

УДК 528.06:551.521

Г. С. Курбасова, О. Є. Вольвач

Лабораторія радіоастрономії Науково-дослідного інституту
«Кримська астрофізична обсерваторія», Кацивелі, Крим

ВЕЙВЛЕТ-АНАЛІЗ НАЗЕМНИХ І КОСМІЧНИХ ВИМІРІВ ЛОКАЛЬНОЇ ІНСОЛЯЦІЇ

Показано можливість застосування вейвлет-аналізу для виділення сигналу із шуму і статистичного аналізу шуму в тривалих послідовностях локальних наземних і космічних спостережень на основі вимірювань інсоляції супутниками НАСА та наземних вимірювань у парку «Нікітський сад» (Крим, широта 44.5°, довгота 34.2°). Обґрунтовано вибір методу аналізу, виду вейвлетів і їхніх параметрів. Проведено аналіз результатів оцінок шуму космічних і наземних спостережень.

ВСТУП

Періодичні зміни локальних кліматичних характеристик пов'язані, в першу чергу, з детермінованими процесами: обертанням Землі навколо осі та її руху на орбіті [1, 7], періодичними процесами на Сонці [2]. Ці зміни передбачувані. Однак на періодичні коливання в атмосфері впливають випадкові процеси, обумовлюючи нерегулярну в часі шумову складову у рядах кліматичних параметрів. Аналіз шуму на тривалих часових інтервалах необхідний для з'ясування причин зміни і корекцій кліматичних характеристик.

Однією з важливих кліматичних характеристик є сонячна радіація — головний ресурс природного середовища. Вона практично визначає умови життя на Землі, є визначальним чинником у формуванні та підтримці стабільності енергетики системи Земля — атмосфера.

Опромінення поверхонь сонячним світлом (сонячною радіацією) прийнято називати інсоляцією. Середню температуру півкулі підтримує рівномірний приплив тепла до земної поверхні від сонячної радіації. Назад у космос відбивається приблизно 30 % цієї енергії. У той же час

через відбивання, розсіювання і поглинання її атмосферними газами і аерозолями лише 47 % всієї енергії досягає поверхні Землі [<http://www.ecomuseum.kz/dieret/solar/solar.html>].

Зміни сонячної активності і місячно-сонячні припливи Землі формують взаємодії суші, моря і атмосфери, які викликають коливання глобальних і регіональних кліматичних характеристик [1, 2]. Однією з найважливіших величин, від якої суттєво залежать всі земні процеси, є швидкість обертання Землі. Земні процеси, у свою чергу, впливають один на одного і можуть змінювати результати зовнішніх впливів і швидкість обертання Землі [7].

Вчені вважають, що сучасне потепління (1920—2035 рр.) — це результат складання теплих фаз кліматичних коливань з періодами 230, 500 і 1000 років [3, 4]. За прогнозами з 2035 р. температура і природні умови поступово стануть нагадувати холодний початок ХХ століття: закриться Північний морський шлях, зміняться умови видобутку нафти на шельфі і у північних районах, впаде середня врожайність. Зменшення сезонних температур для регіонів узбереж і шельфів відбувається у 5—10 разів швидше, ніж зменшення річних температур. Тому вже зараз необхідно створювати додаткові запаси енергоносіїв, розробляти технології максимального використання і збереження сонячної енергії. Для

підтримки наукових і технічних розробок з проблем відновлюваної енергетики в NASA розроблений і успішно функціонує в даний час проект POWER (Prediction Of Worldwide Energy Resource [<http://power.larc.nasa.gov>]). У рамках цього проекту створено базу даних параметрів наземної метеорології і сонячної енергії SSE (Surface meteorology and Solar Energy [<https://eosweb.larc.nasa.gov/sse/>]).

Ця база складається з понад 200 супутникових метеорологічних параметрів і параметрів сонячної енергії. 22-річний кліматичний ряд (липень 1983 р. — червень 2005 р.) забезпечує глобальне охоплення по поверхні Землі (на сітці один градус широти на один градус довготи) безлічі SSE-даних. Більшість наземних станцій вимірювання розташовані поблизу населених районів, які можуть мати природний або техногенний вплив на місцевий клімат. Тому SSE-дані вважаються найбільш точними для прогнозу локального клімату та попереднього техніко-економічного обґрунтування проектів із відновлюваних джерел енергії.

Для точнішого прогнозу необхідний аналіз розбіжностей наземних і космічних вимірювань на тривалих інтервалах часу. Необхідно не тільки встановити, що такі розбіжності мають місце, але і провести аналіз їхніх змін в часі залежно від зовнішніх і внутрішніх факторів. В цьому випадку результати аналізу залежать від обраного методу.

