30	-3.50	-4.30	14.30	3.20	0.50	2.60	5.50

In Ridge, Lasso, and ElasticNet, coefficients in the table constitute the influence magnitude and direction of features on the target variable. Positive coefficients signify a direct relationship, while negative coefficients indicate an inverse relationship. For other algorithms, the table reflects the relative importance of each feature, with individual weights summing up to 100.

Observations reveal the substantial importance of lag variables of IGAE growth and temporal signaling patterns. Notably, the 12-lag of the target variable holds the highest importance among all input variables.

Moreover, the variable representing the median value of nighttime lights in Bolivia, constructed through remote sensing, emerges as the second most important among the 44 features. Variations in luminosity become a relevant predictor, not only validating its inclusion in predicting IGAE percentage changes but also highlighting its potential for economic activity forecasting at various levels.

Expanding on remote sensing variables, the 6-lag of the average NDVI value in sugarcane crops in Santa Cruz ranks as the fourth or fifth most important variable in Ridge and ElasticNet algorithms. Given sugarcane's economic significance, especially in Santa Cruz – main agro-economic center in Bolivia –, these findings underscore its relevance for IGAE prediction.

Similarly, with lower importance, the lag 9 of the average NDVI value in wheat crops in Santa Cruz becomes another noteworthy predictor. In a different context, the lag 9 of the average NDVI value in soybean crops, a significant export product, also exhibits a positive relationship with Bolivia's economic growth, albeit with a comparatively lower contribution.

Built-up area proportion in Bolivia, derived from satellite image data, emerges as a relatively good predictor. Its importance varies across algorithms, with high rankings in Decision Tree, Random Forest, and AdaBoost models.

Conversely, lags in median precipitation for various crops show negative relationships with IGAE variations in Ridge regression. This suggests potential adverse effects of increased precipitation on crop yields, especially in irrigated areas, which is the case of the crop lands in this research.

In terms of conventional features, electricity consumption indices (small and large industry), railway transportation, tin production, aggregate electricity generation, and variations in imports of manufactured products prove to be significant predictors. In particular, it is worth noting that these variables are used in practice for the calculation of IGAE by the National Institute of Statistics of Bolivia.

3.3. Application of nowcasting to the Bolivian case

Building upon the framework established in this study, the predictions from each validated algorithm ($\hat{y}_i;\ \forall\ i = 1, \ldots, 8$) and the aggregated nowcast metrics ($\hat{y}_{\text{AM}}$, $\hat{y}_{\text{WAM}}$, $\hat{y}_{\text{WAM-BEST}}$, $\hat{y}_{\text{WGM}}$, $\hat{y}_{\text{WGM-BEST}}$) serve as inputs to derive a final nowcast measure for the monthly growth rate of the IGAE.

Before addressing the final nowcast measure for the Bolivian monthly economic growth, an associated nowcast interval has been estimated with upper and lower limits defined by the maximum ($\hat{y}_{\max}$) and minimum ($\hat{y}_{\min}$) values of the forecasts generated by the eight validated algorithms. Fig. 1 illustrates the time series of monthly IGAE growth overlaid with the nowcast interval, covering the validation period (January 2019 to September 2022) and the test period (October 2022 to January 2023).

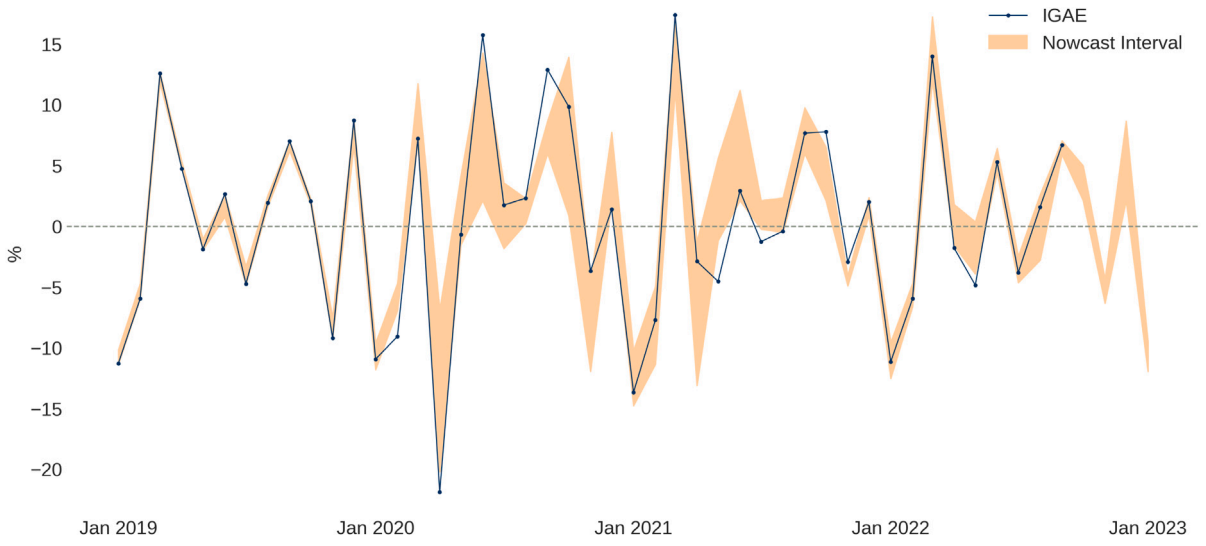


Fig. 1. Observed IGAE monthly growth rate and nowcast interval.
Note: The upper and lower limits of the nowcast interval are determined by the maximum ($\hat{y}_{\max}$) and minimum ($\hat{y}_{\min}$) values of the forecasts generated by the eight algorithms.

In most months within the analysis period, the nowcast interval encompasses the observed growth of the IGAE, adding validity and confidence to the assumption that the proposed nowcasting approach mirrors the performance of the Bolivian economy. The nowcast interval adeptly captures the seasonal behavior and directional variations in the IGAE. Notably, even during atypical events, such as the outbreak of the COVID-19 pandemic, the nowcast interval aligns with the observed performance of IGAE monthly growth.

The IGAE growth exhibits a pronounced and distinct seasonal pattern in terms of the direction (sign) of the percentage variation between months; refer to Appendix F for plots illustrating this seasonal behavior. With that in mind, it is worth mentioning that during pre-pandemic and post-pandemic months when the Bolivian economy adhered to historical seasonal patterns, the nowcast performance seems to be superior, resulting in a narrower nowcast interval that closely reflect observed values.

While the nowcast interval serves as a valuable indicator for the purposes outlined in the preceding paragraphs, stakeholders such as policymakers, researchers, and economic analysts often require specific numerical estimates for economic growth. Therefore, the final nowcast measure is determined by comparing predictive performance, specifically for the validation set, among aggregated nowcast metrics, and selecting the one with the best performance.

Table 3 depicts the RMSE, MAE, and R^2 indicators for each aggregated nowcast metric, as well as for the predictions from the eight validated algorithms. It is noteworthy that, unlike the results in Table 1, the forecast evaluation indicators here are computed with observed and predicted values in the original units of the monthly growth rate of the IGAE.[18]

Clearly, the aggregated nowcast metrics, based on average statistics, outperform the predictions from individual algorithms. In general, both the Weighted Arithmetic Mean (WAM) and Weighted Geometric Mean (WGM), considering only the three best-performing algorithms (Ridge, ElasticNet, and Gradient Boosting Regressor), show the lowest RMSE and MAE values and the highest R^2 scores.

Table 3
Forecast evaluation by aggregated nowcast metrics and algorithms.

