



ИССЛЕДОВАТЕЛЬСКАЯ СТАТЬЯ

Редакционная дата: 23 октября 2026 года

Условная прогностическая ценность спутниковой ночной освещённости для экономики Саудовской Аравии

Методологическая работа и воспроизводимый синтетический вычислительный эксперимент

Conditional Predictive Value of Satellite Nighttime Lights for Saudi Economic Nowcasting: A Critical Evidence Review and Reproducible Synthetic Experiment

Тип работы: гибридная рабочая статья (Hybrid Working Paper): критический обзор доказательств в заданных границах (Critical Evidence Review), фактически выполненный синтетический вычислительный эксперимент (Executed Synthetic Computational Experiment) и предлагаемый протокол эмпирической проверки для Саудовской Аравии (Saudi Empirical Validation Protocol). Это не эмпирическое измерение саудовского ВВП, не исчерпывающий систематический обзор (Systematic Review) и не рецензированное исследование.

Статус: работа завершена в заявленных границах; исследование на реальных данных Саудовской Аравии не проводилось.

Аннотация на русском языке

Мы исследуем, дают ли полученные со спутников показатели ночной освещённости (Satellite Nighttime Lights; NTL) независимую прогностическую информацию для оценки текущей экономической активности Саудовской Аравии (Economic Nowcasting) по сравнению с базовой моделью, использующей традиционные экономические данные, доступные на момент выпуска оценки. Исследование различает измерение энергетической яркости (Radiance Measurement), вывод об общественной активности или энергопотреблении и прогноз официального экономического показателя. Мы рассмотрели основные работы по космической экономике, характеристики продуктов NASA Black Marble, опубликованное в 2026 году глобальное исследование фактических колебаний освещённости и документированный NASA сезонный пример из Саудовской Аравии. Предложенная схема строго соблюдает хронологию доступности информации (Point-in-Time Information Set) и учитывает пропуски наблюдений, ошибки измерения и изменение связи между светом и экономикой (Concept Drift).

Мы провели компьютерную симуляцию 180 месяцев по раскрытой синтетической экономической модели с фиксированным зерном 20261026 и искусственной сменой политики или технологии освещения в месяце 116. До изменения средняя абсолютная ошибка (Mean Absolute Error; MAE) базовой экономической модели составляла 0.7116, а после добавления света — 0.5896. После изменения ошибка выросла с 0.7840 до 0.8620, когда свет добавлялся без учёта смены режима освещения; модель с наблюдаемым индикатором изменения получила 0.6370. Эти результаты условны и относятся к одному механизму генерации. Они не означают, что свет действительно улучшает или ухудшает оценку саудовской экономики. Методологический вывод состоит в том, что ценность спутникового показателя следует проверять **как добавку к экономической базовой модели, на реальных временных версиях выпусков и при меняющихся режимах освещения**, отделяя цель прогноза от наблюдаемого физического явления. В заключение предложено саудовское исследование на открытых спутниковых и официальных экономических данных с реальными сезонами и правилами, способными опровергнуть основную гипотезу.

Ключевые слова: оценка текущего состояния (Nowcasting), ночная освещённость (Nighttime Lights), исторические версии данных (Data Vintages), сдвиг распределения (Distribution Shift), VIIRS, дистанционное наблюдение Земли (Earth Observation), ненефтяная экономика Саудовской Аравии (Saudi Non-oil Economy), утечка данных (Data Leakage).

English abstract

This hybrid working paper asks whether satellite-derived nighttime lights (NTL) offer incremental real-time predictive value for Saudi economic nowcasting beyond a credible conventional-data baseline. We separate measured radiance, inferred human activity, and formal economic targets; review core evidence on NTL measurement error, seasonal cultural light signatures, and high-frequency illumination change; and specify data-vintage and product-quality constraints. We implement a reproducible **wholly synthetic** 180-month simulation, with a declared lighting-regime change at synthetic month 116 and a two-month target-release lag. Before this change, adding NTL improves mean absolute error from 0.7116 to 0.5896 index points. After the change, naive NTL inclusion increases MAE from 0.7840 to 0.8620, while a model incorporating an observed lighting-policy indicator yields 0.6370. These conditional model outputs are not estimates of Saudi activity and cannot establish real-world model superiority. We provide a falsifiable Saudi empirical protocol, point-in-time release rules, alternative explanations, and reproducibility files.

Keywords: economic nowcasting; nighttime lights; VIIRS; Saudi Arabia; real-time backtesting; proxy validity; concept drift; measurement error; synthetic experiment.

1. Вопрос, на который изображение не должно отвечать в одиночку

Главный исследовательский вопрос (Research Question): **снижается ли воспроизводимо, вне обучающей выборки, ошибка оценки ещё не опубликованной ненефтяной экономической активности Саудовской Аравии, если к традиционным показателям, действительно доступным на момент выпуска оценки, добавить данные ночного спутникового наблюдения?** Вопрос не требует доказывать, что изображения содержат сведения о человеческой деятельности: в определённых границах это уже известно. Он спрашивает о фактической дополнительной прогностической ценности в конкретных географических, институциональных и временных условиях.

Предлагаемая единица анализа (Unit of Analysis) в эмпирическом исследовании — саудовский регион × месяц или квартал, в зависимости от доступности независимо определённой и опубликованной целевой переменной с последовательной агрегацией. Если доступен лишь квартальный национальный ВВП, единицей анализа становится Саудовская Аравия × квартал. Нельзя искусственно умножать число наблюдений, выдавая сотни тысяч пикселей за сотни тысяч независимых значений выпуска. Пиксели могут использоваться для построения объясняющих переменных, но сами по себе не создают независимых единиц проверки совокупной цели.

