

Державна служба статистики України

ЗАТВЕРДЖЕНО

Наказ Державної служби
статистики

31 грудня 2014 р. № 420

(зі змінами, внесеними наказом
Державної служби статистики від
19 грудня 2022 р. № 322)

**МЕТОДОЛОГІЧНІ ПОЛОЖЕННЯ
ЩОДО ЗДІЙСНЕННЯ СЕЗОННИХ КОРИГУВАНЬ СТАТИСТИЧНИХ
ПОКАЗНИКІВ**

Київ – 2014

Методологічні положення щодо здійснення сезонних коригувань статистичних показників (далі – Методологічні положення) визначають основні принципи, методи, етапи проведення сезонних коригувань статистичних показників, зокрема оцінку вхідних даних, етапи сезонного корегування, аналіз результатів.

Методологічні положення підготували працівники департаменту статистичної інфраструктури Держстату:

- Товченко А.В. – начальник відділу математичних методів та підтримки метаінформації;
Тихонов Г.В. – головний спеціаліст-економіст відділу математичних методів та підтримки метаінформації,

за участю співробітників Державного підприємства “Науково-дослідний інститут статистичних досліджень”:

- Ященко Л.О. – завідувач відділу методології статистичних досліджень кон’юнктури ринку, к. е. н., с. н. с.
Мотузка О.М. – науковий співробітник відділу методології статистичних досліджень кон’юнктури ринку, к. е. н.

Методологічні положення схвалені Комісією з питань удосконалення методології та звітної документації Держстату (протокол від 25.12.2014 №15).

Методологічні положення затверджені наказом Держстату від 31 грудня 2014 року № 420.

Державна служба статистики України:

- вул. Шота Руставелі 3, Київ, 01601
- телефони: (044) 284-31-32, 289-76-40-
- електронна пошта: office@ukrstat.gov.ua
- веб-сайт: <http://www.ukrstat.gov.ua>

Зміст

I. Загальні положення.....	5
II. Визначення основних термінів і понять.....	7
III. Аналіз вхідних даних	10
1. Підготовка вхідного часового ряду	10
2. Основні компоненти та види декомпозиції часового ряду.....	11
1) Основні компоненти часового ряду	11
2) Види декомпозиції часового ряду	11
3. Ефекти, що впливають на часовий ряд	12
1) Сезонність.....	12
2) Календарні ефекти	13
3) Викиди.....	13
IV. Етапи здійснення сезонного коригування	15
1. Підготовка вхідних даних до проведення сезонних коригувань	15
2. Задання календаря.....	16
3. Відбір специфікацій та регресійних змінних	16
V. Аналіз результатів сезонного коригування	17
1. Діагностика якості.....	17
2. Перегляд сезонно скоригованих даних.....	19
VI. Документування	20
Список використаних джерел.....	21
Додатки	
Додаток 1. Порядок здійснення робіт із запровадження сезонних коригувань основних статистичних показників	22
Додаток 2. Приклад таблиці порівняння результатів сезонного коригування для квартальних даних	23
Додаток 3. Алгоритм сезонного коригування часових рядів у Demetra+.....	24
Додаток 4. Приклад здійснення сезонних коригувань за допомогою програмного забезпечення Demetra+.....	25
Додаток 5. Структура метаданих сезонного коригування	30
Додаток 6. Структура аналітичного звіту	33

Передмова

Згідно зі Стратегією розвитку державної статистики на період до 2017 року, затвердженою розпорядженням Кабінету Міністрів України від 20 березня 2013 р. № 145-р, Держстат здійснює впровадження сезонного коригування основних статистичних показників (далі – сезонне коригування) у статистичну діяльність.

Статистичні дані часто залежать від сезонних коливань та інших календарних ефектів, які можуть заважати виявити відповідні короткострокові й довгострокові зміни в часових рядах і перешкоджають чіткому розумінню економічних явищ.

Основна мета сезонного коригування полягає в тому, щоб нівелювати вплив сезонних коливань і календарних ефектів на динаміку часового ряду та відображення реальних змін, що відбуваються у процесах, які він відображає. Проте зрушення або нерегулярні зміни, що можуть бути пояснені з економічної точки зору (наприклад, наслідки економічної політики, великомасштабних замовлень або страйків тощо), будуть відображені в сезонно скоригованих часових рядах.

Наразі постало питання щодо здійснення сезонних коригувань основних статистичних показників для більш чіткого відображення економічних явищ і співставлення даних на міжнародному рівні.

Методологічні положення забезпечують основу для практичного впровадження сезонного коригування основних статистичних показників органами державної статистики.

I. Загальні положення

Методологічні положення підготовлені відповідно до статті 9 Закону України "Про офіційну статистику".

В основу цих Методологічних положень покладено документи Євростату "Керівництво з сезонних коригувань (2009 року)" [1] та ЄЕК ООН "Практичний посібник з сезонного корегування програмним забезпеченням Demetra+ (2011 року) [2].

Ці Методологічні положення визначають етапи підготовки часових рядів до здійснення сезонних коригувань, проведення власне сезонного коригування та аналізу отриманих результатів, а також етапи впровадження сезонного коригування статистичних показників.

Для стандартизації та уніфікації здійснення сезонного коригування органами державної статистики використовується програмне забезпечення Demetra+, яке рекомендовано в країнах-членах ЄС.