Nowcast Metrics and Algorithms	Full Sample			Train Set			Validation Set			No-COVID Set			COVID Set		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
WAM (3-best-performing algorithms)[a]	1.80	1.04	0.93	0.61	0.46	0.99	2.75	1.92	0.89	1.34	0.97	0.96	4.13	3.49	0.82
WGM (3-best-performing algorithms)[a]	1.80	1.04	0.93	0.61	0.46	0.99	2.75	1.92	0.89	1.34	0.97	0.96	4.13	3.49	0.82
WGM (all algorithms)	1.84	0.98	0.93	0.49	0.38	0.99	2.86	1.90	0.88	1.19	0.79	0.97	4.39	3.72	0.80
WAM (all algorithms)	1.84	0.98	0.93	0.49	0.38	0.99	2.86	1.90	0.88	1.19	0.79	0.97	4.39	3.72	0.80
Gradient Boosting Regressor	1.86	0.74	0.93	0.08	0.06	1.00	2.95	1.76	0.87	1.18	0.75	0.97	4.55	3.43	0.79
Arithmetic Mean (all algorithms)	1.89	0.97	0.93	0.45	0.35	0.99	2.94	1.90	0.87	1.19	0.77	0.97	4.53	3.76	0.79
Ridge	2.00	1.22	0.92	0.88	0.67	0.98	2.97	2.06	0.87	1.70	1.22	0.94	4.31	3.43	0.81
ElasticNet	2.05	1.24	0.91	0.86	0.65	0.98	3.08	2.15	0.86	1.47	1.13	0.96	4.64	3.83	0.78
Extra Trees Regressor	2.11	1.07	0.91	0.45	0.35	0.99	3.29	2.14	0.84	1.40	1.05	0.96	5.05	3.94	0.74
Random Forest Regression	2.20	1.11	0.90	0.55	0.42	0.99	3.42	2.14	0.83	1.37	1.03	0.96	5.28	3.97	0.71
AdaBoost Regressor	2.24	0.91	0.90	0.25	0.15	1.00	3.53	2.06	0.82	1.33	0.84	0.96	5.49	4.07	0.69
Lasso	2.25	1.27	0.89	0.80	0.63	0.98	3.42	2.25	0.83	1.09	0.85	0.98	5.39	4.54	0.70
Decision Tree Regressor	2.73	1.32	0.84	0.74	0.55	0.98	4.23	2.49	0.74	2.01	1.14	0.92	6.38	4.71	0.58

Note: The forecast evaluation is computed with observed and predicted values in the original units of the monthly growth rate of the IGAE, which are percentage points. WAM and WGM stand for Weighted Arithmetic Mean and Weighted Geometric Mean, respectively.
Full Set: April 2013 to September 2022.
Train Set: April 2013 to December 2018.
Validation Set: January 2019 to September 2022.
No-COVID Set: January 2019 to February 2020 and August 2021 to September 2022.
COVID Set: March 2020 to July 2021.
[a] The three best-performing algorithms are Ridge, ElasticNet, and Gradient Boosting Regressor.

Furthermore, Table 3 provides forecast evaluation parameters for various information sets. Notably, forecast performance for the training set (April 2013 to December 2018) excels, as expected with machine learning models. However, the true predictive performance is scrutinized using the validation set (January 2019 to September 2022), revealing that the WAM or WGM, considering the three best-performing algorithms, emerge as the proper choices for a final nowcast measure.

Since the validation set includes months within the COVID-19 pandemic and other non-pandemic periods, it is challenging to find literature suitable for direct comparison of the current nowcasting approach with other methods in terms of RMSE, MAE, or R^2. Marcellino and Sivec (2021) proposed a range of methods for nowcasting GDP growth in a small open economy, with a forecast evaluation period from 2006Q3 to 2020Q3, showing RMSE and MAE values close to those in the validation set in Table 3.[19]

When dividing the validation timeframe into No-COVID and COVID sets, it becomes evident that predictive performance significantly improves under normal (No-COVID) circumstances. The RMSE and MAE for both COVID and No-COVID sets, in Table 3, are relatively comparable to some of the nowcast models tested by Dauphin et al. (2022).[20]

For illustrative purposes, Fig. 2 displays the weighted arithmetic mean of the three best-performing algorithms ($\hat{y}_{\text{WAM-BEST}}$) as the final nowcast measure, overlaid with the observed IGAE monthly growth rate.

3.4. Cumulative nowcast analysis

While originally designed to predict monthly growth rates, Fig. 3 displays the cumulative growth series in the Bolivian economy from January 2022 to January 2023, with the last four months representing the nowcast. This presentation enhances the interpretation and analysis of Bolivia's economic situation during these months.

Historically, the final quarter significantly influences Bolivia's annual economic growth, as shown in Appendix F. This influence arises from the execution of a majority of public resources, particularly public investment, during this period. Additionally, year-end festivities contribute to the dynamism of private economic activities. However, the nowcast for the last quarter of 2022 indicates a deceleration in the Bolivian economy.

[18] The results in Table 1 are calculated with z-score normalized observed and predicted values, making them not directly comparable to the results in Table 3.
[19] It is worth noting that RMSE and MAE in Marcellino and Sivec (2021) tend to be lower than those in this study, as their analysis period includes more No-COVID observations, implying fewer outliers.
[20] A cautious approach is necessary when making comparisons, as the pandemic period in Dauphin et al. (2022) only covers 2020Q1 to 2021Q1, and their non-pandemic time span is limited to data up to 2019Q4.

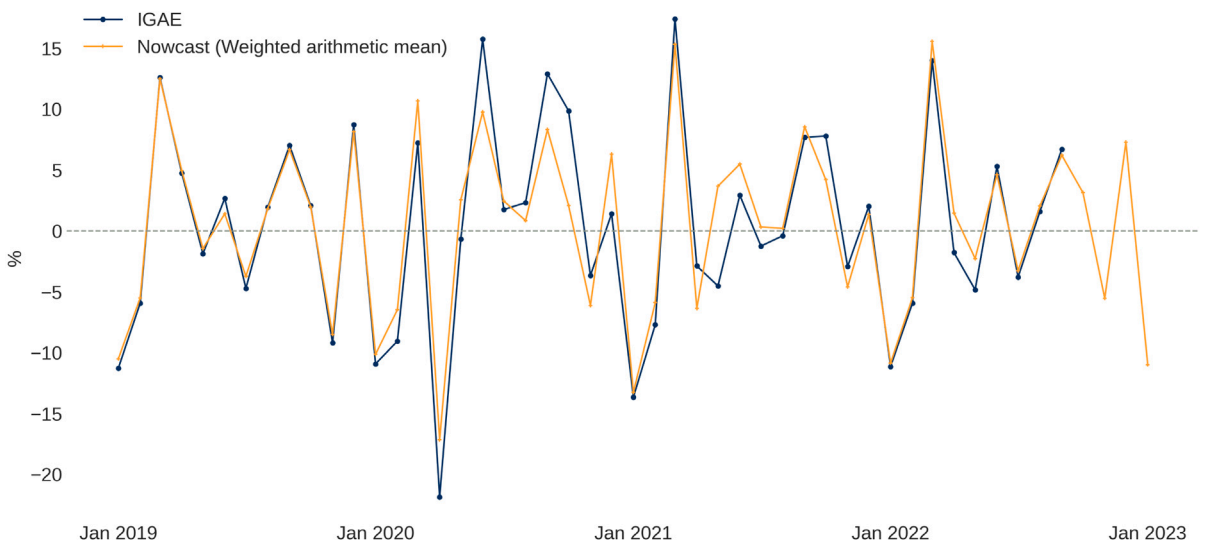


Fig. 2. Observed IGAE monthly growth rate and the final nowcast measure.
Note: The nowcast of the IGAE monthly growth rate is the weighted arithmetic mean of the three best-performing algorithms ($\hat{y}_{\text{WAM-BEST}}$).

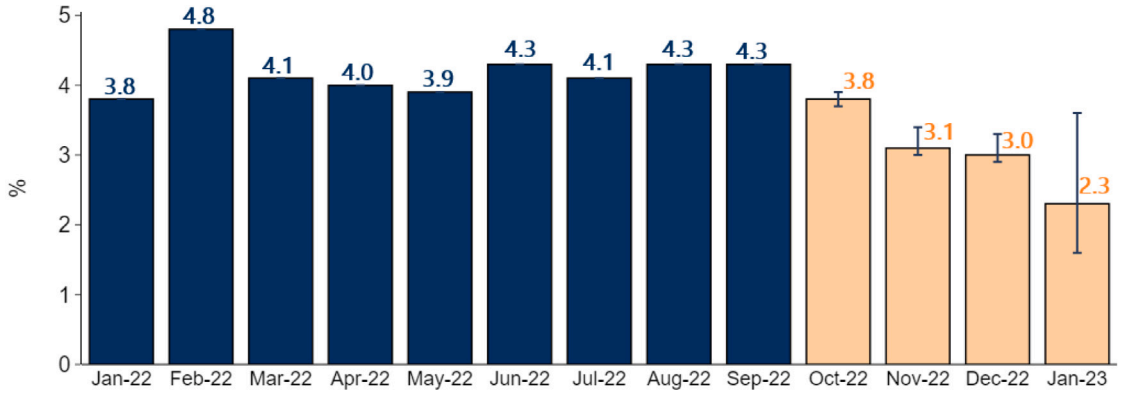


Fig. 3. Cumulative growth rate of the IGAE: Observed and nowcast.
Note: Within the nowcast period (October 2022 to January 2023), the nowcast interval was included alongside the predicted cumulative economic growth.

Conducting the nowcast during these specific months becomes invaluable for testing the predictive capacity of the methodology in the face of atypical events. Notably, from October 22, 2022, for 36 consecutive days, a civic strike occurred in Santa Cruz de la Sierra and neighboring municipalities.[21]

The effects of the civic strike rippled across the Santa Cruz region and the entire country, resulting in economic losses, particularly in commerce and transportation.[22] Santa Cruz, contributing approximately 30% to Bolivia's GDP, specializes in agricultural and agroindustry production, constituting about 35% of its GDP.[23] For instance, industrial production faced challenges due to the shortage of intermediate goods produced in Santa Cruz. Besides, other regions experienced price increases in food products, given Santa Cruz's role as a major supplier. Export activities were also limited due to mobility restrictions on trunk roads.