Работа касается космоса, наблюдения и достоверности экономического знания. Она сопоставляет три потенциально противоречащих друг другу положения: а) спутниковые сигналы добавляют новые сведения об активности, которых нет в традиционных данных; б) эти сведения дорого дублируют более оперативные или более качественные ряды; в) знак дополнительной ценности меняется вместе с режимом освещения, энергопотреблением или структурой экономики. Наличие трёх возможностей — не теоретическое голосование между ними. Ограничить вывод может только внешняя проверка.

Альтернативная гипотеза (Alternative Hypothesis; H1) утверждает, что средняя ошибка вне обучающей выборки на заранее ограниченных периоде и территории снижается при добавлении спутниковой переменной. Практическая нулевая гипотеза (Operational Null; H0) состоит в том, что различие нестабильно или слишком мало, чтобы оправдать расходы на данные и обработку. Сильнейшее конкурирующее объяснение: результат лишь кажется лучше из-за смены выборок, использования продукта, пересмотренного позднее, расхождений между мерами активности или намеренного выбора неоправданно слабой базовой модели.

2. Обзор доказательств: что на самом деле установили исследования?

В работе Хендерсона (Henderson), Сторигарда (Storeygard) и Вейла (Weil) 2012 года предложено объединять сведения о ночной активности со статистикой дохода в модели ошибок измерения (Measurement Error Model), для которой требуется предположение о связи между ошибкой освещённости и ошибкой национальных счетов. Авторы показали, что разные сигналы способны улучшать оценку активности там, где официальное пространственное измерение затруднено, особенно в приложениях для Африки южнее Сахары. Перенос вывода на современную Саудовскую Аравию требует отдельного обоснования: качество статистической системы, доли нефти и услуг, интенсивность освещения и объём визуально невидимого труда различаются от места к месту [S04](#).

Работа Всемирного банка (World Bank) «Измерение квартального экономического роста из космоса» расширила анализ неопределённости света, используя число пригодных ночных наблюдений, пространственные и временные факторы для понимания ошибки измерения. В одной из регрессионных спецификаций временная связь между изменением освещённости и выпуска была статистически незначимой при фиксированных эффектах страны без контроля временных эффектов, но усиливалась после их добавления. Это не прогностическое сравнение будущих периодов для

Саудовской Аравии. Это эмпирическое предостережение: коэффициент корреляции на международной панели нельзя превращать в устойчивую способность распознавать месячные изменения на местном уровне S05.

В исследовательском материале Всемирного банка о Марокко во время пандемии «COVID-19» (COVID-19) показано, что необработанный свет с примесью временных сигналов не имел значимой прогностической связи с выпуском в проверочный период; после очистки наблюдений связь усилилась, а после введения ограничений в 2020 году ночная освещённость резко снизилась. Эти результаты следует читать в первоисточнике с учётом местных определений очистки, этапов ограничений и отчётности. Они показывают, что способ выделения сигнала меняет вывод, но не дают коэффициента пересчёта, применённого к другой экономике S12.

Роман (Román) и Стоукс (Stokes), напротив, изучили сезонную освещённость городов Саудовской Аравии, Египта, Турции и других стран и обнаружили изменения, связанные с общественными и религиозными обычаями. NASA сообщало, что в отдельных измерениях такие саудовские города, как Эр-Рияд и Джидда, во время Рамадана были примерно на 60–100% ярче. Это **свидетельство чувствительности показателя к социальным привычкам**, а не измерение рамаданских расходов и не доказательство полного разрыва между светом и экономической активностью: оба могут двигаться вместе, но в разной степени S02S01.

В апреле 2026 года Ли (Li) и его соавторы представили высокочастотную глобальную оценку изменений искусственного ночного света (Artificial Light At Night; ALAN) в 2014–2022 годах: на одних территориях общий свет рос, на других сокращался. Среди описанных механизмов — расширение человеческой деятельности, смена технологий освещения, меры по снижению энергопотребления, кризисы и войны. Для нашего вопроса особенно существенен вывод о **непостоянстве предполагаемых коэффициентов связи** между светом и активностью во времени, даже если датчик остаётся физически откалиброванным S08.

Руководство NASA по продуктам Black Marble второй коллекции (Collection 2.0) описывает продукты наблюдения и коррекции и качество оценки, включая ежедневные значения, которые восстановлены из прежних наблюдений, а не получены в новый день. Поэтому следует использовать слои последнего высококачественного наблюдения и флаги качества пикселя, а не считать все заполненные во времени значения (Gar-Filled Values) непрерывным потоком новой информации. Качественные данные объясняют то, **чего не наблюдали**, не меньше, чем измеряют увиденное S06S07.

Граница кабинетного исследования: мы не проводили исчерпывающий систематический обзор по стандарту PRISMA 2020 и не вычисляли единую величину влияния NTL на точность оценки экономики королевства. Были отобраны

основополагающие работы, первичный технический источник, два прикладных исследования и документированный общественный пример, непосредственно относящиеся к вопросу измерения. До даты отсечения нам не удалось найти опубликованную и проверенную нами работу, которая объединяла бы все условия предлагаемого саудовского дизайна. Отсутствие находки не доказывает отсутствия научной работы.