Програмне забезпечення Demetra+ було розроблено Національним банком Бельгії у співпраці з Євростатом. У програмному забезпеченні Demetra+ реалізовано два методи сезонного коригування:

X-12-ARIMA (ARIMA розшифровується як Autoregressive Integrated Moving Average), розроблений Бюро перепису США;

TRAMO-SEATS (TRAMO розшифровується як Time Series Regression with ARIMA Noise, Missing Observations and Outliers, а SEATS — Signal Extraction in ARIMA Time Series), запропонований банком Іспанії.

Хоча методи сезонного коригування X-12-ARIMA та TRAMO-SEATS мають значні відмінності методологічно, в їх основі лежить модель ARIMA.

Алгоритм реалізації X-12-ARIMA представлено на рис. 1 [3].

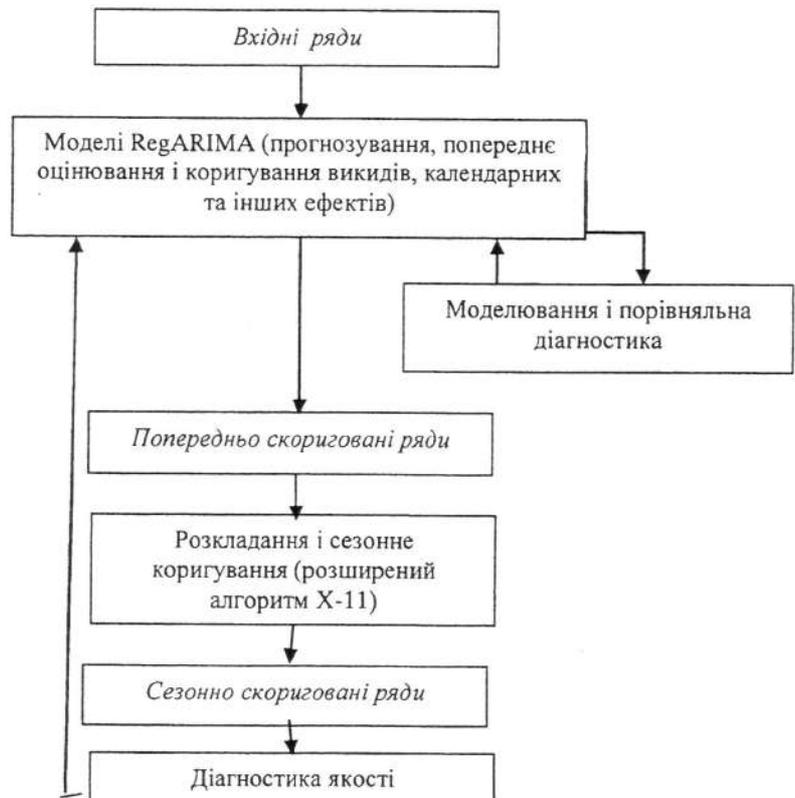


Рис. 1. Алгоритм реалізації X-12-ARIMA

Алгоритм реалізації TRAMO-SEATS наведено на рис. 2.

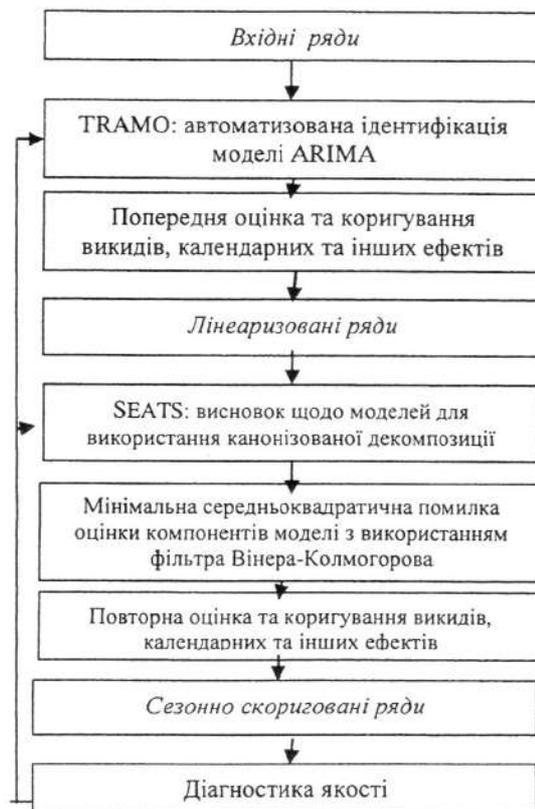


Рис. 2. Алгоритм реалізації TRAMO-SEATS

Сезонні коригування здійснюються для статистичних показників, щодо яких спостерігаються сезонні коливання та вплив календарних ефектів, у результаті чого може приховуватися справжня динаміка ряду (наприклад, ВВП, оборот роздрібною торгівлі, індикатори ділових очікувань тощо).

Роботи із запровадження сезонного коригування статистичних показників здійснюються відповідно до порядку, наведеному в додатку 1 до цих Методологічних положень.

Методологічні положення призначені для використання працівниками органів державної статистики, а також можуть бути корисними для фахівців інших державних органів, наукових організацій та установ, які проводять дослідження в галузі економіки або користуються результатами таких досліджень, інших зацікавлених користувачів.