При виборі методу аналізу спостережень інсоляції нами була прийнята до уваги неоднорідність у часі кліматичних характеристик. Однак традиційне перетворення Фур'є, основане на тому що функція перетворення покриває всю часову вісь, не дає уявлення про локалізацію процесу в часі. Навпаки, функція одновимірного вейвлет-перетворення добре локалізована і в часі, і по частоті. Вейвлет-аналіз шуму в неоднорідних рядах спостережень дає інформацію про зміни в навколишньому природному середовищі і космічному просторі на різних інтервалах часу. Результати проведеного авторами вейвлет-аналізу вимірів сонячної інсоляції SSE і наземних вимірів в ЦГМ «Нікітський сад» обговорюються в даній роботі.

ВИХІДНІ ДАНІ, ПЕРЕВАГИ ВЕЙВЛЕТ-АНАЛІЗУ

Для аналізу використовувалися щоденні усереднені дані:

- послідовності супутникових вимірів SSE: swv_x — середня інсоляція на горизонтальній поверхні (в одиницях $\text{kB} \cdot \text{г} \cdot \text{м}^{-2} \cdot \text{доба}^{-1}$); послідовність даних являє собою часовий ряд з інтервалом між відліками 1 доба (8036 відліків з липня 1983 р. по червень 2005 р.). Джерело даних: <http://eosweb.larc.nasa.gov/sse/>;

- послідовності вимірів у парку «Нікітський сад», Крим (період 1983—2011 рр.): sr — сумарна інсоляція за добу ($\text{МДж} \cdot \text{м}^{-2} \cdot \text{доба}^{-1}$). Для порівняльного аналізу ці дані перераховано в одиниці $\text{kВт} \cdot \text{год} \cdot \text{м}^{-2} \cdot \text{доба}^{-1}$. Місце розташування парку «Нікітський сад»: широта 44.5° , довгота 34.2° .

Насамперед будемо виходити з відсутності інформації про статистичні і спектральні характеристики вихідних даних. Для встановлення спектральної структури послідовностей вихідних даних застосуємо традиційний метод Фур'є-аналізу, в основі якого лежить розкладання сигналу на синусоїдальні хвилі різних частот. Синусоїди не мають обмеженого терміну — вони простягаються від мінус до плюс нескінченності. Тому Фур'є-аналіз дає надійну інформацію про періодичні складові сигналу, які не схильні до локальних змін.

Одна з головних переваг вейвлетів — це можливість виконувати локальний аналіз, тобто аналіз локалізованої області більшого сигналу: вейвлет-аналіз дозволяє використовувати тривалі проміжки часу, де ми хочемо отримати точнішу низькочастотну інформацію і короткі проміжки для виявлення високочастотної інформації.

Локальний вейвлет-аналіз здатний виявити аспекти даних, які іншими методами аналізу сигналів не виявляються: тенденції, розбивку точок, розривів у вищих похідних і самоподібності. У місцях неоднорідностей даних синусоїди Фур'є-перетворення гладкі і передбачувані, а сплески, як правило, нерегулярні і асиметричні.

На відміну від звичайних методів, вейвлет-розкладання визначає сімейство ієрархічно організованих розкладів. Вибір відповідного рівня

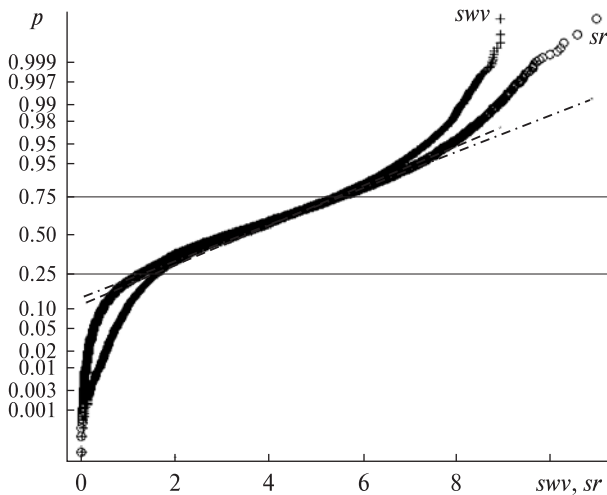


Рис. 1. Графіки нормального розподілу ймовірностей в даних *sr* і *swv*

ієрархії залежатиме від характеристик даних і постановки конкретної задачі: перший рівень розкладання може бути використаний для усунення більшої частини шуму, викиди і неоднорідності частіше видно на вищому рівні розкладу.