3. От измерения к выводу: трёхуровневая модель

Разделим переменные, которые нельзя смешивать. $Y_{r,t}$ — целевая величина экономической активности в регионе r и периоде t , как её позднее определит официальная статистика; $A_{r,t}$ — потенциально связанная человеческая активность, энергопотребление или общественная практика; $L_{r,t}^{obs}$ — ночная энергетическая яркость, полученная спутниковым измерительным процессом после обработки. Свет может быть связан с экономикой через человеческую деятельность, но он также реагирует на технологию освещения T , календарь C , качество наблюдения Q , пространственные характеристики G и источники света F , не связанные напрямую с целью:

$$L_{r,t}^{obs} = f(A_{r,t}, T_{r,t}, C_{r,t}, G_r, F_{r,t}, Q_{r,t}) + \epsilon_{r,t}^L$$

Из этого не следует, что $A_{r,t}$ полностью определяется выпуском: разные виды деятельности преобразуются в добавленную стоимость с разными коэффициентами. Завод способен увеличить производство без нового освещения, а общество — продлить часы ночной активности без удвоения зарплат. В практической модели требуется проверить:

$$\widehat{Y}_{r,t}^B = g_B(I_{r,t;\tau}^B), \quad \widehat{Y}_{r,t}^S = g_S(I_{r,t;\tau}^B, L_{r,t;\tau}^{available}, Q_{r,t}, Z_{r,t})$$

Здесь τ — момент выпуска оценки; $\mathcal{I}^B$ — набор традиционных переменных, **доступных в момент τ** ; Z — факторы изменения режима освещения, если их можно было наблюдать в то время. В $L^{available}$ не должен попадать продукт, опубликованный после τ , каким бы старым ни был указанный на архивном снимке момент съёмки.

Механизмы ожидаемого смещения

1. **Смещение факторов (Confounding):** сезоны паломничества влияют на свет и некоторые виды расходов, но сила воздействия различается, а привычки освещения могут меняться независимо от расходов.
2. **Смена технологии (Technology Shift):** светодиоды или политика затемнения способны изменить наблюдаемую яркость без равнозначного изменения услуг или дохода.

3. **Несоответствие прокси-переменной (Proxy Mismatch):** снимки освещения иногда хорошо отражают расположение застройки или доступность электричества, но хуже показывают производство финансовых и программных услуг.
4. **Отбор по качеству наблюдения (Observation Selection):** пригодные ночи не всегда распределены случайно; облачность, время и сезоны связаны с географией и жизнью города.
5. **Пространственная зависимость пикселей (Spatial Dependence):** различия миллионов пикселей не дают миллионов независимых экономических наблюдений; поведение, погода и свет распространяются по взаимосвязанным территориям.
6. **Загрязнение временной версии (Vintage Contamination):** использование окончательно пересмотренных значений выпуска или спутниковых продуктов, которых исторически ещё не было, превращает оценку текущего состояния в ретроспективное объяснение.

4. Дизайн выполненного вычислительного эксперимента

4.1 Почему использованы синтетические данные?

Научная цель — проверить **возможность смены знака прогностической пользы**, когда независимый от активности источник света меняется, и строго отделить прежнюю информацию от последующей. Для этой работы не загружался ряд VNP46A2 с покрытием Саудовской Аравии, не собирались реальные исторические версии дат публикации ВВП и не проводилось фактическое сравнение для саудовской экономики. Создавать саудовские числа без этих данных означало бы их выдумать. Поэтому мы публикуем небольшой эксперимент с открытыми предпосылками, позволяющий увидеть, что вычислительный инструмент способен доказать, а что остаётся вне его возможностей.

4.2 Генерация рядов и введение изменения

Мы создали $N = 180$ последовательных месячных периодов — синтетических чисел без дат реальных событий. Случайное зерно зафиксировано как 20261026; использован генератор случайных чисел NumPy (NumPy Random Generator). Истинная скрытая активность y_t порождается сезонным авторегрессионным процессом (Autoregressive Process):

$$y_t = 0.56y_{t-1} + 1.00\epsilon_t^0 + 0.52 \sin(2\pi t/12), \quad y_0 = 0$$

Затем создаётся традиционный показатель, доступный в текущем месяце:

$$x_t = y_t + 1.55\epsilon_t^1$$

До обработки качества создаётся гипотетический спутниковый показатель:

$$\ell_t = 0.95y_t + 0.98\epsilon_t^2 + 0.60 \sin[2\pi(t + 2)/12] - 4 \mathbf{1}_{t \geq 116}$$

Величины ϵ извлекаются из независимых стандартных нормальных распределений в порядке, закреплённом кодом. Синтетический бинарный флаг качества задаётся как $q_t = \mathbf{1}(\epsilon_t^3 > -0.55)$; к наблюдениям низкого качества добавляется шум $0.90\epsilon_t^4$. Этот флаг не является флагом контроля качества NASA (NASA QA) и не отражает реальную долю облачности. Он лишь показывает необходимость различать качество наблюдения и экономическую связь. Значение -4 также не означает, что замена ламп на светодиоды (Light-Emitting Diode; LED) снижает освещённость Саудовской Аравии на такую величину. Это экспериментальный коэффициент, выбранный для создания измеримого структурного изменения.