II. Визначення основних термінів і понять

Терміни, що висвітлюються в цих Методологічних положеннях застосовуються у значеннях наведених у Практичному посібнику із сезонного коригування програмним забезпеченням Demetra +, а також Статистичному словнику та Керівництві із сезонних коригувань:

автокореляція – це лінійна залежність між значеннями різних періодів стаціонарної змінної. Стаціонарна змінна – це ряд, статистичні властивості якого, такі як математичне сподівання, дисперсія і автокореляція, залишаються незмінними з часом [2];

авторегресійний спектр – це спектральний графік, що використовується для попередження про присутність залишкових сезонних ефектів і ефектів операційних днів [2];

адитивні відхилення – це імпульсний викид, який впливає тільки на значення одного спостереження [2];

асиметрія – це міра якості результатів сезонного коригування. Чим менша асиметрія залишків, тим якісніші результати [2];

викиди – аномальні значення ряду [2];

вхідний ряд – це ряд, який має всі характеристики часового ряду без будь-яких коригувань або скорочень. Вхідний ряд означає "сирі дані" і може бути виражений у вигляді абсолютних значень або індексів [2];

декомпозиція – це розділення часового ряду на складові частини (компоненти) за допомогою статистичних методів. Зазвичай сезонне коригування виділяє тренд-циклічну компоненту, сезонну компоненту і нерегулярну компоненту із вхідного ряду [2];

ефекти операційних днів – це повторювані ефекти дня тижня й ефекти високосного року в щомісячних (або щоквартальних) часових рядах. Ефекти операційних днів оцінюються шляхом підрахунку їх пропорції за місяць. Таким чином, ефект операційних днів означає, що рівень вимірюваної діяльності варіюється залежно від дня тижня. Наприклад, обсяг продажів може бути більш високим по п'ятницях, ніж по понеділках [2];

залишки – різниця між вхідними значеннями часового ряду та значеннями, що були скориговані. Аналіз залишків допомагає виявити викиди та непостійність дисперсії часового ряду [2];

залишкова сезонність – це небажана ознака сезонно скоригованого ряду і залишків. Наявність залишкової сезонності є ризиком, який може негативно вплинути на інтерпретацію сезонно скоригованих даних [2];

зсув рівня – це вид викиду, який відноситься до зміни в бік більш високих або низьких значень, з більш-менш постійною зміною рівня часового ряду [2];

календарні ефекти – це будь-які економічні ефекти, пов'язані з календарем, які впливають на часовий ряд. Календарні ефекти видаляються перед декомпозицією ряду. До календарних ефектів, як правило, відносяться: різна кількість робочих днів у певний місяць або період, склад робочих днів, ефект високосного року та перехідні свята, такі як, наприклад, Великдень. Як правило, календарні ефекти включають у себе ефекти робочих днів або ефекти операційних днів [2];

ковзне середнє – це метод для згладжування часового ряду шляхом обчислення середнього значення (з зважуванням або без) для фіксованої кількості послідовних періодів [2];

коригування на операційні/робочі дні – отримання скоригованого ряду, значення якого не залежать від тривалості та складу днів місяця або кварталу (робочих та неробочих днів, операційних та неопераційних днів) [2];

короткочасні зміни – це вид викиду, а саме точковий стрибок, за яким іде гладке повернення до вхідних значень [2];

логарифмічне перетворення – це обчислення логарифма кожного значення в ряду. Логарифм схильний об'єднувати великі значення і збільшувати невеликі значення, які можуть зменшувати проблеми, пов'язані з асиметрією, викидами або коливанням. Це робить ряд більш стаціонарним [2];

модель ARIMA – модель, що використовують для моделювання та прогнозування даних часового ряду. Модель ARIMA містить три види параметрів: авторегресії (p), одиничної різниці (d) та ковзного середнього (q). (p , d , q) – порядок параметрів у моделі. Ряд із сезонністю має два набори цих параметрів: для регулярної складової (p , d , q) та сезонної складової (P , D , Q) [2];

непрямий метод сезонного коригування – сезонно скориговані показники агрегуються в узагальнюючий показник [1];

нерегулярна компонента – це компонента, яка відображає короткострокові коливання, що залишилися в часовому ряді, і є несистематичними і непередбачуваними [2];

одночасне коригування – це одна зі стратегій оновлення сезонно скоригованого часового ряду, яка означає повний перегляд усіх параметрів, задіяних у коригуванні даних, щоразу при появі даних нового спостереження (щомісячно, щоквартально) або якщо дані вхідного ряду були переглянуті [2];

перегляди – визначаються в широкому сенсі як будь-які зміни, внесені до значень часового ряду [2];

перехідні свята – це свята, які відзначаються кожен рік, але їх час святкування змінюється [2];

періодограма – це спектральний графік, що використовується для попередження про присутність залишкових сезонних ефектів і ефектів операційних днів [2];

поточне коригування – це одна зі стратегій оновлення сезонно скоригованого часового ряду, яка означає, що всі параметри, задіяні у коригуванні даних, залишаються без змін та повторно оцінюються в періоди контролю (наприклад, один раз на рік), які були встановлені заздалегідь [2];

прямий метод сезонного коригування – статистичні показники агрегуються в узагальнюючий показник, який потім сезонно коригується [1];

регресійна змінна – це змінна, що при сезонному коригуванні описує вплив викидів, перехідних свят, операційних та робочих днів [2];