Вейвлет-аналіз є лінійний метод: вейвлет-коефіцієнти лінійної комбінації двох сигналів $\alpha s^{(1)} + \beta s^{(2)}$ дорівнює лінійній комбінації їх вейвлет-коефіцієнтів $\alpha C_{j,k}^{(1)} + \beta C_{j,k}^{(2)}$, де s — послідовність відліків в даних (сигнал); індекс j — масштаб, k — зсув по осі часу. Те ж саме справедливо для відповідних наближень. Математичний опис одновимірного вейвлета має вигляд

$$\frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right),$$

де $a > 0$, $b \in R$, R — область визначення функції сигналу. Коефіцієнти C обчислюються в першу чергу шляхом згортання вейвлет-функції Ψ і сигналу s з використанням звичайних формул. Шляхом аналізу вейвлета, так, якщо b це був сигнал, відбувається самовідтворення функції $U(a, b)$:

$$U(a, b) = \int \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right) \Psi(x) dx, \quad (1)$$

Етапи відновлення збуреного сигналу за допомогою вейвлет-перетворення описані в численній літературі, алгоритми реалізовані у безлічі машинних програм.

Вибір того чи іншого класу вейвлетів залежить від статистичної та спектральної структури даних. Тому вейвлет-аналіз почнемо з оцінки ймовірності та спектральних моделей даних.

ПЕРЕВІРКА ГІПОТЕЗИ ПРО НОРМАЛЬНИЙ РОЗПОДІЛ ТА СПЕКТРАЛЬНИЙ СКЛАД ДАНИХ

Будемо виходити з припущення, що дані *sr* і *swv* вийшли з сукупностей з нормальним розподілом, тому нормальні ймовірнісні графіки можуть забезпечити деяку гарантію того, що припущення виправдовується, або видати попередження про проблеми з припущенням. Аналіз статистичної структури даних почнемо з побудови нормальних ймовірнісних графіків та перевірки гіпотези на нормальність за допомогою обчислення тестів.

На рис. 1 емпіричні ймовірності позначені хрестиками (дані *swv*) і кружками (дані *sr*). Суцільні лінії проведено через 25-й і 75-й процентілі в даних. По осі ординат значення ймовірностей від нуля до одиниці, масштаб нелінійний: відстані між поділами на цій осі відповідають відстані між квантилями нормального розподілу.

Для кількісної перевірки гіпотези нормального розподілу ми використали критерій нормальності *lillietest* [10], який не передбачає попередніх знань про статистику даних. Відповідно до цього тесту гіпотеза про нормальний розподіл відкидається на рівні значимості 5%. Логічний нуль означає спробу відкинути нульову гіпотезу при тому, що в даних має місто нормальний розподіл. *Lillietest* для обох послідовностей *sr* і *swv* видає значення одиницю, що вказує на відхилення від нормального розподілу.

Через те що справжня модель даних нам не відома, будемо виходити з базової моделі для сигналу з шумом, яка має в основному такий вигляд:

$$S(n) = f(n) + \sigma e(n), \quad (2)$$

де n — рівновіддалені відліки часу. У простій моделі ми припускаємо, що $e(n)$ є гауссовим білим шумом $N(0, 1)$, а рівень шуму σ повинен дорівнювати одиниці. Складова $f(n)$ моделі (2) — сигнал, спектральна структура якого має бути встановлена.

З метою встановлення спектральної структури даних sr і swv були обчислені графіки періодограм за допомогою швидкого перетворення Фур'є (рис. 2). Видно, що в даних є найбільш значуще коливання з періодом 1 рік. У роботі [5] були зіставлені довгоперіодичні коливання в даних sr і swv методом двоканального спектрального аналізу авторегресії, та встановлено когерентність основних (рис. 2) коливань з періодом 1 рік: квадрат модуля когерентності більше 90 %.

Таким чином, результати попереднього аналізу такі: 1) гіпотеза про приналежність даних sr і swv до сукупностей з нормальним розподілом не підтверджується; 2) у спектральному складі даних є одне найбільш значне коливання.

ВИБІР ТИПУ ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ, СІМЕЙСТВО ВЕЙВЛЕТІВ ДОБЕШИ

Послідовності даних sr і swv — дискретні, тому не всі значення вейвлет-розкладання необхідні для розподілу вихідного сигналу на складові: відновлений сигнал і шум. З двох типів вейвлет-перетворення (безперервного і дискретного) в цьому випадку доречно застосувати дискретне вейвлет-перетворення. При цьому вейвлет повинен задовольняти деяким умовам допустимості [6, с. 105]. Ці умови задовольняють всі вейвлети сімейства Добеши, показано на рис. 3.