4.3 Правила сравнения и запрет знания будущего

При выпуске оценки для синтетического месяца t официальные значения активности доступны лишь до $t - 2$. Поэтому обучающие данные заканчиваются в $t - 2$, даже если сегодня в сгенерированном файле уже известно значение $t - 1$. В каждом эксперименте используется метод наименьших квадратов (Ordinary Least Squares; OLS) с расширяющимся окном обучения. Оценивание начинается в месяце $t = 65$ и заканчивается в месяце 178, чтобы недействительная версия с «утечкой» могла для демонстрации экспериментального загрязнения прочитать $t + 1$. Во всех моделях есть константа, x_t и y_{t-2} . Различаются только дополнительные сигналы:

Модель	Входные данные сверх общей базовой модели	Допустимость информации на момент выпуска
Традиционная базовая модель (Baseline)	Нет	Допустима в рамках предпосылок
Простое добавление света (Satellite)	ℓ_t^{obs}	Допустимо в рамках предпосылок
Свет с поправкой на режим (Regime-adjusted)	ℓ_t^{obs} и индикатор изменения $\mathbf{1}(t \geq 116)$	Допустимо, если сведения о смене освещения были доступны своевременно
Намеренная проверка утечки (Leakage)	$0.65\ell_t^{obs} + 0.35\ell_{t+1}^{obs}$	Недопустима ; использует будущее наблюдение

Третья модель не знает величину эффекта -4 в начале изменения. Она оценивает коэффициент новой переменной по мере того, как в обучающем окне накапливается достаточно наблюдений после изменения. Код объявляет источник каждой переменной, её тип и предполагаемую дату доступности. Все результаты цикла записаны в `data/synthetic_monthly_predictions.csv`; по ним рассчитаны показатели следующих разделов.

5. Фактически вычисленные результаты

Средняя абсолютная ошибка (Mean Absolute Error; MAE), корень из средней квадратической ошибки (Root Mean Squared Error; RMSE) и среднее смещение (Mean Bias) рассчитаны как разность между каждой оценкой и истинным синтетическим значением. Доверительные интервалы (Confidence Intervals) и тесты статистической значимости (Statistical Significance Tests) не рассчитывались, поскольку один механизм генерации не позволяет делать выводы о генеральной совокупности за пределами этой симуляции. В таблице приведены вычисленные кодом значения, округлённые до четырёх знаков, в **пунктах синтетического индекса**.

Период	Модель	Число оценок	MAE	RMSE	Смещение
До изменения	Традиционная	51	0.7116	0.9227	+0.0087
До изменения	Простой свет	51	0.5896	0.7266	-0.0525
До изменения	Свет с поправкой на режим	51	0.5896	0.7266	-0.0525
После изменения	Традиционная	63	0.7840	0.9488	-0.0295
После изменения	Простой свет	63	0.8620	1.0987	-0.5693
После изменения	Свет с поправкой на режим	63	0.6370	0.8018	+0.1177
Все периоды	Традиционная	114	0.7516	0.9372	-0.0124
Все периоды	Простой свет	114	0.7402	0.9504	-0.3381
Все периоды	Свет с поправкой на режим	114	0.6158	0.7691	+0.0416
Все периоды	Недействительная модель с утечкой	114	0.7570	0.9700	-0.3132

Результаты показывают три разных механизма. Во-первых, синтетический спутниковый показатель действительно содержит сигнал о y_t , поэтому полезен, пока связь остаётся стабильной. Во-вторых, после снижения освещённости в показании появляется систематическое смещение, хотя уравнение генерации активности не

меняется. Наивная модель из-за этого сильнее занижает активность и работает хуже, несмотря на прежнюю связь. В-третьих, если факт изменения известен и учтён, модель может адаптироваться. Но добавленный бинарный показатель знает о самом событии — а в действительности собрать такие сведения исчерпывающе может быть трудно. Знание о событии не раскрывает величину его эффекта и не гарантирует поправку на все виды спектральных или общественных изменений.

По этим результатам относительное улучшение MAE против базовой модели до изменения равно $(0.7116 - 0.5896)/0.7116$, то есть ошибка снижается примерно на **17,1%**. После изменения простое добавление света увеличивает ошибку на $(0.8620 - 0.7840)/0.7840$, или примерно **9,9%**, тогда как ошибка модели с отмеченным событием снижается на $(0.7840 - 0.6370)/0.7840$, то есть на **18,8%**. Это доли **специально заданного синтетического эксперимента**; переносить их на возможную точность модели в королевстве нельзя.



Результаты двухэтапной синтетической проверки

Рисунок R1. Средняя абсолютная ошибка на двух этапах по данным расчётного файла; за столбцами не стоят саудовские спутниковые или экономические измерения.

Чего этот запуск не доказал? Он не доказал, что освещённость предсказывает саудовский ВВП, что более сложные модели лучше или что замена ламп имеет определённую величину эффекта в реальном мире. Многократная симуляция Монте-Карло (Monte Carlo) с разными случайными зёрнами для оценки разброса не проводилась. Без неё результат нельзя обобщить даже на все ряды, создаваемые теми

же уравнениями. Синтетический традиционный показатель x_t напрямую связан с истинной активностью и может быть значительно лучше одних реально доступных данных и хуже других. Поэтому эксперимент следует понимать как **вычислительное доказательство возможности смены сравнительного вывода**, а не как свидетельство определённой силы этой смены или превосходства какого-либо коммерческого исследовательского продукта.

Недействительная версия, использовавшая часть наблюдения $t + 1$, не превзошла здесь базовую модель. Тем не менее это по-прежнему недопустимый эксперимент, если называть его прогнозом в момент t . Даже численный успех не изменил бы методологической оценки. Будущая информация не становится законной частью модели оттого, что улучшает результат, и не становится менее опасной лишь потому, что в данном вычислительном примере не принесла награды.