сезонна компонента – це компонента часового ряду, яка відображає коливання, що спостерігаються протягом року та повторюються на досить постійній основі з року в рік [2];

сезонне коригування (згладжування сезонних коливань) – це процес оцінки та послідовного видалення з часового ряду ефектів, які є систематичними і пов'язаними з календарем. У математичному моделюванні розрізняють два види сезонних коливань: 1) мультиплікативні, коли розмах коливань пропорційний середньому рівню часового ряду (рівню тренда); адитивні, коли розмах сезонних коливань для відповідних фаз циклу приблизно постійний і не залежить від середнього рівня ряду. У випадку мультиплікативного сезонного ефекту сезонні коливання виключають із ряду шляхом поділу значень ряду на відповідні сезонні коефіцієнти; при адитивному сезонному ефекті сезонні коефіцієнти віднімаються від значень ряду [4]. Методи сезонного коригування – TRAMO-SEATS та X-12-ARIMA. Методи TRAMO-SEATS та X-12-ARIMA мають спільні функції. По-перше, вони спочатку виконують попереднє коригування на ефект робочих або операційних днів і викиди за допомогою регресійної моделі. По-друге, вони виявляють та оцінюють тренд-циклічну, сезонну та нерегулярну компоненти [2];

специфікація – формалізований опис характеристик, що впливають на часовий ряд [2];

стаціонарність – це набір умов, які дозволяють оцінити параметри моделі, властивості якої є стандартними. Іншими словами, стаціонарний часовий ряд легше моделювати і прогнозувати. Стаціонарний часовий ряд повинен мати постійне математичне сподівання, дисперсію і автокореляції з плином часу [2];

стратегія оновлення – варіант оновлення сезонно скоригованого часового ряду у випадку, коли накопичуються нові або переглянуті спостереження. Вибір стратегії оновлення впливає на частоту переглядів і точність сезонно скоригованих даних [2];

тренд-циклічна компонента – це компонента часового ряду, яка включає в себе довгострокові й середньострокові зміни [2];

часовий ряд – сукупність спостережень за змінною протягом часу. Змінна вимірюється в регулярні інтервали часу, наприклад, щомісячно або щоквартально [2];

часткове одночасне коригування – це одна зі стратегій оновлення сезонно скоригованого часового ряду, яка означає, що частина параметрів, задіяних у коригуванні даних, переглядається щоразу при здійсненні сезонного коригування (щомісячно, щоквартально), а всі параметри повторно оцінюються в періоди контролю (наприклад, один раз на рік), які були встановлені заздалегідь [2].

III. Аналіз вхідних даних

Якість сезонних коригувань залежить від якості вхідних даних. Тому необхідно забезпечувати умови, яким повинен відповідати вхідний часовий ряд для того, щоб він був придатний для коригування.

Під час аналізу вхідного часового ряду вивчають такі аспекти:

- підготовка часового ряду;
- компоненти часового ряду;
- ефекти, що впливають на часовий ряд.

1. Підготовка вхідного часового ряду

Якість вхідних даних впливає на якість здійснення сезонного коригування. Необхідно перевірити вхідні дані на їх точність та співставність методів складання протягом часу. Якщо існують певні неузгодженості, їх необхідно виявляти та вирішувати до здійснення коригування.

Сезонне коригування є методом перетворення часового ряду в дані, які легше інтерпретувати. Часовий ряд вимірюють регулярними часовими інтервалами, наприклад, місяць, квартал. Для здійснення сезонного коригування часовий ряд повинен вимірюватися часовими інтервалами до одного року, так як сезонні коливання відбуваються протягом року і мають тенденцію повторюватися із року в рік.

Для здійснення коригування довжина щомісячного ряду повинна бути щонайменше три роки, або 36 спостережень. Для повноцінного сезонного коригування необхідно більш як 5 років, але не більше 10 років. Якщо часовий ряд охоплює період у 10 років, може виникнути нестійкість сезонного коригування серед даних. Якщо часовий ряд виявиться занадто довгим, то інформація про сезонність, яка мала місце багато років тому, може виявитися незначною для поточного періоду часу, особливо якщо при цьому вносилися зміни в концепцію, визначення та методологію.

Якщо ряд не є співставним з точки зору виміру та розрахунку, то краще його скоротити в цілях виявлення сезонної моделі. Якщо виявляються досить суттєві неспівставності в часовому ряді, одним з виходів є одночасне представлення довгого вхідного часового ряду та скороченого сезонно

скоригованого ряду із застереженнями відносно якості даних. Іншою альтернативою є представлення двох окремих сезонно скоригованих рядів: один за останній період та інший за попередній період.

2. Основні компоненти та види декомпозиції часового ряду

1) Основні компоненти часового ряду

На вхідний часовий ряд впливають як випадкові, так систематичні чинники. Він може вміщувати як сезонні ефекти, так і ефекти, пов'язані з календарем.

Сезонна компонента представляє собою коливання, що спостерігаються протягом року та повторюються регулярно із року в рік. Час, направлення й амплітуда компоненти є більш менш стійкими.

Тренд-циклічна компонента містить довгострокові та середньострокові зміни. Тренд відображає довгострокові зміни з плином декількох десятиліть, тобто структурні зміни, і зазвичай являє собою результат приросту кількості населення, інфляції цін, технологічних змін та загального економічного розвитку. Циклічна компонента являє собою досить гладку змінну навколо довгострокового тренду від зростання до падіння економічної активності, що можуть продовжуватися від 3 до 10 років.