Інгрід Добеши винайшла те, що називається фінітними ортонормованими вейвлетами, тим самим роблячи дискретний вейвлет-аналіз практичним. На рис. 3 наведені графіки сімейства вей-

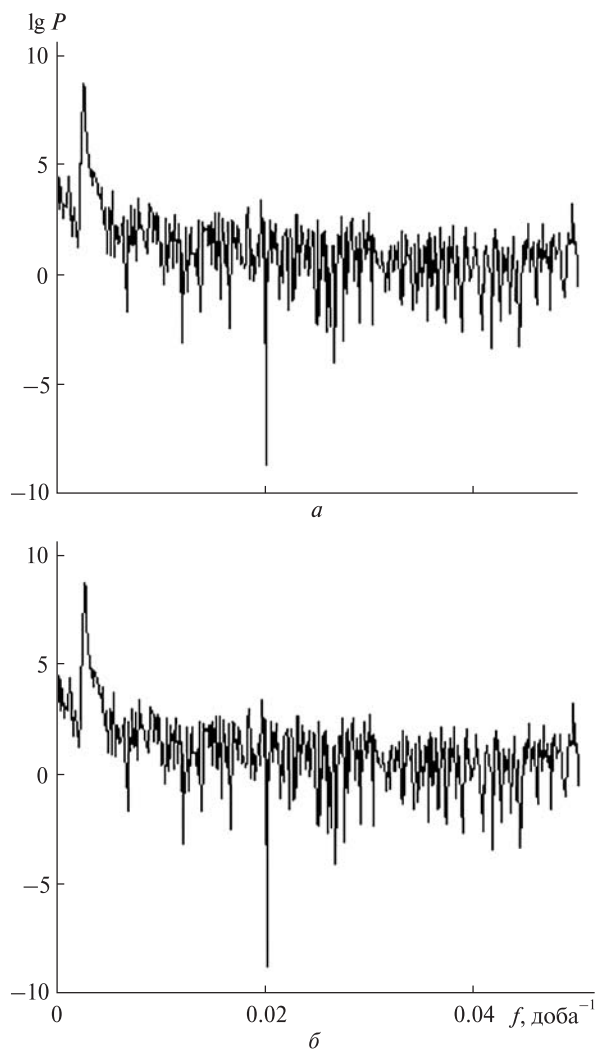


Рис. 2. Періодограми даних sr і swv

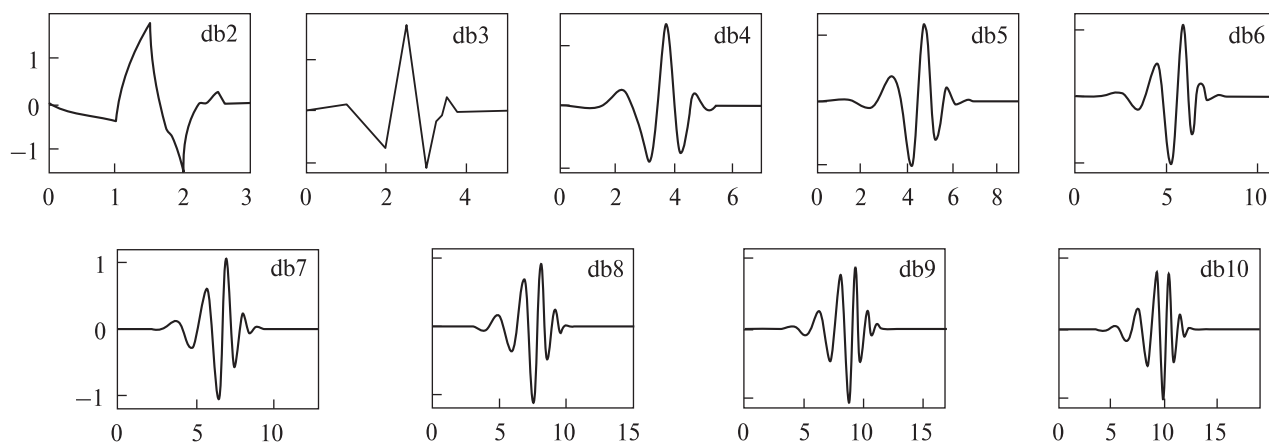


Рис. 3. Сімейство вейвлетів Добеши dbN

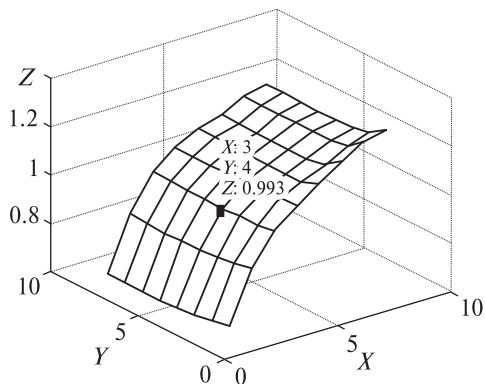


Рис. 4. Оптимальний розклад: db4, $L = 3$. Дані sr і swv

влетів Добеши: db — символи, походження яких пов'язане з прізвищем автора вейвлет-функцій, N — порядковий номер в сімействі вейвлетів Добеши. Ці сплески не мають аналітичного виразу.