6. Как превратить эту проверку в реальное исследование Саудовской Аравии?

До получения национального результата следует разработать саудовский протокол. Предлагаемые единица и рамка могут иметь вид «регион × месяц», если на этом уровне существует опубликованный ряд реальной активности, либо «королевство × квартал», если целью служит реальный ВВП. Нельзя смешивать уровни агрегации или измерять рынок отдельной отрасли по свету всего города. Мы предлагаем три уровня цели по мере доступности: агрегированный показатель производственной активности региона или отрасли; совокупный показатель торговли и перевозок для группы пунктов; национальный экономический показатель из официальных счетов. Для каждой цели нужна отдельная модель и собственное окно выпуска.

6.1 Доступ к данным освещённости

Первичный источник — ночные продукты NASA. Следует сохранять полное название продукта, коллекцию (Collection), датчик и орбитальную платформу, период сцены, покрытие, размер ячейки, каналы и сведения о качестве. В VNP46A2 слой ночной энергетической яркости с поправкой DNB BRDF (DNB BRDF-Corrected NTL) читается вместе с заполненной версией, слоем последнего высококачественного наблюдения и обязательными флагами качества, облаков и снега. Месячный продукт VNP46A3 допустимо использовать после проверки смысла месячного композита и качества числа ночей, а не как сырой средний показатель, равнозначный любому самостоятельно составленному агрегату ежедневного продукта [S06S07](#).

Риски, связанные со временем доступности, принципиальны. Руководство описывает почти оперативный контур (Near-Real-Time; NRT) и архив научного качества (Science-Quality Archive) с разными сроками и процедурами. Исследователь должен отличать быстро выпускаемый продукт с ограниченным применением от научного продукта, который стал доступен позднее. Если оценка вышла утром определённого дня, нельзя использовать версию, попавшую в систему лишь через несколько дней или обработанную повторно через месяцы. Для каждого файла в `data_vintage_manifest` следует хранить дату его доступности, а не только момент наблюдения. Некоторые современные интерфейсы загрузки не раскрывают автоматически дату первого появления исторической версии. В таком случае проверку нужно назвать **псевдооперативной**, а не выдавать за имитацию реального выпуска.

6.2 Выбор подходящей модели ошибки света

До агрегирования пикселей нужно определить правила работы с маской облаков (Cloud Mask), отрицательными и пропущенными значениями, факелами, пожарами, судами, отражённым лунным светом и временными источниками. Отсутствующий пиксель нельзя превращать ни в нулевую освещённость, ни в нулевую экономическую активность: отсутствие наблюдения не равно темноте. Вместе с переменной света можно показывать долю качественных наблюдений для каждого региона и месяца, заранее установить минимально допустимое покрытие и проверить чувствительность к порогу. Изменить порог после просмотра результатов ради более красивого коэффициента означало бы воспользоваться степенями свободы исследователя (Researcher Degrees of Freedom).

Маски городской территории и водных поверхностей следует отделять от административных границ и сравнивать сумму энергетической яркости (Sum of Radiance), среднюю яркость по заданной площади (Mean Radiance) и долю освещённой площади (Lit Area Share). Первый показатель чувствителен к размеру региона, второй — к изменению типа застройки и освещения, третий — к выбору порога. Нельзя говорить об «индексе света» как об одной величине с неизменным смыслом. Поскольку освещённые зоны пересекают границы, нужно проверить чувствительность к краям разбиения и буферной зоне (Boundary Buffer).

6.3 Экономическая базовая модель, которую трудно превзойти

Базовая модель должна использовать сведения, доступные на момент выпуска: опубликованные к этой дате официальные счета с задержкой, показатели торговли и производства, а при надлежащих правах — агрегированные данные о транзакциях или показатели спроса с известным временем публикации. Нельзя включать закрытые

данные организации без разрешения и приписывать их открытому источнику. Следует построить и допустимый наивный ориентир: продолжение последнего опубликованного темпа (Last-Observation Baseline), сезонный фактор, оценённый лишь до конца обучающего периода, а затем богатую экономическую базовую модель на традиционных переменных. Если спутник превзошёл слабую модель, но не превзошёл сильную базовую, это не доказывает дополнительной инвестиционной ценности.

Если квартальная цель строится из месячных переменных, необходимо точно документировать способ агрегации месяцев (Temporal Aggregation) и не давать модели сведений обо всём квартале, когда прошёл только один месяц. В каждом квартале можно выпускать три оценки в заранее заданные даты. Для каждой даты будет свой набор информации, и сравнивать следует лишь выпуски с одинаковым временем. Официальная предварительная цель (Flash GDP) и позднее пересмотренная цель (Revised GDP) — две версии целевой переменной. Нужно решить, какую из них прогнозируют и когда вторую разрешено использовать для проверки после публикации.

6.4 Саудовский календарь входит в определение модели

Рамадан, хадж и праздники включаются в саудовские проверки, поскольку первичные исследования уже выявили связь режимов освещения с этими сезонами в регионе S02S01. Дизайну нужны даты начала и окончания сезона по правильному календарю в каждом году, а не неподвижная григорианская сезонная функция. Но любой сезонный эффект нельзя называть «шумом, который надо удалить». Гостиницы, транспорт и розничная торговля могут действительно влиять на измеряемую цель. Поэтому предлагается сравнить три спецификации: надёжный по качеству необработанный световой показатель; показатель с поправкой на календарные сезоны; показатель со взаимодействием (Interaction Term) сезона и функции региона. Вывод определяется результатом вне обучающей выборки для каждого класса целей.