Despite recording a cumulative economic growth of 4.3% until September 2022, the Bolivian economy would experience a slowdown in October (3.8%) and November 2022 (3.1%), according to the IGAE nowcast. This implies negative year-on-year variations for October and November, with contractions of −0.5% and −2.6%, respectively.

Considering the last month of 2022 in light of the IGAE nowcast results, it is inferred that the Bolivian economy may conclude the period with a growth rate of approximately 3%. In terms of year-on-year growth, December would witness a recovery, with economic activity approximately 2% higher than that recorded in 2021.

[21] Civic representatives advocated for an updated census, as the last one occurred in 2012, impacting resource allocation in the regions.

[22] Government and private sector representatives estimated losses ranging from USD 500 to USD 1000 million, equivalent to 1.25% and 2.5% of the GDP, respectively.

[23] Data sourced from the National Institute of Statistics.

Based on the January 2023 nowcast indicator, the Bolivian economy is estimated to have grown by 2.3% compared to the same period in 2021. This suggests sustained product expansion, slightly surpassing the estimated year-on-year growth in December 2022.

Furthermore, the nowcast interval for January 2023 exhibits greater variability, ranging from a minimum growth of 1.6% to a maximum growth of 3.6%. The increased variability may indicate variations in predictor performance within the overall group, suggesting that certain predictors exhibit more favorable growth compared to others.

3.5. Robustness

3.5.1. Nowcasting vs. Econometric methods

The preceding sections have proven initial evidence of the efficacy of the proposed methodological approach in nowcasting monthly economic growth in Bolivia, utilizing machine learning algorithms incorporating information from both conventional sources and satellite imagery. However, a complementary evaluation involves comparing the predictive power of this approach with conventional econometric models typically used for time series forecasting.

This section introduce a comparative analysis of the predictive performance of the final nowcast measure with four alternative econometric models:

- **SARIMA:** A Seasonal Autoregressive Integrated Moving Average model designed for time series with complex seasonal patterns, relying solely on IGAE data.
- **Stepwise Least Squares:** An iterative model refinement method, initially using the 44 variables from both conventional and remote-sensing sources (listed in Appendix B).
- **Auto-Gets:** A general-to-specific variable selection technique following Sucarrat and Escribano (2012), starting with the 44 study variables.
- **BSVAR:** A Bayesian Structural Vector Autoregressive model based on Arias et al. (2018), considering a coherent structural economic rationale for endogenous variables.

Refer to Appendix G for detailed methodological aspects regarding the comparison econometric models.

Table 4 facilitates a comparison of forecast performance between the final nowcast measure and the econometric models, using indicators such as RMSE, MAE, and R^2 for the period from January 2029 to September 2022 (validation timeframe). Overall, these results suggest that the final nowcast measure, developed in this research, exhibits superior predictive performance compared to the four evaluated econometric methods. While Stepwise Least Squares or Auto-Gets outperform other econometric models, their RMSE and MAE (R^2) are relatively higher (lower) than those reported for the nowcast.

Table 4
Forecast evaluation: Nowcasting vs. Econometric models.

Model	RMSE	MAE	R2
Nowcast[a]	2.75	1.92	0.89
Stepwise	3.77	2.44	0.79
GETS	4.00	2.48	0.76
BSVAR	5.04	2.67	0.62
SARIMA	5.77	3.34	0.51

Note: Forecast evaluation is computed with observed and predicted values in the original units of the monthly growth rate of the IGAE, expressed in percentage points.

[a] Corresponds to the Weighted Arithmetic Mean of the 3 best-performing algorithms.

Furthermore, Fig. 4 depicts the distribution of forecast errors (observed minus predicted values of the monthly growth of the IGAE) for the final nowcast measure and the evaluated econometric models. The errors from the nowcast measure exhibit a more symmetrical distribution with a relatively narrow interquartile range compared to the errors from econometric models.

Conversely, econometric models, especially BSVAR and SARIMA, show greater variability in the distribution of forecast errors. The maximum errors, both positive and negative, are considerably higher in the econometric models compared to the proposed nowcast using machine learning algorithms.

In conclusion, the methodology proposed in this paper for nowcasting with machine learning algorithms, incorporating data from conventional sources and satellite images, appears to surpass evaluated econometric models in precision for forecasting the monthly growth rate of the IGAE.

3.5.2. Importance of remote-sensing predictors

The results presented thus far underscore the effectiveness of the proposed nowcasting approach, not only when compared to observed IGAE growth rates but also in contrast to other econometric methods. A distinctive feature of this approach is its integration of variables derived from remote sensing.

To further assess the robustness of the model, I conduct an exercise by evaluating predictive performance over datasets without remote-sensing variables. Table 5 facilitates this evaluation by comparing MSE, RMSE, and R^2 for the entire dataset against restricted versions excluding remote-sensing-based features. The overall trend suggests higher predictive power when incorporating features based on remote-sensing data.

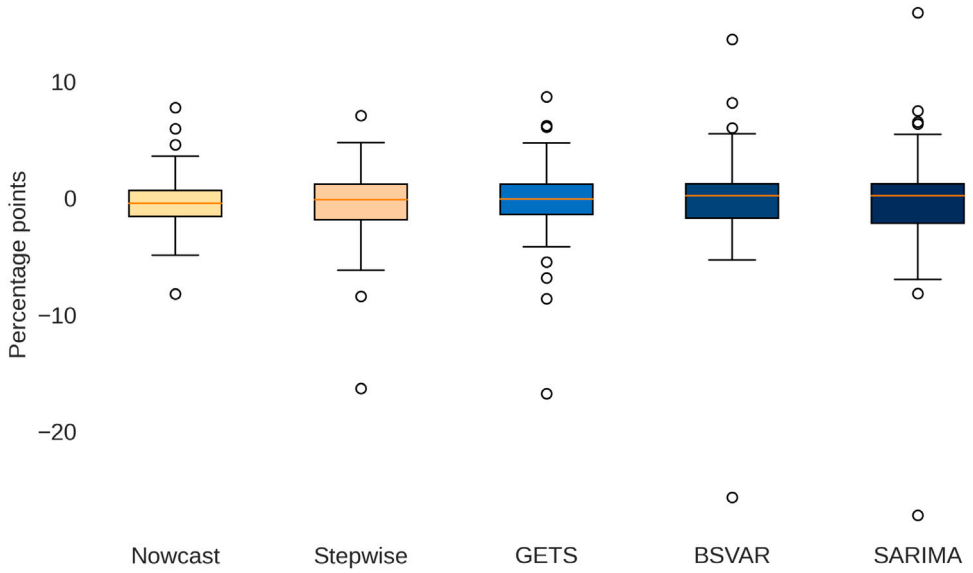


Fig. 4. Distribution of forecast errors by model.

Table 5
Forecast evaluation: All variables vs. No remote-sensing variables.

Algorithm	All Variables			No Remote-Sensing		
	MSE	RMSE	MAE	MSE	RMSE	MAE
Ridge	0.297	0.545	0.357	0.337	0.581	0.405
Lasso	0.421	0.649	0.378	0.567	0.753	0.430
ElasticNet	0.350	0.592	0.377	0.426	0.653	0.422
Decision Tree	0.712	0.844	0.420	0.687	0.829	0.404
AdaBoost	0.524	0.724	0.332	0.556	0.746	0.346
Gradient Boosting	0.390	0.625	0.292	0.439	0.663	0.300
Random Forest	0.531	0.729	0.359	0.587	0.766	0.382
Extra Trees	0.496	0.704	0.356	0.567	0.753	0.380

Note: Results are calculated with z-score normalized observed and predicted values. Hyperparameters are those of Table 1.

To delve into the importance of remote-sensing-based variables, the Relative MSE (Rel-MSE$_{i,d}$) is computed for algorithm i and dataset d, representing the ratio between the MSE when remote-sensing-based variables are excluded and the MSE when all variables are considered. Focusing on the top 3-performing algorithms (Ridge, ElasticNet, and GBR), Table 6 reveals that excluding all remote-sensing variables results in MSE 1.13 to 1.22 times higher than those when all variables are included.

Table 6
Relative MSE.

Excluded Variables	Ridge	Lasso	EN	DT	ADA	GBR	RF	ET
All Remote-Sensing	1.135	1.347	1.217	0.965	1.061	1.126	1.105	1.143
Median Nighttime Lights	1.186	1.281	1.207	0.979	1.050	1.026	1.103	1.159
Share of Built-Up Pixels	0.990	1.002	0.999	0.946	0.982	1.150	1.005	1.031
Median NDVI in Crops	1.037	0.984	1.004	1.002	1.101	1.025	1.008	0.996
Median Precipitation in Crops	0.964	1.005	1.002	0.968	1.030	1.058	0.999	1.011

Note: Results are calculated with z-score normalized observed and predicted values. Hyperparameters are those of Table 1.