Вибір типу вейвлета з приведенного сімейства зроблено емпірично (рис. 4).

Для вирішення задачі відновлення сигналу і аналізу складової шуму нами використовувалися програми пакетного вейвлет-аналізу зі складу програмного забезпечення MatLab 7.6.

ПАКЕТНИЙ ВЕЙВЛЕТ-АНАЛІЗ

Запровадження пакетів полягає в наступному. Для даної ортогональної вейвлет-функції створюється бібліотека баз, названа вейвлет-базисом пакета. Кожна з цих баз містить особливий спосіб кодування сигналів, збереження глобального енергетичного рівня і способу реконструкції конкретної функції. Вейвлет-пакети можуть бути використані для численних розширень аналізу даного сигналу.

Практична реалізація вейвлет-пакетів потребує вибір ентропії. Ентропія E повинна бути адитивною функцією, яка відповідає умовам $E(0) = 0$ і $E(s) = \sum_i s_i$, де s_i — відліки даних.

Виходячи з поставлених нами цілей і попередніх характеристик даних ентропія була задана у вигляді [8]:

$$E(s) = -\sum_i s_i \lg(s_i^2). \quad (3)$$

Це нормована ентропія, пов'язана з логарифмом відліків s_i .

Наступним кроком є вибір рівня L , при якому здійснюється розкладання сигналу. Пакет-аналіз представляє можливість вибору на основі кри-

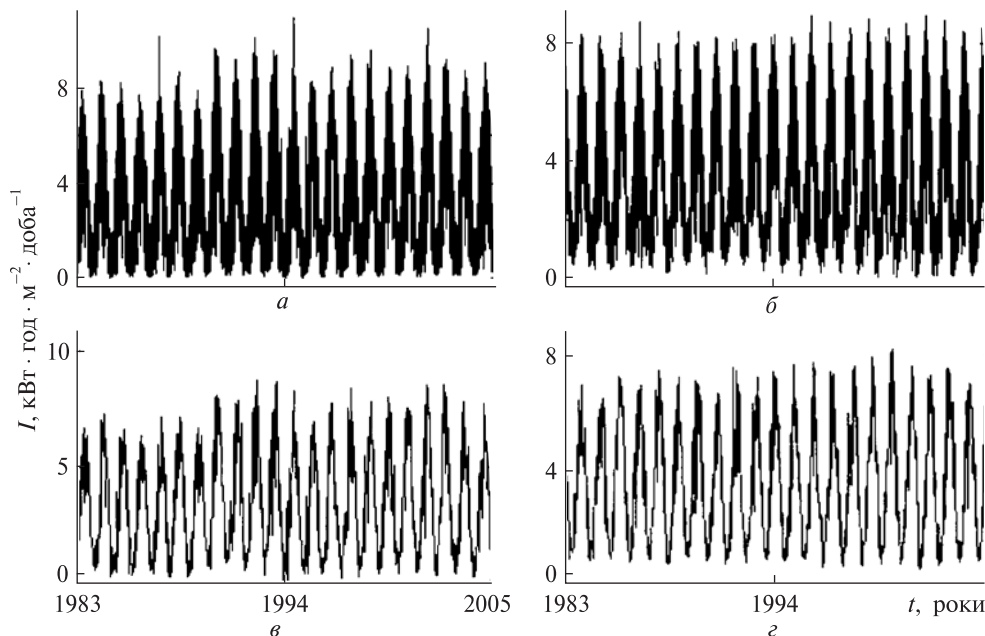


Рис. 5. Результати відновлення сигналу: а — дані sr , б — дані swv , в — відновлений сигнал у даних sr , г — відновлений сигнал у даних swv

терію [8], найбільш відповідного щодо ентропії рівня розкладання вихідного сигналу.

У ортогональній процедурі вейвлет-розкладання спільний крок розділяє коефіцієнти апроксимації на дві частини. Після первісного поділу отримуємо наближений вектор коефіцієнтів апроксимації і вектор коефіцієнтів деталізації. Наступний крок полягає в розщепленні нового вектора коефіцієнтів наближення. На рис. 4 приведені можливі поєднання вейвлетів сімейства Добеши (по осі Y) і відповідні їм рівні (по осі X); по осі Z відкладені стандартні відхилення. Обраною оптимальному розкладанню відповідають: $db4$, $L = 3$, $\sigma = 0.993$.

РЕЗУЛЬТАТИ ВЕЙВЛЕТ-АНАЛІЗУ

Основне завдання проведеного вейвлет-аналізу вихідних даних (вихідний сигнал) полягало в усуненні шуму (відновлення сигналу) і встановленню спектрального складу і статистик шумової компоненти. При підборі оптимальних параметрів для відновлення сигналу за основу було прийнято базову модель (2).