Официальное сообщение о **1 707 301 паломнике в 2026 году** и сведения о странах прибытия не доказывают величину энергетической яркости Мекки в том сезоне. Они позволяют определить проверочный период и сопоставить официальные данные о движении людей с подходящими областями наблюдения. Если нет однородного пригодного ряда освещённости или наземных показателей хозяйственной деятельности, недостаток проверки нужно объявить прямо. Число фотонов нельзя вывести из числа паломников интуитивно S03.

7. Предлагаемый план измерения и проверки

Мы предлагаем перспективный протокол, который до открытия последнего проверочного окна фиксирует следующее:

Элемент	Предварительное методологическое решение	Что его опровергает или ограничивает
Цель	Официально измеренное изменение ненефтяной активности, регион × месяц, если данные позволяют	Отсутствие надёжного измерения с такой детализацией требует перейти к кварталу/королевству
Спутниковый признак	Освещённость VNP46A2/A3 с коррекцией и масками флагов качества	Пространственные/временные пробелы или задержка публикации, несовместимая с оценкой текущего состояния
Базовая модель	Официально запаздывающая величина + обычные показатели, доступные в то время	Её нельзя заменять более слабой моделью лишь для увеличения разницы
Разделение	Расширяющиеся окна обучения, затем проверка на последующих периодах	Любое случайное смещение месяцев превращает результат в невременную проверку
Основная ошибка	Средняя абсолютная ошибка (MAE) в единицах цели	Сравнение чисел с разными единицами или покрытием
Дополнительные ошибки	Корень из средней квадратической ошибки (RMSE), смещение и границы неопределённости	Одна средняя оценка скрывает региональные ошибки и смену знака результата
Критерий успеха	Заранее определённое устойчивое снижение после учёта стоимости сигнала, с календарным и пространственным сравнением	Успех одного года или выбор лучшей настройки после просмотра результата
Контрольная проверка	Документированные периоды смены освещения, разные сезоны и территории со слабой связью со светом	Недостаток метаданных расширяет неопределённость

Величина улучшения, оправдывающая стоимость продукта, определяется до оценки вместе с реальными пользователями прогноза: она зависит от решения и его чувствительности, а не от предпочтения исследователя. Если доступны реальные цены и конкретные решения, можно построить **кривую ценности информации** (Value of Information Curve), связывающую снижение ошибки со стоимостью обработки, задержкой доступности и чувствительностью решения. В этой работе такая кривая не рассчитана из-за отсутствия операционных финансовых данных.

Устойчивость предлагается проверять на нескольких периодах, включая годы роста цифровой экономики и годы изменения цен на энергию или освещение, не предполагая заранее их влияния на свет. Для неоднородных городов можно построить иерархическую модель (Hierarchical Model) с разными коэффициентами связи по регионам, регуляризацией (Regularization) против подгонки к шуму и по возможности независимыми региональными проверочными выборками. При сборе чувствительных данных следует ограничиться разрешёнными агрегированными показателями, не раскрывая критически важные объекты и рабочие маршруты.

8. Проверки устойчивости и альтернативные объяснения

На полевом этапе следующие опровергающие проверки обязательны для экономического утверждения, но **в этой работе они не проводились на данных Саудовской Аравии:**

1. **Исключение признака (Feature Ablation):** сравнить модель со светом с той же моделью без света при неизменных прочих условиях, затем отдельно проверить признаки качества и календаря.
2. **Устойчивость к заполнению пробелов (Gap-Fill Robustness):** исключить восстановленные ночи или снизить их вес, сравнить результат со всеми наблюдениями и отделить реальный облачный шум от изменения света.
3. **Чувствительность к технологии освещения (Lighting Technology Sensitivity):** по мере доступности включать известные места и даты замены ламп и проверять, дрейфует ли после них коэффициент.
4. **Пространственная отложенная выборка (Spatial Holdout):** не допускать попадания соседних пикселей или связанных районов одновременно в обучение и проверку, когда модель переносится с одного региона на другой.
5. **Воспроизведение исторической версии (Historical Vintage Replay):** воссоздать пакет данных, доступный в момент каждой оценки, без исправленных снимков и экономических показателей, опубликованных позднее.
6. **Чувствительность к цели (Target Sensitivity):** разделить ненефтяную добавленную стоимость, продажи и логистические потоки и определить, для какой цели сигнал информативен, а для какой нет.
7. **Плацебо-период (Placebo Period):** использовать период до предполагаемого изменения освещения и убедиться, что «эффект» не возникает раньше события, если метод оценки предполагает такое условие.
8. **Добавление радиолокационного наблюдения (SAR Addition):** сопоставить агрегированный радиолокационный сигнал изменения застройки или порта,

полученный по спецификациям наблюдения, со световыми и наземными данными, не приравнивая обратное рассеяние радара к финансовой активности.

Уверенность модели нельзя выводить из «уверенности» функции генерации естественного языка. Если модель выдаёт вероятностное распределение, проверяют покрытие интервалов на будущих наблюдениях и различия результата между регионами, а не просто показывают цветовую шкалу. Можно построить систему предупреждения о дрейфе понятия (Concept Drift Alert) на тестах изменения остатков прогноза во времени. Но ей нужна схема сигнализации, допускающая ложные тревоги и различающая изменение цели, датчика и светового поведения.