As highlighted in Table 2, Bolivia's median nighttime lights significantly contribute to forecasting. Likewise, the Rel-MSE when excluding this variable disclose losses in forecast capacity. Omitting luminosity variations in the proposed nowcasting framework would compromise predictive power more than any other remote-sensing exclusion strategy outlined in Table 6.

Other variables generated by remote-sensing processes, such as the share of built-up areas, and vegetation and precipitation statistics for major Bolivian crops (rice, sugar cane, soybean, and wheat), contribute to predictive power. Their exclusion results in some loss of forecasting accuracy.

Finally, while features like the median value of nighttime lights and the proportion of built-up areas were considered for the entire territory of Bolivia, similar time series were constructed for more specific geographic areas, such as the metropolitan regions of La Paz, Kanata, and Santa Cruz. Appendix H dispenses forecast evaluations for variables on median luminosity and built-up areas

independently for these three metropolitan regions. However, the results show no significant improvement in predictive power, leading to the decision to retain remote-sensing variables for the national territory for simplicity and parsimony.

4. Concluding remarks

The paper aimed to enrich the GDP nowcasting literature by presenting an innovative strategy tailored for a developing country context, where timely and reliable economic data is often elusive. Bolivia, in particular, experiences a significant delay of up to six months in releasing its monthly economic growth indicator, the Global Index of Economic Activity (IGAE). A key result from this study is the reduction of the gap between real-time estimates and official IGAE reports from six months to just two months.

This improvement is achieved by demonstrating the efficacy of combining machine learning techniques with data encompassing indicators from traditional sources (administrative records from productive sectors, trade, commodities, etc.) and statistics derived from satellite imagery, including nighttime lights, vegetation indices, built-up areas, and precipitation.

By addressing data constraints, this study not only advances the understanding of economic dynamics in Bolivia but also lays the groundwork for similar analyses in regions facing challenges related to timely information availability.

The study employs eight machine learning algorithms, with Ridge, ElasticNet, and Gradient Boosting Regressor emerging as the top-performing models. Furthermore, the paper introduces aggregated nowcasting metrics, combining individual forecasts through averaging methods, and a nowcast interval to outline the potential range of values for the IGAE growth rate.

The nowcast interval, defined by the upper and lower limits of forecasts from the eight algorithms, mirrors the observed growth of the IGAE throughout validation periods. Notably, even during atypical events such as the COVID-19 pandemic, the nowcast interval aligns closely with the actual performance of IGAE monthly growth, underscoring the adaptability of the proposed method.

Additionally, the aggregated nowcast metrics, based on average statistics, surpass prediction power from individual algorithms. Specifically, both the weighted arithmetic mean and weighted geometric mean,[24] considering only the three best-performing algorithms (Ridge, ElasticNet, and Gradient Boosting Regressor), demonstrate the lowest RMSE and MAE values and the highest R2 scores.

The paper evaluates the nowcasting approach using various criteria, indicating superior performance compared to naive benchmarks like conventional econometric models. Likewise, the robustness of nowcasting is tested by assessing sensitivity to different feature sets, with and without remote-sensing-based data.

As of the completion of this research, official reports provided IGAE time series up to September 2022, and nowcast estimates were generated for the last three months of 2022 and January 2023. This specific timeframe becomes particularly valuable for testing the proposed methodology in the face of atypical events, because of Bolivia's economic challenges due to social conflicts from October to November 2022.

While this study contributes to the field of economic nowcasting for developing countries, several avenues for future research emerge. Firstly, the nowcast proposed here relies on the availability of mainly contemporaneous information, posing a challenge for extending the forecast horizon beyond current data availability. Secondly, although the study utilizes a set of features with strong correlations to the target variable, exploring additional variables or transformations to enhance predictive performance remains a pending task for further investigation. Lastly, the choice of machine learning algorithms is crucial, and future extensions will consider testing additional algorithms or ensemble methods to expand the model pool.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Potential predictors from conventional sources

See Table A.1.

Appendix B. Correlation between target variable and features

See Table B.1.

Appendix C. Examples of remote sensing imagery

See Figs. C.1 and C.2.

[24] Weights are the inverse mean square errors from individual algorithms.

Table A.1
Potential predictors from conventional sources.

Sector	Potential predictors	Source
Production	The generation of electrical power, the manufacture of cement, and the extraction of oil, natural gas, and various minerals including tin, copper, lead, zinc, tungsten, silver, antimony, and gold.	National Institute of Statistics
Services	Electricity consumption across various sectors, including large and small industries, mining, residential, public urban illumination, and rural areas; water and transport consumption indices, as well as construction permits.	National Institute of Statistics
Fiscal	Domestic and customs tax revenues.	Ministry of Economy and Public Finance[a]
Monetary	Monetary aggregates, base and monetary emission, and interest rates.	Central Bank of Bolivia
Financial	Deposits and credits, and reserve requirements.	Central Bank of Bolivia[b]
Trade	Numerous disaggregations of both exports and imports.	National Institute of Statistics
Prices	Consumer Price Index (CPI) in its entirety, as well as its various categories, alongside commodity prices.	National Institute of Statistics

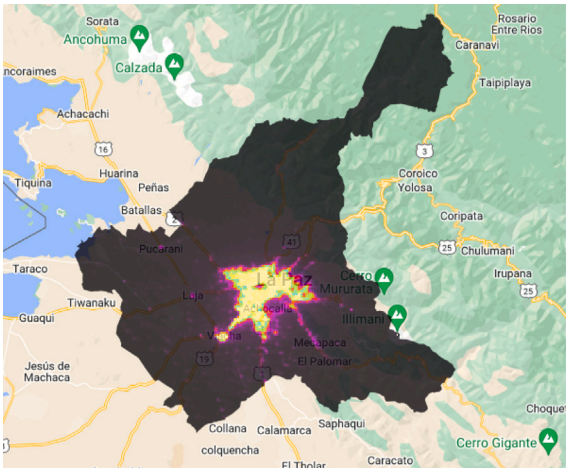
[a] Tax revenue time series are constructed with information from publications and other online sources of Ministry of Economy and Public Finance, National Tax Service, and National Customs.
[b] Financial data is also extracted from publications from the Financial System Supervisory Authority.

Table B.1
Correlation between target variable and features.

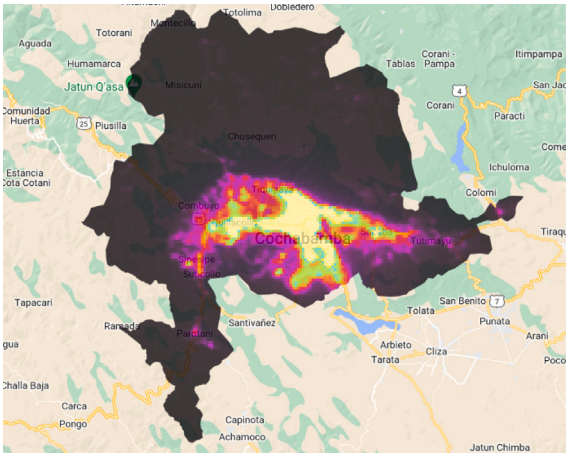
Features	Correlation	P-value	Source
IGAE (lag 12)	0.88	0.0000	Conventional
Bolivia: Median Nighttime Lights	0.71	0.0000	Remote sensing
Generation of electrical power	0.68	0.0000	Conventional
Sine monthly-time signaling variable	0.58	0.0000	Conventional
Imports: Manufacture of clothing	0.53	0.0000	Conventional
Imports: Raw and intermediate products for agriculture	0.49	0.0000	Conventional
Generation of electrical power (lag 12)	0.49	0.0000	Conventional
Imports; Chemical substances and products	0.48	0.0000	Conventional
Imports: Manufacture of rubber and plastic products	0.46	0.0000	Conventional
Production: Liquefied petroleum gas	0.46	0.0000	Conventional
Imports: Manufacturing industries	0.45	0.0000	Conventional
Imports: Suitcases, handbags, footwear, etc.	0.45	0.0000	Conventional
Export: Mineral extraction	0.42	0.0000	Conventional
Exports: Wood and wood products	0.42	0.0000	Conventional
Imports: Durable Consumer Goods	0.42	0.0000	Conventional
IGAE (lag 6)	0.42	0.0000	Conventional
Imports: Manufacture of food and beverage products	0.41	0.0000	Conventional
Production: Tin	0.39	0.0000	Conventional
Imports: Raw and intermediate products for industry	0.39	0.0000	Conventional
Exports: Zinc ore	0.39	0.0000	Conventional
Imports: Non-metallic mineral manufactured products	0.38	0.0000	Conventional
Small industries: Electricity consumption	0.37	0.0000	Conventional
Construction permits	0.37	0.0000	Conventional
Rail transport consumption index	0.36	0.0001	Conventional
Large industries: Electricity consumption	0.35	0.0001	Conventional
Exports: Agriculture, livestock, forestry and fishing	0.35	0.0001	Conventional
Santa Cruz: Median NDVI in wheat crops (lag 9)	0.34	0.0002	Remote sensing
Commodity price: Gold (lag 2)	0.33	0.0003	Conventional
Export: Chia	0.32	0.0004	Conventional
Santa Cruz: Median NDVI in soybean crops (lag 9)	0.30	0.0010	Remote sensing
Santa Cruz: Average NDVI in sugar cane crops (lag 6)	0.28	0.0025	Remote sensing
Commodity price: Oil (lag 1)	0.27	0.0027	Conventional
Bolivia: Share of built-up pixels	0.26	0.0053	Remote sensing
Rural Cochabamba: Median precipitation (lag 1)	0.16	0.0867	Remote sensing
Cosine yearly-time signaling variable	0.15	0.1144	Conventional
IGAE (lag 9)	0.15	0.1179	Conventional
IGAE (lag 3)	0.11	0.2362	Conventional
Rural Cochabamba: Median precipitation (lag 2)	0.03	0.7725	Remote sensing
Cosine monthly-time signaling variable	0.00	0.9787	Conventional
Sine yearly-time signaling variable	−0.05	0.5695	Conventional
IGAE (lag 1)	−0.06	0.4979	Conventional
Santa Cruz soybean crops: Median precipitation (lag 2)	−0.22	0.0199	Remote sensing
IGAE (lag 2)	−0.32	0.0004	Conventional
Tarija sugar cane crops: Median precipitation (lag 3)	−0.33	0.0003	Remote sensing