Шумова складова визначалася як різниця вхідного і відновленого сигналів. Ефективність усунення шуму забезпечена оптимальним вибором типу ентропії (3) і етапів відновлення збуреного сигналу за допомогою вейвлет-перетворення (1).

На рис. 5 наведено графіки: a — дані sr , b — дані swv , v — відновлений сигнал в даних sr , z — відновлений сигнал в даних swv . Зіставлення графіків вихідних даних до видалення шумових складових (рис. 5, a , b) і після (рис. 5, v , z) показує, що видалення шуму проведено без шкоди для чіткості деталей вихідного сигналу. У цьому перевага вейвлет-аналізу.

Нижче наведені графіки складової шуму в даних sr (рис. 6, a) і swv (рис. 6, b). Візуальний аналіз показує наявність окремих сплесків в серіях даних. Оптимальна процедура видалення шуму вимагає більш тонкого підходу, названого пороговим. Це передбачає відкидання тільки частини з деталей сигналу, які перевищують певну межу [9]. Графіки на рис. 6 обчислені з автоматичним підбором оптимального порогу, передбаченого пакет-аналізом.

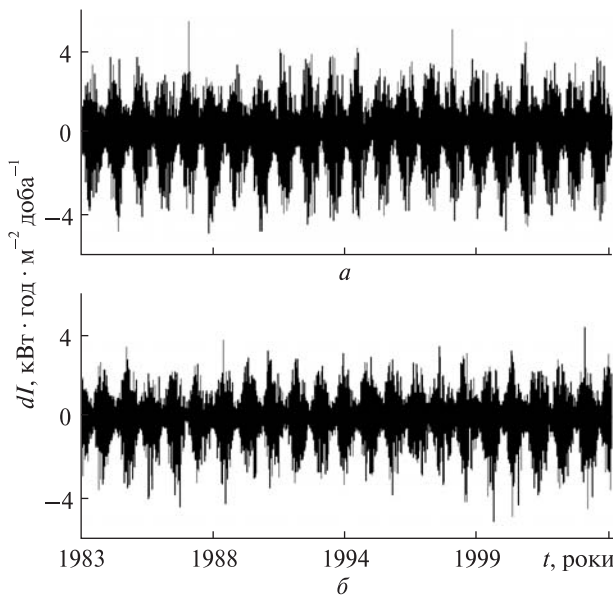


Рис. 6. Складова шуму: a — в даних sr , b — в даних swv

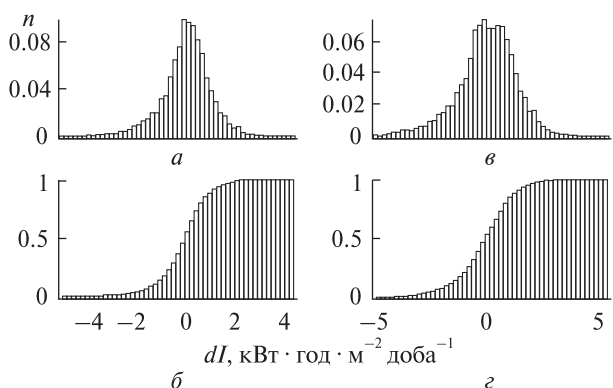


Рис. 7. Звичайні і кумулятивні гістограми (складова шуму): a , b — дані swv ; v , z — дані sr

Інформацію про розподіл складових шуму (рис. 6) відображають наведені на рис. 7 діаграми (гістограми і кумулятивні гістограми) розподілу частот. Наведені на рис. 7 гістограми свідчать про наближення розподілу відліків в складових шуму dI (рис. 6) до нормального. Більш докладна кількісна інформація про статистику складової шуму міститься в наведеній нижче таблиці.

Характеристики складових шуму доповнюють спектральний і кореляційний аналізи.

На рис. 8 наведені графіки автокореляції шумової складової в даних swv (рис. 8, a) і в даних sr (рис. 8, b). У широкому діапазоні зміни лагів τ