Тренд характеризує зміну всього часового ряду і може бути оцінений лише за даними тривалого періоду. Якщо часовий ряд короткий, оцінити тренд важко. Часто тренд оцінюється разом із циклом.

Нерегулярна компонента відображає залишкові короткотермінові коливання, які є несистематичними та непередбачуваними. Нерегулярна компонента залишається після видалення з вхідного ряду сезонної та тренд-циклічної компонент.

При здійсненні сезонного коригування вхідний ряд розподіляється на три основні компоненти (сезонна, тренд-циклічна та нерегулярна) з ціллю виявлення закономірностей часового ряду. Методи TRAMO-SEATS та X-12-ARIMA автоматично розбивають ряд на три компоненти.

2) Види декомпозиції часового ряду

При здійсненні сезонного коригування виявляють компоненти часового ряду. Декомпозиція припускає, що компоненти ряду є незалежними один від одного. Існує два види декомпозиції: адитивна та мультиплікативна. У Demetra+ у методах TRAMO-SEATS та X-12-ARIMA автоматично виявляється необхідний вид декомпозиції.

При адитивній декомпозиції амплітуда сезонних коливань не залежить від рівня тренд-циклічних змін. Адитивна модель означає, що загальна варіація ряду може бути розкладена на адитивні компоненти, тобто часовий ряд

представлений у вигляді суми сезонної, тренд-циклічної та нерегулярної компонент:

$$O_t = T_t + S_t + I_t = A_t + S_t \quad (1)$$

де O_t – вхідний ряд у момент часу t ,
 T_t – тренд-циклічна компонента в момент часу t ,
 S_t – сезонна компонента в момент часу t ,
 I_t – нерегулярна компонента в момент часу t ,
 A_t – сезонно скоригований ряд в момент часу t .

Мультиплікативна декомпозиція означає, що в міру зростання тренда ряду амплітуда сезонних коливань також зростає. Більшість часових рядів виявляють мультиплікативне відношення між компонентами. Якщо амплітуда сезонних коливань зростає при зростанні рівня ряду та навпаки, а ряд не набирає від'ємних значень, можна використовувати мультиплікативну декомпозицію:

$$O_t = T_t \cdot S_t \cdot I_t = A_t \cdot S_t \quad (2)$$

Декомпозиція адитивних компонент застосовується у тих випадках, коли амплітуда сезонних коливань залишається практично незмінною. Мультиплікативна декомпозиція застосовується у випадках, коли амплітуда сезонних коливань рівномірно збільшується з року в рік.

3. Ефекти, що впливають на часовий ряд

1) Сезонність

Сезонність означає закономірність часового ряду, в якому дані підлягають регулярним та передбачуваним змінам, які повторюються із року в рік. Сезонність являє собою періодичні коливання. Наприклад, тенденція роздрібною торгівлі досягнення піку під час новорічних свят. Сезонність дуже часто трапляється в часовому ряді, що відображає економічні зміни.

Використання терміну "сезонність" зумовлене фактом, що сезони здійснюють помітний вплив на економічну і соціальну активність. Сезонність відображає традиційну поведінку, пов'язану з календарем. Сезонні зміни відображаються в сезонній компоненті. Сезонні коливання даних ускладнюють аналіз того, чи являються зміни дійсно підвищенням або спадом активності, чи вони є частиною регулярних змін.

Поправка на робочі дні – це додаткова процедура, яка допомагає виявляти приховану динаміку. Така процедура призначена для отримання ряду, значення якого не залежать від тривалості та складу днів місяця. Перехідні свята також спричиняють проблеми при інтерпретації даних. При коригуванні на перехідні свята перевіряються та виправляються ефекти, що варіюються із року в рік.

Зазвичай установити сезонність можна за допомогою візуального аналізу даних. Співставляючи квартальні або місячні вхідні ряди за декілька років, можна побачити деякі закономірності, що повторюються.

2) Календарні ефекти

Сезонна компонента включає в себе також календарні ефекти, в тому числі перехідні свята. Ці ефекти впливають на рівень ділової активності за місяць або квартал.

Календарні ефекти являють собою будь-які економічні явища, пов'язані з календарем, з різною кількістю операційних днів і перехідних свят протягом місяця або кварталу. Вони видаляються з часового ряду перед здійсненням сезонного коригування та декомпозиції часового ряду. Ефект високосного року, різна кількість днів у місяцях, склад робочих та неробочих днів, а також перехідні свята можуть змінювати рівень ділової активності, що вимірюється в певний момент часу.

Календарними ефектами є варіюючі будні дні та ефект високосного року. Ці ефекти складаються з будніх днів, коливань робочих та неробочих днів, а також перехідних свят. Високосний рік впливає на часовий ряд, оскільки він додає один день кожні чотири роки, і цей день може припасти як на робочий, так і на вихідний день.

Місяці бувають різними за складом робочих та неробочих днів, операційних та неопераційних днів. Рівень активності може бути різний у різні дні. Наприклад, обсяг виробництва може бути нижчим у суботу та неділю, ніж у будні дні. Таким чином, обсяг виробництва може бути більшим у певний місяць, ніж в інший, за рахунок більшої кількості робочих днів у цьому місяці (ефект робочих днів).