Note: Pearson correlation coefficients and p-values are calculated with z-score normalized variables.

(a) Metropolitan area of La Paz



(b) Metropolitan area of Kanata



(c) Metropolitan area of Santa Cruz

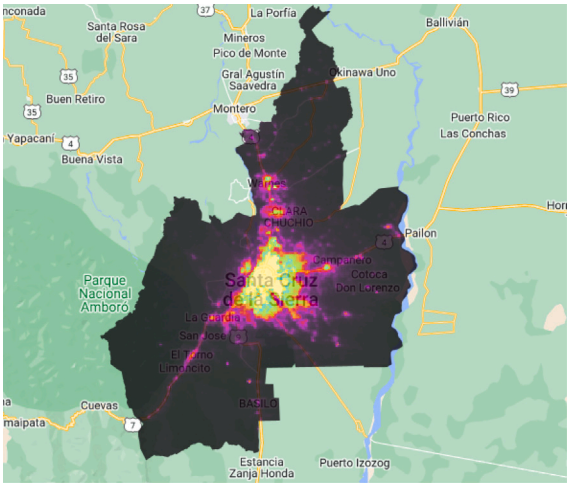


Fig. C.1. Median luminosity intensity for the metropolitan areas of La Paz, Kanata, and Santa Cruz in the year 2022,
Note: The intensity of luminosity in each pixel is defined by a color palette of black, purple, red, orange, green, and yellow. The lowest levels are represented by black, while the highest levels are represented by yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(a) Satellite imagery

(b) NDVI

Fig. C.2. Satellite and NDVI mapping soybean crop lands.
Note: Regarding the NDVI image, crops are highlighted in a yellowish color, while areas with abundant vegetation tend to be fluorescent green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Appendix D. Built-up classifier random forest model

To predict economic growth, it is important to possess time series data on the changes in built-up areas for Bolivia. To address this objective, a Random Forest algorithm is trained using GHSL images to classify 1 km pixels into built-up or non-built-up areas, using data from the Landsat −8 satellite's blue, green, red, near-infrared, and shortwave infrared 1 and 2 bands and pixel-level light brightness levels.

The GHSL images classify areas according to their degree of urbanization, such as uninhabited, rural, and urban clusters of low and high density. However, as this study focuses on a binary classification of built-up areas, urban clusters of low and high density are defined as built-up areas for algorithm training, while uninhabited and rural areas are defined as non-built-up areas. The images used in the training stage are from the 2015, which is the latest update of GHSL. The pixels within the images are randomly separated into a training set (70%) and a validation set (30%).

To evaluate the classifier algorithm's performance, Accuracy (ACC) and the Matthews Correlation Coefficient (MCC) are used as metrics. The ACC measures the proportion of correct predictions made by the classifier, while the MCC measures the correlation between the predicted and actual labels, taking into account both true positive and true negative rates. The performance of the algorithm on the training sample is usually higher than that on the validation sample, because the algorithm has learned from the training sample and has optimized its parameters to fit the training data. Therefore, the ACC and MCC on the training sample are usually higher than those on the validation sample.

MCC is more appropriate for classifying built-up areas since most GHSL pixels in Bolivia are non-built-up areas; MCC outperforms ACC and other metrics in binary classification evaluation (Chicco and Jurman, 2020). The region of interest for classification encompasses zones around Bolivia's main urban areas, with a buffer extension to ensure inclusion of representative built-up and non-built-up areas from different regions of the country and varied geographical features. The region of interest contains a total of 5722 1 km pixels (i.e., 5722 observations; 4005 in the training set and 1717 in the validation set).

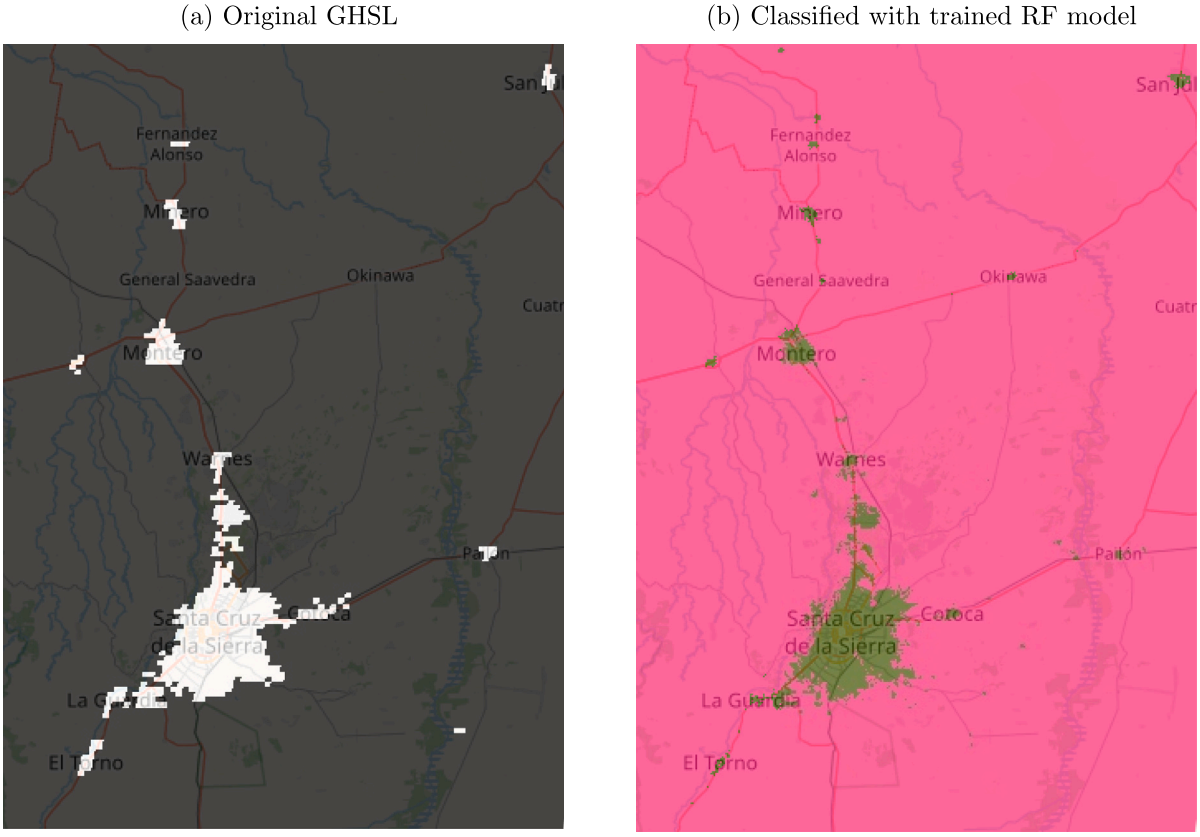


Fig. D.1. Comparison between GHSL and Random Forest classified built-up areas.
Note: In the image on the left, built-up areas are identified with white color and non-built-up areas with black color. In the image on the right, built-up areas are identified with green color and non-built-up areas with pink color. Due to the spatial resolution of Landsat −8 and VIIRS night lights images, the built-up area classification image is less pixelated and, therefore, can provide more granular inputs regarding the identification of this type of area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In the training sample, the values of ACC and MCC are considerably high, measuring 0.995 and 0.917, respectively; these results indicate an exceptional performance in terms of classification within this specific subset. Regarding performance on the validation sample, the ACC remains considerably high (0.987); however, the MCC is relatively lower (0.702) compared to that observed in the

training sample. This signals that the algorithm may have overfitted to the training data and is not generalizing well to new data. Nonetheless, in general, an MCC score of 0.5 or above is considered a good performance, while a score of 0 indicates a completely random classifier. A comparison between GHSL and Random Forest classified built-up is presented in Fig. D.1.