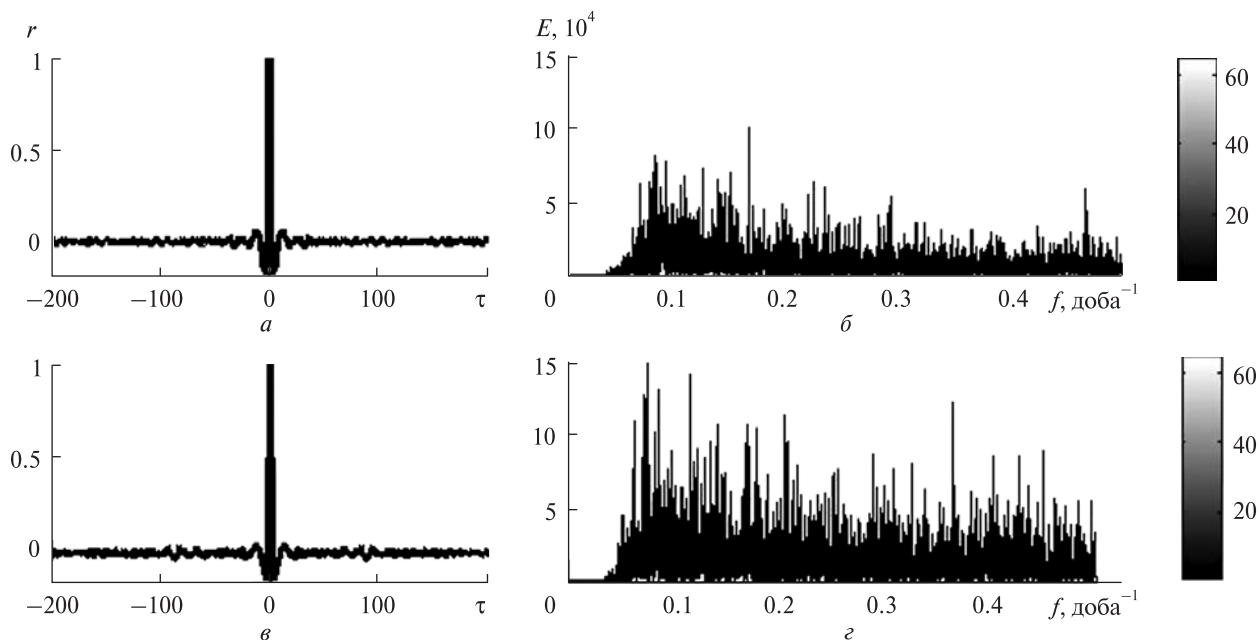


Рис. 8. Коефіцієнти автокореляції r і спектри Фур'є (складова шуму E): $a, б$ — дані swv ; $в, г$ — дані sr

Статистики шуму в даних sr, swv

Дані	$\bar{x}, 10^{-5}$	M	MM	MAX	MIN	σ	$L1$	$L2$	LL
sr	-2.7	0.07851	0.4248	5.388	-5.172	1.272	10370	130.8	5.388
swv	-4.424	0.06007	-01065	4.379	-5.150	0.993	5966	88.99	5.150

(рис. 8, $a, в$) коефіцієнт автокореляції r близький до нуля. Теоретична коррелограма при нормальному розподілу в цьому випадку є пряма лінія, що проходить через нуль. Такій умові відповідають часові ряди, породжені випадковим процесом.

Аналіз графіків періодограм, наведених на рис. 8, виявляє втрату інформації у високочастотній області вихідних даних. У разі необхідності аналізу коливань з періодами менше 20 діб, необхідне проведення більш гнучкої порогової фільтрації. Поглиблення аналізу високо-частотного компонента вихідних сигналів sr і swv необхідне з метою пояснення розбіжності, яка має місце між наземними і космічними вимірами. На наявність цих розбіжностей вказують графіки усереднених даних. На рис. 9 наведено графіки усереднених середньомісячних вимірів сонячної інсоляції на горизонтальну поверхню з космосу (SSE) і наземні для парку «Нікітський сад» з липня 1983 р. по червень 2005 р.

Кількісні характеристики статистик складових шуму в даних sr і swv наведені в таблиці. Вони включають в себе тенденції (середнє \bar{x} , мода MM , медіана M , максимальне і відхилення MAX , MIN від середнього значення) і дисперсію (стандартне відхилення σ , верхній і нижній рівень $L1$ і $L2$ відхилення від нормального розподілу, максимальне відхилення LL від нормального розподілу).

ВИСНОВКИ

1. Аналіз довгоперіодичної складової вихідних сигналів sr і swv виявляє їхню узгодженість на 22-річному інтервалі часу. Квадрат модуля когерентності основного коливання з періодом 1 рік складає більше 90 %.