Деякі календарні ефекти є сезонними і тому є частиною сезонної компоненти. Наприклад, кількість днів у місяцях є сезонним явищем. Тому календарні ефекти складаються тільки з частини несезонних календарних ефектів.

Державні свята також впливають на кількість робочих днів. Коли свята припадають на вихідні дні, можуть бути додаткові компенсовані вихідні дні. Також слід звертати увагу на перехідні свята, такі як Великдень.

3) Викиди

Часовий ряд може мати деякі екстремальні значення, які значно відхиляються від тренду і виходять за межі очікуваного спектру на основі характерної закономірності часового ряду. Такі екстремальні значення називаються викидами. Деякі несприятливі природні умови або виробничі конфлікти можуть спричинити появу викидів. Зокрема, прийняття нового закону, введення нового типу мита, військові конфлікти можуть зумовлювати їх появу.

Існує три види викидів:

адитивні відхилення – це одноточковий стрибок часового ряду;

короткочасні зміни – це точковий стрибок, за яким іде плавне повернення назад до початкового рівня ряду;

зсув рівня – це стійка зміна рівня ряду.

Адитивні відхилення впливають тільки на значення одного спостереження.

Вони можуть бути зумовлені випадковими явищами (страйками, негодами, заздалегідь оголошеними підвищенням цін тощо).

При короткочасних змінах після надзвичайно низьких або високих значень спостереження йде поступове зниження або підвищення до тих пір, поки часовий ряд не повернеться до початкового рівня. Короткочасні зміни можуть виникати, наприклад, через відхилення у середньомісячних умовах погоди. Якщо взимку погода раптово стає дуже холодною, то зазвичай споживання енергії зростає. Коли погода поступово повертається до середнього рівня, споживання енергії також має повернутися до середнього значення.

Зсув рівня є зміною значень на більш високий або низький рівень, починаючи з певного періоду, і означає, що рівень часового ряду зазнає стійкої зміни. Зсув рівня може бути спричинений різкими економічними спадами або зростаннями. Наприклад, якщо заробітна плата в деяких професіях підвищується, рівень такого часового ряду є постійно високим, проте сезонна закономірність не змінюється.

Причиною зсуву рівня може бути також зміна в методології. Тому необхідно підтримувати співставність ряду, уникаючи зсуву рівня через зміну в методології.

Викиди за останній місяць або квартал важко виявити. Поки не отримано даних щодо наступних спостережень, дуже важко розрізнити зсув рівня або адитивний викид. Щоб виявити причину виникнення викиду, проводять детальний аналіз часового ряду.

Велика кількість викидів у часовому ряді може значно ускладнити виявлення сезонної моделі. Тому під час підготовки статистичних даних необхідно виявити усі викиди та виправити можливі помилки перед тим, як здійснювати сезонні коригування.

Відсутні дані можуть спричинити перешкоди. Часовий ряд із великою кількістю відсутніх даних може викликати значні проблеми оцінки та коригування. Зважаючи на це, необхідно виявити та виправити відсутні дані, використовуючи альтернативні джерела даних або застосовуючи статистичні методи при відсутності вхідних даних.

У методів TRAMO-SEATS і X-12-ARIMA, що реалізовані у Demetra+, є автоматична процедура з виявлення викидів і внесення поправки на їх вплив. У програмі Demetra+ позначаються: адитивні відхилення AO (additive outlier), короткочасні зміни TS (transitory change) і зсуви рівня LS (level shift) [2].

Викидами особливо важко управляти в кінці ряду, коли їх важко відрізнити від поворотної точки.

Часовий ряд, де занадто багато викидів (більше 10%), створить проблеми з розрахунками. Тому число викидів повинно бути відносно невеликим та не надмірно сконцентрованим навколо одного періоду року [2]. Скорочення відрізка часу може допомогти для гарного моделювання викидів.

IV. Етапи здійснення сезонного коригування

Виконання сезонного коригування в практичному сенсі здійснюється пошуковим методом. Універсальний підхід цього методу полягає в тому, щоб почати з певної специфікації для отримання перших результатів та надалі змінювати певні параметри, а результати порівнювати між собою. Приклад таблиці порівняння результатів наведено в додатку 2.

Demetra+ дозволяє здійснювати сезонне коригування пошуковим методом. У кожному циклі коригування можна змінювати модель та будь-яку специфікацію і одразу побачити результати. Рекомендується перевірити всі опції та їх взаємозв'язок для того, щоб обрати найкращу модель. Загальний алгоритм сезонного коригування часових рядів наведено в додатку 3.

1. Підготовка вхідних даних до проведення сезонних коригувань

Відповідно до пункту 1.4 цих Методологічних положень сезонні коригування статистичних показників органи державної статистики здійснюють за допомогою програмного забезпечення Demetra+ (далі – Demetra+) як такого, що рекомендований ЄЕК ООН та Євростатом.

Demetra+ підтримує різні формати даних, такі як Excel, SDMX, Tramo/Seats, xml, ODBC та інші, але має специфічні вимоги до побудови вхідних даних.

Так, для роботи з форматом Excel файл із вхідними даними повинен відповідати таким вимогам:

- електронна таблиця повинна бути побудована вертикально;
- крайня верхня ліва клітинка A1 повинна бути пустою;
- у першому рядку таблиці зазначаються назви рядів;
- у першому стовпчику наводяться дати у такому форматі: мм-рррр, де мм – місяць, рррр – рік;
- часові ряди представлені у вигляді чисел;
- пусті клітини мають означати відсутні значення (за виключенням початку та кінця ряду).