Appendix E. Methodological characteristics and execution of the machine learning algorithms

E.1. Ridge, lasso and ElasticNet

Ridge, Lasso, and ElasticNet are three popular regularization techniques used in linear regression models to prevent overfitting. These techniques add a penalty term to the objective function, which controls the complexity of the model and reduces the impact of irrelevant features.

Ridge regression, also known as L2 regularization, adds a penalty term to the sum of squared errors (SSE) objective function. The penalty term is proportional to the square of the L2 norm of the coefficients vector. The objective function for Ridge regression is given by:

$$minimize\ J(\beta) = SSE + \lambda \sum_{j=1}^{p} \beta_j^2$$

where β is the vector of coefficients, p is the number of features, and λ is the regularization parameter that controls the strength of the penalty term.

Lasso regression, also known as L1 regularization, adds a penalty term to the SSE objective function. The penalty term is proportional to the L1 norm of the coefficients vector. The objective function for Lasso regression is given by:

$$minimize\ J(\beta) = SSE + \lambda \sum_{j=1}^{p} |\beta_j|$$

ElasticNet is a combination of Ridge and Lasso regression techniques. It adds both L1 and L2 penalty terms to the SSE objective function. The objective function for ElasticNet regression is given by:

$$minimize\ J(\beta) = SSE + \lambda \left(\alpha \sum_{j=1}^{p} |\beta_j| + (1-\alpha) \sum_{j=1}^{p} \beta_j^2 \right)$$

The hyperparameter λ controls the strength of the regularization, with larger values of λ leading to stronger regularization. The hyperparameter α controls the mixing parameter between L1 and L2 regularization, with values between 0 and 1. A value of 1 corresponds to L1 regularization, while a value of 0 corresponds to L2 regularization. Intermediate values correspond to a combination of both L1 and L2 regularization. In summary, Ridge regression shrinks the coefficients towards zero, Lasso regression sets some coefficients to zero, and ElasticNet combines both techniques to overcome their limitations.

E.2. Decision tree regressor

The Decision Tree Regressor works by recursively partitioning the feature space into smaller and smaller regions based on the values of the input features. Let X be the input feature matrix of size $n \times m$, where n is the number of observations and m is the number of features. Let y be the corresponding target variable vector of size n. The algorithm works as follows:

1. Choose a feature j and a threshold value t that best splits the data into two subsets R_1 and R_2 based on the mean squared error (MSE) loss function:

   $$MSE(R) = \frac{1}{|R|} \sum_{i \in R} (y_i - \bar{y}_R)^2$$

   where $|R|$ is the number of observations in region R and $\bar{y}_R$ is the mean target value in region R.
2. Repeat step 1 recursively on each subset R_1 and R_2 until a stopping criterion is met, such as a maximum depth ($MaxDepth$) or minimum number of observations in a leaf node ($MinSamplesLeaf$).
3. Assign the mean target value of the training observations in each leaf node as the predicted value for new observations that fall within that region.

The algorithm can be represented as a binary tree, where each internal node represents a split on a feature and threshold value, and each leaf node represents a predicted target value. The tree can be trained using an iterative algorithm such as CART (Classification and Regression Trees).

E.3. AdaBoost regressor

AdaBoost Regressor is a boosting algorithm used for regression tasks. It trains a series of weak learners on weighted versions of the training data, with the weights updated at each iteration to focus on the data points that were poorly predicted in previous iterations. The final model is an ensemble of the weak learners, weighted by their performance.

Let $D = (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be the training data, where x_i is a vector of input features and y_i is the corresponding output label. Let $h(x)$ be a weak learner that takes an input x and outputs a prediction $\hat{y}$.

The algorithm proceeds as follows:

1. Initialize the sample weights $w_i = 1/n$ for $i = 1, 2, \ldots, n$.
2. For $t = 1, 2, \ldots, T$ do:
   a. Train a weak learner $h_t(x)$ on the training data D with weights w_i.
   b. Compute the weighted error ϵ_t of the weak learner:

$$\epsilon_t = \frac{\sum_{i=1}^{n} w_i \left| y_i - h_t\left(x_i\right) \right|}{\sum_{i=1}^{n} w_i} \tag{E.1}$$

   c. Compute the learning rate α_t:

$$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right) \tag{E.2}$$

   d. Update the sample weights w_i:

$$w_i \leftarrow w_i \cdot \exp\left(-\alpha_t \left(y_i - h_t\left(x_i\right)\right)\right) \tag{E.3}$$

   e. Normalize the weights:

$$w_i \leftarrow \frac{w_j}{\sum_{j=1}^{n} w_j} \tag{E.4}$$

3. Output the final model:

$$f(x) = \sum_{t=1}^{T} \alpha_t h_t(x) \tag{E.5}$$

In this context, T denotes the number of weak learners, also referred to as base estimators, while ϵ_t measures the error of the weak learner $h_t(x)$ on the training data. The learning rate α_t governs the contribution of each weak learner to the final ensemble model, with a small value chosen in the case of high error rates, and a large value chosen when the error rate is low. The loss function utilized in AdaBoost Regressor is typically the exponential loss function, which places greater emphasis on data points that are difficult to predict.

E.4. Gradient boosting regressor

Gradient Boosting Regressor (GBR) is an ensemble method that combines multiple weak learners to create a strong predictor. Let y_i be the actual value of the target variable for the ith observation, and let $\hat{y}_i$ be the predicted value of the target variable. Then, the MSE between the predicted and actual values is given by:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

The GBR algorithm works by sequentially adding new regression trees to the ensemble, with each tree attempting to correct the errors of the previous trees. Let $f_m(x)$ be the prediction of the mth tree for the input vector x. Then, the prediction of the ensemble after M trees is given by:

$$\hat{y}_i^{(M)} = \sum_{m=1}^{M} \gamma_m f_m(x_i)$$

where γ_m is the learning rate for the mth tree, which controls the contribution of each tree to the final prediction. The learning rate is typically set to a small value, such as 0.1, to prevent overfitting.

To train the GBR, it starts with an initial prediction $\hat{y}_i^{(0)}$, which is typically set to the mean value of the target variable. Then iteratively add new trees to the ensemble by minimizing the residual errors of the previous trees. Specifically, at each iteration m, fit a new regression tree $f_m(x)$ to the negative gradients of the MSE loss function:

$$r_{im} = -\frac{\partial}{\partial \hat{y}_i^{(m-1)}} L(y_i, \hat{y}_i^{(m-1)})$$

where $L(y_i, \hat{y}_i^{(m-1)})$ is the loss function used to measure the error between the predicted and actual values. The negative gradients are then used as the target values for the new tree, and the learning rate is applied to control the contribution of the new tree to the ensemble. The final prediction of the GBR is obtained by summing the predictions of all the trees in the ensemble:

$$\hat{y}_i = \sum_{m=1}^{M} \gamma_m f_m(x_i)$$

E.5. Random forest regression

Random Forest Regression is an ensemble learning method that combines multiple decision trees to make a more accurate prediction. In this algorithm, a random subset of features is selected at each node of the decision tree, and the best split is made based on the selected features. The final prediction is made by averaging the predictions of all the decision trees in the forest.

Let X be the input data with n samples and m features, and y be the target variable. The Random Forest Regression algorithm can be described as follows:

1. Initialize the number of decision trees T, the number of features to be selected at each node m', and the maximum depth of each tree d.
2. For $t = 1$ to T:
    a. Randomly select m' features from the m features.
    b. Construct a decision tree using the selected features with a maximum depth of d.
3. For a new input x', predict the target variable y' as:

$$y' = \frac{1}{T} \sum_{t=1}^{T} f_t(x')$$

where $f_t(x')$ is the prediction of the tth decision tree.

E.6. Extra trees regressor

The Extra Trees Regressor algorithm is similar to the Random Forest algorithm, but with some key differences in the way the trees are constructed. Let X be the input features and y be the target variable. The Extra Trees Regressor algorithm works as follows:

1. Randomly select a subset of the input features for each tree to use.
2. For each tree, randomly select a subset of the training data to use for training.
3. For each split in each tree, randomly select a subset of features to use for determining the best split.
4. Split the data based on the selected feature and threshold value that minimizes the mean squared error (MSE).
5. Repeat steps 3–4 until the tree reaches a maximum depth or the number of samples in a leaf node is below a certain threshold.