2. Аналіз спектральних і статистичних характеристик складової шуму в даних sr і swv виявляє розбіжність, яка не може бути випадковою. Як показано на рис. 9, графіки усереднених серед-

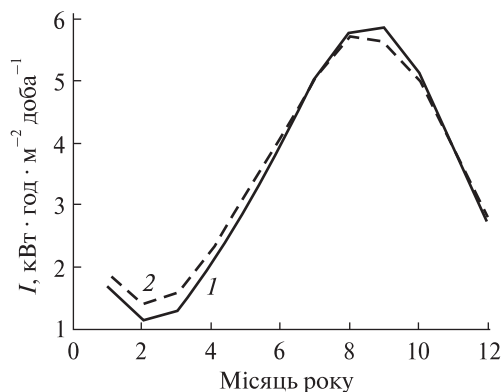


Рис. 9. Середньомісячна інсоляція на горизонтальну поверхню

ньомісячних вимірів сонячної інсоляції на горизонтальну поверхню з космосу (SSE) і наземні для парку «Нікітський сад» за період липень 1983 — червень 2005 рр. мають зсув: максимальні значення розбіжностей припадають на лютий, березень (інсоляція на горизонтальну поверхню за даними наземних вимірювань більша, ніж за космічними вимірюваннями) та вересень (зворотна залежність).

У русі Землі по орбіті місяці березень і вересень особливі: Земля перетинає площину екліптики в дні весняного і осіннього рівнодення. Залежність притоку сонячної енергії від положення Землі на орбіті — встановлений факт. Найбільші розбіжності між супутниковими і наземними вимірюваннями інсоляції припадають на період максимального притоку сонячної енергії.

3. Проведення подальшого розширення досліджень про наявність розбіжностей космічних і наземних вимірювань інсоляції включає: вживання гнучкого порогового підходу при проведенні вейвлет-аналізу складових в даних sr і swv ; встановлення зв'язку появи і розміру розбіжностей з положенням Землі на орбіті та зміною швидкості її обертання; встановлення зв'язку характеристик розбіжностей з локальними умовами клімату.

1. Авсюк Ю. Н. Приливные силы и природные процессы. — М.: Объединённый институт физики Земли им. О. Ю. Шмидта РАН, 1996. — 186 с.
2. Берри Б. Л. Спектр солнечной системы и модели геофизических процессов // Геофизика. — 2006. — № 3. — С. 64—68.

3. Берри Б. Л. Управление климатом, его прошлое и будущее // Холод'ОК. — 2008. — № 1 (6). — С. 72—78.
4. Берри Б. Л. Жить в тепле и уюте нам осталось недолго. Великая Эпоха (Epoch Times International) — международный информационный проект: <http://www.epochtimes.ru/content/view/8960/5/>
5. Вольвач А. Е., Корсакова С. П., Курбасова Г. С. та ін. Порівняння притоку сонячної радіації на верхню межу атмосфери і на поверхні Землі в пункті «Нікітський сад» за останні 30 років // Abstracts. 13th Ukrainian conference on Space Research (Yevpatoria, Crimea, Ukraine, September, 2—6, 2013). — Yevpatoria, 2013. — P. 31.
6. Добеши И. Десять лекций по вейвлетам. — Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001. — 464 с.
7. Сидоренков Н. С. Атмосферные процессы и вращение Земли. — СПб.: Гидрометеиздат, 2002. — 200 с.
8. Coifman R. R., Wickerhauser M. V. Entropy-based algorithms for best basis selection // IEEE Trans. Inform. Theory. — 1992. — 38, N 2. — P. 713—718.
9. Donoho D. L. De-Noising by soft-thresholding // IEEE Trans. Inform. Theory. — 1995. — 41, N 3. — P. 613—627.
10. Lilliefors H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown // J. Amer. Statist. Assoc. — 1967. — 62. — P. 399—402.

Стаття надійшла до редакції 24.04.14

Г. С. Курбасова, А. Е. Вольвач

ВЕЙВЛЕТ-АНАЛИЗ НАЗЕМНЫХ И КОСМИЧЕСКИХ ИЗМЕРЕНИЙ ЛОКАЛЬНОЙ ИНСОЛЯЦИИ

Показана возможность применения вейвлет-анализа для выделения сигнала из шума и статистического анализа шума в длительных рядах локальных наземных и космических наблюдений на основе измерений инсоляции спутниками НАСА и наземных измерений в парке «Никитский сад» (Крым, широта 44.5°, долгота 34.2°). Обоснован выбор метода анализа, вида вейвлетов и их параметров. Проведен анализ результатов оценок шума космических и наземных наблюдений.

G. S. Kurbasova, A. E. Volvach

WAVELET ANALYSIS FOR GROUND-BASED AND SPACE MEASUREMENTS OF LOCAL INSOLATION

We show the possibility to use the wavelet analysis for the selection of the signal from the noise and for the noise statistical analysis in long sequences of local ground-based and satellite observations on the basis of the NASA satellite measurements of the insolation and ground-based measurements in the «Nikitsky Garden» (Crimea, the latitude is 44.5° and the longitude is 34.2°). Our choice of the analysis method, type of wavelets and their parameters are justified. An analysis of the results of noise evaluations for space and ground-based observations is performed.