Вигляд таблиці з вхідними даними у форматі Excel для здійснення сезонного коригування в Demetra+ наведений у пункті 1 додатку 4 до цих Методологічних положень.

2. Задання календаря

Задання календаря передбачає встановлення дат усіх національних свят та періоди їх дій, включаючи дати перехідних свят за кожен рік, що включений до даних, що коригуються.

Методи TRAMO-SEATS і X-12-ARIMA мають вбудовані процедури для виявлення і виправлення ряду з урахуванням календарного ефекту. Метою календарного коригування є більш точна оцінка сезонної компоненти та покращення якості сезонно скоригованого ряду.

У програмному забезпеченні Demetra+ можна скоригувати дані на ефект операційних або робочих днів, ефект високосного року, Великодня та інших святкових днів [2].

Згідно з Кодексом законів про працю в Україні відмічають такі святкові дні (встановлюються як фіксовані свята у програмному забезпеченні Demetra+), що можуть коригуватися відповідно до чинного законодавства:

- 1 січня – Новий рік;
- 7 січня – Різдво Христове;
- 8 березня – Міжнародний жіночий день;
- 1 і 2 травня – дні міжнародної солідарності трудящих;
- 9 травня – День Перемоги;
- 28 червня – День Конституції України;
- 24 серпня – День Незалежності України.

Вихідні також у дні релігійних свят – на Великдень та Трійцю (встановлюються як перехідні свята у програмному забезпеченні Demetra+).

Деякі ряди можуть не піддаватися впливу ефекту операційних днів, наприклад квартальні дані. А от перехідне свято Великдень може впливати на квартальні дані.

3. Відбір специфікацій та регресійних змінних

Після задання календаря необхідно обрати певну специфікацію.

У Demetra+ методи TRAMO/SEATS та X-12-ARIMA автоматично створюють регресійні змінні операційних днів, робочих днів, високосного року та Великодня в залежності від обраних специфікацій.

Demetra+ пропонує заздалегідь задані регресійні змінні для календарних ефектів:

- шість змінних для перевірки на ефекти операційних днів;
- одну змінну для перевірки будніх проти вихідних днів;
- змінну для високосного року;
- змінну для Великодня.

Можна обрати специфікацію, яка вже налаштована за замовчуванням, або налаштувати власну. Demetra+ пропонує набір стандартних специфікацій під різні варіанти параметрів. У правому вікні "Workspace" знаходиться перелік специфікацій, які можуть бути використані для обробки даних. Наприклад, при

обранні методу X-12 у програмі вже встановлено 6 специфікацій цього методу, але за потреби специфікації з власними налаштуваннями можна створити самостійно. Аналіз розпочинають з однієї з специфікацій RSA1c (логарифм/рівень, виявлення викидів), RSA2c (логарифм/рівень, робочі дні, Великдень, виявлення викидів), RSA3c (логарифм/рівень, виявлення викидів, автоматичне виявлення моделі), RSA4c (логарифм/рівень, робочі дні, Великдень, виявлення викидів, автоматичне виявлення моделі) та RSA5c (логарифм/рівень, операційні дні, Великдень, виявлення викидів, автоматичне виявлення моделі).

Рекомендовано розпочинати зі специфікації RSA4c для квартальних даних або RSA5c для місячних даних (дивись пункти 4 та 5 додатка 4 до цих Методологічних положень).

Обрана специфікація дозволяє у Demetra+ здійснювати перевірку на операційні дні, робочі дні, високосного року та інші. За допомогою тестування різних специфікацій можна проаналізувати вплив тих чи інших регресійних змінних і обрати найкращий варіант.

V. Аналіз результатів сезонного коригування

Ціль аналізу є виявлення значущих недоліків у результатах сезонного коригування для уникнення неправильної інтерпретації даних про розвиток економіки.

Сезонне коригування є складною статистичною процедурою, якій необхідний точний контроль і аналіз отриманих результатів. Для забезпечення прийнятної якості сезонно скориговані дані повинні бути ретельно перевірені.

Перевірку сезонно скоригованих даних виконують за допомогою декількох графічних, описових, непараметричних і параметричних критеріїв, докладний набір яких пропонує Demetra+.

Результат здійснення сезонного коригування Demetra+ представляє статистичні та математичні властивості певної моделі та її компонент. Особливу увагу приділяють таким характеристикам якості сезонно скоригованого ряду:

- відсутність залишкової сезонності;
- відсутність залишкових календарних ефектів;
- відсутність надмірного коригування сезонних і календарних ефектів;
- стійкість сезонної компоненти.

1. Діагностика якості

Діагностика якості передбачає отримання детальної інформації щодо процедури сезонного коригування.

Demetra+ автоматично виявляє якість результатів та висвітлює кожен із оцінених характеристик відповідним кольором та написом щодо якості.