The predicted output value for a new input sample x can then be obtained by averaging the output values of all the trees in the ensemble:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^{N} f_i(x) \tag{E.6}$$

where N is the number of trees in the ensemble and $f_i(x)$ is the output value of the ith tree for the input sample x.

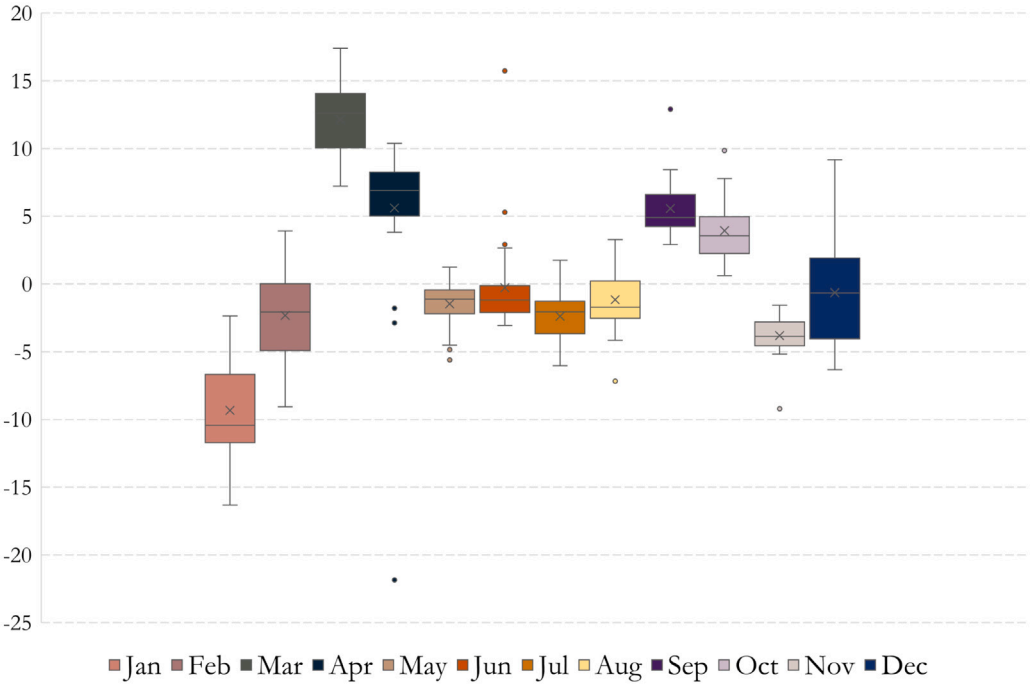
Appendix F. IGAE seasonal patterns

See Fig. F.1.

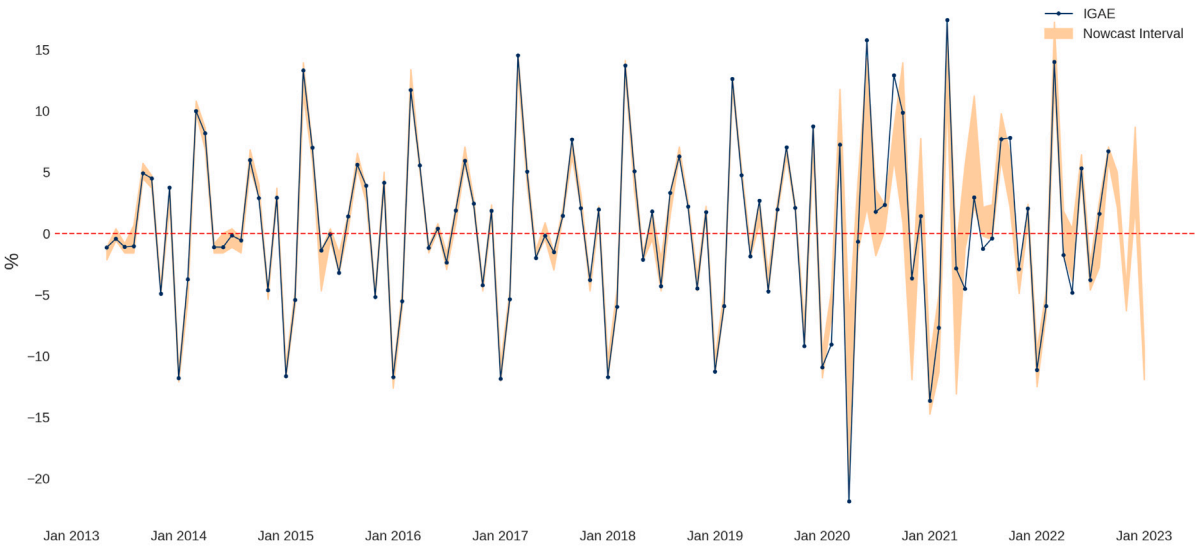
Appendix G. Econometric models for comparison

- **Univariate Model:** I use a Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The optimal specification of the model that minimizes the value of the information criteria was obtained using the *"auto.arima"* function in the pmdarima library in Python.[25] Four SARIMA models are developed, each with a 1-year projection horizon, to compare the predictive capacity of the models over the validation set period (January 2019 to September 2022). For instance, the first model is trained with data from 1990M2 to 2018M12 and its projection covers the period 2019M1-2019M12. The second model is trained with data from 1990M2 to 2019M12 and its projections cover the period 2020M1 to 2020M12, and so on.

[25] The specification is SARIMAX(1, 0, 2)x(2, 0, [1], 12).



(a) Box Plot: IGAE monthly growth rate by month



(b) Observed IGAE monthly growth rate and Nowcast Interval

Fig. F.1. (a) depicts a box plot constructed using IGAE data encompassing the period January 1990 to September 2022. This visualization summarizes the distribution of the IGAE within this timeframe.
(b) presents the monthly growth rate of the IGAE alongside its associated nowcast interval. The data depicted in this figure corresponds to the training, validation, and testing periods employed in this study, spanning January 2013 to January 2023.

- **Variable Selection Models:** These models are trained using data from April 2013 to December 2018, followed by a conditional forecast for the validation period from January 2019 to September 2022.
  - The first model in this category is Stepwise Least Squares, a method that systematically refines a model by iteratively adding or removing variables based on statistical criteria, primarily minimizing the sum of squared differences between observed and predicted values. The procedure combines forward selection and backward elimination, achieving a balance between model parsimony and explanatory power. The overarching goal is to strike an equilibrium by incorporating the most relevant variables while avoiding overfitting. Variable selection considers all 44 variables developed in this research.
  - The second model in this category is Auto-Gets, short for Automated General-to-Specific modeling, presenting a comprehensive approach in econometrics for model selection and specification. Auto-Gets aims to produce models that not only accurately represent the data but also adhere to principles of simplicity and interpretability. Following Sucarrat and Escribano (2012), the modeling process starts with the 44 variables developed in the current research.

- **Multivariate Structural Model:**
  Following Arias et al. (2018), I estimate a Bayesian Structural Vector Autoregressive (BSVAR) model, allowing for the imposition of restrictions on the identification of zeros, signs, and magnitudes for contemporary and subsequent periods associated with structural shocks. The endogenous variables in this model include the year-on-year growth of IGAE, inflation, legal reserve requirements surplus, and the price of oil, whose dynamics are explained by supply, demand, monetary policy, and commodity price shocks. The identification strategy is illustrated in Table G.1.
  The forecasts of this econometric model, transformed into monthly growth rates, are included in the comparison with the nowcast indicator. This advanced model for time series forecasting captures the dynamics (posterior probability distributions) of the main shocks determining the behavior of the Bolivian economy, thus expecting a good predictive performance. The BSVAR model is trained on a sample of monthly data from January 2007 to December 2018, and the growth of IGAE is forecasted for the months of the validation set conditional on the observed values of the other endogenous variables.

Table G.1
Structural shocks identification strategy.

Shocks	IGAE growth	Inflation	Reserve requirements	Oil price
Supply	+	−	.	0
Demand	+	+	.	0
Monetary Policy	.	.	+	0
Commodity price	.	.	.	+

Note: Supply shocks are distinguished from demand shocks as they generate reductions in prices, associated with improvements in productivity (i.e., lower production costs). The prices of commodities are exogenous to internal conditions (i.e., zero effect of domestic supply and demand on international prices of commodities), consistent with the characteristics of the Bolivian economy. Expansionary monetary policy shocks increase the liquidity of the financial system (i.e., through a larger surplus of legal reserves), allowing the model to determine contemporary effects on output and prices. (.) denotes no imposed condition.

Appendix H. Forecast evaluation: Remote-sensing features at national and regional level

See Table H.1.