9. Границы применимости и переход к решению

В работе есть два разных вида доказательств: **опубликованные результаты других авторов** о роли света и механизмах его изменения и **результаты нашей симуляции** с выбранным процессом генерации. Первые не предрешают работу саудовской обсерватории после 2026 года; вторые не описывают географическую или экономическую действительность. Здесь нет реальной саудовской проверки модели, измерения точности спутника над определённым городом или оценки размера неформальной экономики. Соблюдение этих границ сохраняет возможность будущего исследования, не закрывая вопрос технической риторикой.

Есть и проблема причинности. Добавление света может улучшить прогноз официального экономического показателя, не делая свет «причиной» роста или снижения выпуска. Верный прогноз нельзя автоматически использовать для оценки вмешательства, например проекта замены освещения. Если город погасил часть вывесок, а стоимость бизнеса выросла, совпавшее по времени отключение не произвело рост. Если новый город засветился и выпуск изменился, прежде чем назвать причину, нужно изучить инвестиции, производство, население и другие факторы.

Прогностические модели (Prediction) и модели причинной оценки вмешательства (Causal Evaluation) должны храниться и обсуждаться отдельно.

Официальные экономические данные тоже не делятся на «истину» и «ошибку».

Статистическое ведомство может пересмотреть предварительное значение после поступления новых сведений или изменения отраслевого распределения.

Исследователь должен решить, прогнозирует ли он первую опубликованную версию или последнюю пересмотренную. Обе цели уместны для разных задач, но проверять модель с помощью сведений из будущего выпуска и называть результат предварительным прогнозом нельзя. Это особенно видно на примере оперативной оценки за второй квартал 2026 года, которую GASTAT уже опубликовало 30 июля S10. Эта дата не примечание на полях, а часть уравнения доступной информации.

С точки зрения инвестиционной стратегии на этом этапе мы не рекомендуем создавать новый официальный индекс выпуска на основе света. Сначала нужен воспроизводимый испытательный процесс с ясным определением цели, версии входных данных и стоимости каждого обновления — и лишь затем решение о покупке дополнительных данных или привязке оценки к крупному решению. Если несколько периодов покажут, что бесплатных открытых сигналов достаточно, коммерческого смысла переходить на частные снимки только ради более высокого разрешения нет. Если окажется, что они не видят нужную ненефтяную активность, исследование лучше направить к показателям, более близким к рассматриваемой экономической функции.

10. Следующий решающий эксперимент

Мы предлагаем историческую проверку в форме перспективного протокола, напоминающего предварительную регистрацию (Prospective Pre-registered-like Protocol), не утверждая, что нынешний протокол официально зарегистрирован. Она должна охватить десятки реальных выпусков данных по саудовским регионам с достаточно длинной предшествующей историей. До просмотра результатов проверочного периода фиксируются названия продуктов, территории, календарные сезоны и правила качества. Для каждой точки выпуска создаётся хранилище данных со снимком каждого пакета, а при публикации сохраняется официальная целевая величина. Команда, не занимающаяся настройкой признаков, независимо проверяет воспроизводимость файлов оценок и ошибок.

Свидетельством в пользу гипотезы будет повторяющееся улучшение против сильной экономической базовой модели, сохраняющееся при разных режимах освещения, сезонах и регионах, пользу которого можно вычислить до публикации официального показателя. Свидетельством против — исчезновение улучшения после учёта календаря или предотвращения утечки данных, провал в сервисных регионах, разрыв связи после технологической смены или задержка доступности, из-за которой оценка приходит позже официального числа. Тогда область применения следует изменить на городское или сервисное наблюдение, соответствующее природе света, либо прекратить этот путь экономического прогнозирования.

Стратегический вывод не состоит в том, что Саудовской Аравии нужен «спутник, читающий ВВП». Она может развивать **науку диагностики альтернативных сигналов** (Alternative Signal Diagnostics): программы, учитывающие качество датчика, культуру места, историю доступности данных и силу экономического ориентира, а затем уверенно отказывающиеся приписывать изображению ценность, которой оно не даёт. Теоретически такая компетенция способна вырасти в отрасль исследовательского

программного обеспечения для других рынков. Но гипотеза экспорта требует собственных клиентов, независимых испытаний и документированной коммерческой отдачи, а не предварительного заявления.

11. Воспроизводимость, данные и рабочие файлы

Программа [analysis/reproduce.py](#) генерирует все ряды, число наблюдений и результаты моделей без загрузки внешних файлов, используя только NumPy. Она записывает `data/simulation_protocol.json` с зерном, масштабами шума, датой синтетического изменения и описанием недействительной модели. Для каждого периода и каждой модели в `data/synthetic_monthly_predictions.csv` сохраняются оценка, цель, ошибка и флаг временной допустимости, после чего [data/backtest_metrics.csv](#) сводит результаты по трём окнам. График `figure_02_regime_backtest` строится по таблице метрик без ручного повторного ввода значений. `figure_01_global_light_balance` показывает опубликованные величины из исследования Ли (Li) и соавторов [S08](#), а не результат загрузки спутниковых снимков в этой работе.

В материалы входит книга таблиц XLSX с основной таблицей, показателями качества, вычислительными результатами и некоторыми производными связями в формулах, доступных для проверки. Книга не заменяет программу генерации рядов или проверку реальной экономики; её вычислительная версия служит дополнительным ориентиром воспроизводимости. Расчёт можно повторить, установив зависимости из `analysis/requirements.txt` и запустив `python analysis/reproduce.py` внутри каталога статьи. В пакете нет закрытых саудовских данных и реальных значений целевой экономической переменной.