Наведемо критерії якості сезонного коригування та інтерпретацію підсумку діагностики (від "невизначеного" до "добре") на нижченаведеному прикладі:

Summary (загальна оцінка якості)

Good

basicchecks (основна перевірка)

definition: Good (0,000) (чим менше за 0,05, тим краще)

annual totals: Good (0,000) (чим менше за 0,05, тим краще)

visualspectralanalysis (візуальний спектральний аналіз)

spectral seas peaks: Good

spectral td peaks: Good

regarimaresiduals (залишки)

normality: Good (0,834) (чим більше за 0,051, тим краще)

independence: Good (0,629) (чим більше за 0,051, тим краще)

spectral td peaks: Good (0,211) (чим більше за 0,051, тим краще)

spectral seas peaks: Good (0,457) (чим більше за 0,051, тим краще)

residualseasonality (залишкова сезонність)

on sa: Good (0,775) (чим більше за 0,051, тим краще)

on sa (last 3 years): Good (0,981) (чим більше за 0,051, тим краще)

onirregular: Good (1,000) (чим більше за 0,051, тим краще)

outliers (викиди)

number of outliers: Good (0,020) (чим менше за 0,05, тим краще)

Таблиця 1

Інтерпретація підсумку діагностики (від "невизначеного" до "добре")

Оцінка	Значення
Невідоме (Undefined)	Якість невідома: необроблений тест, незначущий тест, помилка в розрахунку тесту
Помилка (Error)	Результати містять помилку. Необхідно відхилити обробку (наприклад, вона містить викривлені значення або не виконані деякі числові обмеження)
Серйозна помилка (Severe)	У результатах логічні помилки відсутні, але їх не можна приймати через деякі статистичні причини
Погане (Bad)	Якість результатів погана за яким-небудь конкретним критерієм, але реальної помилки немає, а результати можна використовувати
Невизначене (Uncertain)	Результат перевірки невизначений
Добре (Good)	Результат тесту добрий

Demetra+ розділяє діагностику на п'ять основних частин: наявність сезонності, спектральні графіки, історія переглядів, ковзні середні та аналіз стабільності.

Перша частина діагностики розглядає результати тестів на наявність сезонності у вхідному часовому ряді та залишках. Ряд повинен містити сезонну закономірність, в іншому випадку немає необхідності здійснювати сезонне коригування. Серед тестів на наявність сезонності є тест Фрідмана, тест

Краскела-Уолліса, тест на наявність сезонності, що приймає стабільний характер, тест на залишкову сезонність і комбінований тест на наявність сезонності. Якщо р-значення в тестах менше за 0,05, то ряд має певну сезонність і можна здійснювати сезонне коригування.

Друга частина діагностики розглядає спектральні графіки, а саме періодограму й авторегресійний спектр. Ці спектральні графіки перевіряють залишки, нерегулярну компоненту та сезонно скоригований ряд на залишкові сезонні або календарні ефекти. Якщо присутня залишкова сезонність, необхідно змінювати модель, специфікацію моделі, регресійні змінні або інтервал часу, що використовується для моделювання. Demetra+ містить інформацію про залишки, тобто частину даних, які моделювання не пояснює. Залишки повинні приблизно мати нормальний розподіл, бути випадковими і незалежними. Залишки являють собою стаціонарну змінну, тобто не повинні мати лінійної структури. Для цього залишки перевіряються на відсутність автокореляції. Наприклад, за допомогою тестів Льюнг-Бокса і Бокса-Пірса аналізують наявність сезонності в залишках. За цими тестами відхилена нульова гіпотеза свідчить про наявність автокореляції в залишках. Це вказує на те, що залишається лінійна, небажана структура в рядах [2].

Третя частина діагностики розглядає історію переглядів часового ряду. Тут аналізуються перегляди, які визивають додаткові спостереження в кінці ряду. Розглядаються графіки як для сезонно скоригованого ряду, так і для тренд-циклічної компоненти. Чим ближче точки початкового спостереження до кривої, що базується на інших доступних спостереженнях, тим вища якість.

Четверта частина діагностики аналізує ковзні середні. Цей метод діагностики є дуже корисним для рядів зі змінною сезонністю або великою кількістю викидів. Отримані результати є стабільними, якщо додавання або видалення спостережень не викликає великих змін результату.

П'ята частина діагностики виявляє стабільність моделі, що розраховує параметри ARIMA і коефіцієнти регресійних змінних у різні періоди часу. При оцінюванні стабільності моделі можна вважати нестабільними значення, що перевищують 3% від порогового значення. Чим далі розташовані точки від осі абсцис на графіку, тим менша стабільна модель.

2. Перегляд сезонно скоригованих даних

Важливим питанням при здійсненні сезонних коригувань є визначення стратегії оновлення. Оскільки сезонні коригування базуються на оцінках та припущеннях, вони підлягають перегляду. Оновлення та накопичення даних забезпечують кращі оцінки сезонних закономірностей.

У Demetra+ визначено такі стратегії оновлення: поточне коригування, часткове одночасне коригування і одночасне коригуванням. Найбільш поширеною стратегією є часткове одночасне коригування.

При здійсненні перегляду має бути визначена стратегія, визначений графік переглядів та має здійснюватися відповідне документування перегляду.

VI. Документування

Перелік статистичних показників, які піддають сезонним коригуванням, методи, якими здійснюються коригування, а також необхідність і частота переглядів скоригованих рядів повинні бути наведені за стандартною уніфікованою структурою, що відповідає стандартному опису метаданих здійснення сезонних коригувань ЄСС. Уніфікована структура наведена в додатку 5 та має заповнюватися при кожному раунді здійснення сезонного коригування.

Директор департаменту статистичної
інфраструктури Держстату



Ю. М. Остапчук