Table H.1
Forecast evaluation: Remote-sensing features at national and regional level.

Algorithm	Bolivia		3 light vars		3 built-up vars	
	MSE	MAE	MSE	MAE	MSE	MAE
Ridge	0.297	0.357	0.293	0.347	0.301	0.358
Lasso	0.421	0.378	0.415	0.373	0.422	0.379
ElasticNet	0.350	0.377	0.337	0.367	0.353	0.379
Decision Tree	0.712	0.420	0.633	0.358	0.741	0.456
AdaBoost	0.524	0.332	0.515	0.344	0.527	0.344
Gradient Boosting	0.390	0.292	0.391	0.295	0.382	0.288
Random Forest	0.531	0.359	0.530	0.354	0.531	0.358
Extra Trees	0.496	0.356	0.463	0.339	0.470	0.344

Note: Results are calculated with z-score normalized observed and predicted values. Hyperparameters are those of Table 1.
3 light vars: 3 time series on median luminosity for the metropolitan regions of La Paz, Kanata, and Santa Cruz.
3 built-up vars: 3 time series on built-up areas for the metropolitan regions of La Paz, Kanata, and Santa Cruz.

References

Arias, J.E., Rubio-Ramírez, J.F., Waggoner, D.F., 2018. Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications. Econometrica 86 (2), 685–720.
Bajari, P., Nekipelov, D., Ryan, S.P., Yang, M., 2015. Machine learning methods for demand estimation. Amer. Econ. Rev. 105 (5), 481–485.
Balcilar, M., Gabauer, D., Gupta, R., Pierdzioch, C., 2022. Uncertainty and forecastability of regional output growth in the UK: Evidence from machine learning. J. Forecast. 41 (6), 1049–1064.

Bańbura, M., Giannone, D., Modugno, M., Reichlin, L., 2013. Now-casting and the real-time data flow. In: Handbook of Economic Forecasting. Vol. 2, Elsevier, pp. 195–237.
Belly, G., Boeckelmann, L., Caicedo Graciano, C.M., Di Iorio, A., Istrefi, K., Siakoulis, V., Stalla-Bourdillon, A., 2023. Forecasting Sovereign risk in the Euro area via machine learning. J. Forecast. 42 (3), 657–684.
Bok, B., Caratelli, D., Giannone, D., Sbordone, A.M., Tambalotti, A., 2018. Macroeconomic nowcasting and forecasting with big data. Annu. Rev. Econ. 10, 615–643.
Bonato, M., Çepni, O., Gupta, R., Pierdzioch, C., 2022. El Niño, La Niña, and forecastability of the realized variance of agricultural commodity prices: Evidence from a machine learning approach. J. Forecast..
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Chen, X., Nordhaus, W.D., 2011. Using luminosity data as a proxy for economic statistics. Proc. Natl. Acad. Sci. 108 (21), 8589–8594.
Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21, 1–13.
Corbane, C., Syrris, V., Sabo, F., Politis, P., Melchiorri, M., Pesaresi, M., Soille, P., Kemper, T., 2021. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. Neural Comput. Appl. 33, 6697–6720.
Dauphin, M.J.-F., Dybczak, M.K., Maneely, M., Sanjani, M.T., Suphaphiphat, M.N., Wang, Y., Zhang, H., 2022. Nowcasting GDP-A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies. International Monetary Fund.
Dietterich, T.G., 2000. Ensemble methods in machine learning. In: International Workshop on Multiple Classifier Systems. Springer, pp. 1–15.
Donaldson, D., Storeygard, A., 2016. The view from above: Applications of satellite data in economics. J. Econ. Perspect. 30 (4), 171–198.
Duro, D.C., Franklin, S.E., Dubé, M.G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens. Environ. 118, 259–272.
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63, 3–42.
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. Open Geogr. J. 3 (1).
Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. J. Monet. Econ. 55 (4), 665–676.
Gogas, P., Papadimitriou, T., Sofianos, E., 2022. Forecasting unemployment in the euro area with machine learning. J. Forecast. 41 (3), 551–566.
Henderson, J.V., Storeygard, A., Weil, D.N., 2012. Measuring economic growth from outer space. Am. Econ. Rev. 102 (2), 994–1028.
Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12 (1), 55–67.
Hu, M., Xia, B., 2019. A significant increase in the normalized difference vegetation index during the rapid economic development in the Pearl River Delta of China. Land Degradat. Develop. 30 (4), 359–370.
Jiang, S., Wei, G., Zhang, Z., Wang, Y., Xu, M., Wang, Q., Das, P., Liu, B., 2020. Detecting the dynamics of urban growth in Africa using DMSP/OLS nighttime light data. Land 10 (1), 13.
Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., Landivar-Bowles, J., 2021. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. Curr. Opin. Biotechnol. 70, 15–22.
Keola, S., Andersson, M., Hall, O., 2015. Monitoring economic development from space: using nighttime light and land cover data to measure economic growth. World Dev. 66, 322–334.
Kotsiantis, S., 2011. Feature selection for machine learning classification problems: a recent overview. Artif. Intell. Rev. 42 (1), 157–176.
Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sens. Lett. 14 (5), 778–782.
Liu, L., Chen, C., Wang, B., 2022. Predicting financial crises with machine learning methods. J. Forecast. 41 (5), 871–910.
Liu, Y., Zuo, R., Dong, Y., 2021. Analysis of temporal and spatial characteristics of urban expansion in Xiaonan District from 1990 to 2020 using time series Landsat imagery. Remote Sens. 13 (21), 4299.
Loh, W.-Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 1 (1), 14–23.
Mantovani, R.G., Horváth, T., Cerri, R., Junior, S.B., Vanschoren, J., de Carvalho, A.C.P.d.F., 2018. An empirical study on hyperparameter tuning of decision trees. arXiv preprint arXiv:1812.02207.
Marcellino, M., Sivec, V., 2021. Nowcasting GDP growth in a small open economy. Natl. Inst. Econ. Rev. 256, 127–161.
Medeiros, M.C., Vasconcelos, G.F., Veiga, Á., Zilberman, E., 2021. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. J. Bus. Econom. Statist. 39 (1), 98–119.
Melchiorri, M., Florczyk, A.J., Freire, S., Schiavina, M., Pesaresi, M., Kemper, T., 2018. Unveiling 25 years of planetary urbanization with remote sensing: Perspectives from the global human settlement layer. Remote Sens. 10 (5), 768.
Milunovich, G., 2020. Forecasting Australia's real house price index: A comparison of time series and machine learning methods. J. Forecast. 39 (7), 1098–1118.
Mullainathan, S., Spiess, J., 2017. Machine learning: An applied econometric approach. J. Econ. Perspect. 31, 87–106.
Pinkovskiy, M., Sala-i Martin, X., 2016. Lights, camera... income! Illuminating the national accounts-household surveys debate. Q. J. Econ. 131 (2), 579–631.
Plakandaras, V., Papadimitriou, T., Gogas, P., 2015. Forecasting daily and monthly exchange rates with machine learning techniques. J. Forecast. 34 (7), 560–573.
Román, M.O., Wang, Z., Sun, Q., Kalb, V., Miller, S.D., Molthan, A., Schultz, L., Bell, J., Stokes, E.C., Pandey, B., et al., 2018. NASA's Black Marble nighttime lights product suite. Remote Sens. Environ. 210, 113–143.
Sharifi, A., 2021. Yield prediction with machine learning algorithms and satellite images. J. Sci. Food Agric. 101 (3), 891–896.
Storeygard, A., 2016. Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. Rev. Econ. Stud. 83 (3), 1263–1295.
Sucarrat, G., Escribano, A., 2012. Automated model selection in finance: General-to-specific modelling of the mean and volatility specifications. Oxf. Bull. Econ. Stat. 74 (5), 716–735.
Tang, B., Liu, Y., Matteson, D.S., 2022. Predicting poverty with vegetation index. Appl. Econ. Perspect. Policy 44 (2), 930–945.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 58 (1), 267–288.
Wiegand, C., Richardson, A., Escobar, D., Gerbermann, A., 1991. Vegetation indices in crop assessments. Remote Sens. Environ. 35 (2–3), 105–119.
Wolpert, D.H., 1992. Stacked generalization. Neural Netw. 5 (2), 241–259.
Zhang, R., Tian, Z., McCarthy, K.J., Wang, X., Zhang, K., 2023. Application of machine learning techniques to predict entrepreneurial firm valuation. J. Forecast. 42 (2), 402–417.
Zhou, M., Lu, L., Guo, H., Weng, Q., Cao, S., Zhang, S., Li, Q., 2021. Urban sprawl and changes in land-use efficiency in the Beijing–Tianjin–Hebei region, China from 2000 to 2020: A spatiotemporal analysis using earth observation data. Remote Sens. 13 (15), 2850.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 67 (2), 301–320.