12. Реестр результатов и нерешённых гипотез

Вычисленный результат R-01: в симуляции до изменения освещения добавление света уменьшило ошибку конкретной базовой модели. **Граница:** один механизм генерации, одно зерно и объявленные коэффициенты. **Почему результат не переносится на Саудовскую Аравию:** отсутствуют реальные данные об активности и свете и ряд исторических версий выпусков.

Вычисленный результат R-02: после искусственного снижения света без изменения уравнения активности наивная световая модель стала хуже той же базовой по средней абсолютной ошибке. **Граница:** выбранный тип изменения и воздействующий коэффициент величиной четыре единицы.

Вычисленный результат R-03: модель, которой сообщили о факте изменения, может оценить коэффициент и избежать части ухудшения. **Граница:** наличие переменной события, известной своевременно; её полнота и доступность во всём мире не гарантированы.

Опубликованное наблюдение E-01: реальные глобальные различия в причинах роста и снижения освещённости в 2014–2022 годах по данным Nature S08. **Чего оно не доказывает:** направление роста саудовской экономики или точность оценки с орбиты.

Опубликованное наблюдение E-02: изменение освещённости саудовских городов во время Рамадана в исследовании NASA S02S01. **Чего оно не доказывает:** темп изменения выпуска, совокупного потребления или благосостояния.

Рабочая гипотеза H-01: оценки некоторых видов пространственно локализованной активности в Саудовской Аравии могут улучшиться после добавления качественных световых сигналов с учётом календаря и изменений освещения. **Основание для отказа:** отсутствие устойчивого улучшения против сильной базовой модели на реальных последующих проверочных периодах.

Рабочая гипотеза H-02: спутниковые свидетельства могут быть полезны для наблюдения за пространственными или операционными изменениями, даже если слабо предсказывают национальный ВВП. **Основание для отказа в конкретном применении:** прокси-переменная не подтверждает пригодность для этой цели при сопоставлении с независимыми наблюдениями.

Эти записи не являются вероятностными прогнозами и не присваивают сценариям численные вероятности. Работа не регистрировала датированный прогноз будущего экономического результата. Когда начнётся фактический этап работы обсерватории, для каждого прогноза следует сохранять дату, правило урегулирования и версию доступной до исхода информации.

13. Научные источники и основные документы

- **S04** Henderson, J. Vernon; Storeygard, Adam; Weil, David N. (2012). *Measuring Economic Growth from Outer Space*. American Economic Review 102(2), 994–1028.
<https://doi.org/10.1257/aer.102.2.994>
- **S05** World Bank. (2022). *Measuring Quarterly Economic Growth from Outer Space*. Research paper.
<https://documents1.worldbank.org/curated/en/692011641480491613/pdf/Measuring-Quarterly-Economic-Growth-from-Outer-Space.pdf>

- **S12** World Bank Development Impact. (2021). *Tracking Economic Activity in Response to the COVID-19 Crisis Using Night Lights: The Case of Morocco*.
<https://blogs.worldbank.org/en/developmenttalk/tracking-economic-activity-response-covid-19-crisis-using-night-lights-case-morocco>
- **S02** Román, Miguel O.; Stokes, Eleanor C. (2015). *Holidays in Lights: Tracking Cultural Patterns in Demand for Energy Services*. *Earth's Future* 3, 182–205.
<https://doi.org/10.1002/2014EF000285>
- **S01** NASA Earth Observatory. (2015). *The Lights of Ramadan and Eid al-Fitr*.
<https://science.nasa.gov/earth/earth-observatory/the-lights-of-ramadan-and-eid-al-fitr-84923/>
- **S08** Li, Tian et al. (2026). *Satellite Imagery Reveals Increasing Volatility in Human Night-time Activity*. *Nature* 652, 379–386. <https://doi.org/10.1038/s41586-026-10260-w>
- **S06** NASA Goddard Space Flight Center. *Black Marble Product*.
<https://viirsland.gsfc.nasa.gov/Products/NASA/BlackMarble.html>
- **S07** NASA. (October 2024). *Black Marble User Guide, Collection 2.0*.
https://viirsland.gsfc.nasa.gov/PDF/BlackMarbleUserGuide_Collection2.0.pdf
- **S03** Saudi Press Agency. (2026). *GASTAT: Total Number of Pilgrims for Hajj 2026 Reaches 1,707,301*. <https://www.spa.gov.sa/en/N2599493>
- **S10** General Authority for Statistics. (30 July 2026). *GDP: Flash Estimate Q2/2026*.
<https://www.stats.gov.sa/en/search-publication?category=publication&delta=4&q=الناتج+المحلي+الإجمالي>
- **S11** European Space Agency. *Sentinel-1: Radar Vision for Copernicus*.
https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1
- **S09** Bará, Salvador; Castro-Torres, José J. (2024). *Diverging Evolution of Light Pollution Indicators: Can the Globe at Night and VIIRS-DNB Measurements Be Reconciled?* Author manuscript. <https://arxiv.org/abs/2409.13111>

Этическое и процедурное заявление: текст создан с применением автоматизированных средств под документированной редакционной ответственностью. Реальные исследователи из упомянутых организаций не участвовали в работе; интервью и симуляции реальных потрясений для Саудовской Аравии не проводились; доступа к чувствительным операционным данным не было; внешнего рецензирования не проводилось. Реестр доказательств, ссылки и код предоставлены для проверки воспроизводимости и исправления обнаруженных ошибок.