

*А.Ю. Дорошенко, В.М. Шпиг, О.М. Овдій*

## ПРОГРАМНА СИСТЕМА АНАЛІЗУ ХМАРНОСТІ ЗА ДАНИМИ СУПУТНИКОВИХ СПОСТЕРЕЖЕНЬ

У даній роботі за допомогою онлайн-ового діалогового конструктору для автоматизації проектування робочих процесів ОКРП було спроектовано робочий процес для обробки архіву супутникових знімків з метою аналізу показників хмарності на території України. Робочий процес для реалізації поставленої задачі було розроблено для системи Apache Oozie, призначеної для управління роботами розподіленої платформи Apache Hadoop. З метою підтримки масштабованості та оптимізації обробки для реалізації аналізу знімків було використано фреймворк Apache Spark. Проведено експерименти, що довели ефективність обраного підходу та відповідність отриманих результатів даним фактичних спостережень.

Ключові слова: розподілені обчислення, робочі процеси, Apache Hadoop, Apache Spark, аналіз зображень, моніторинг, хмарність.

### Вступ

В області метеорології задача обробки та аналізу великих обсягів даних є надзвичайно актуальною. Це зумовлено наявністю величезних об'ємів історичних та поточних даних метеорологічних величин, великим різноманіттям технічних засобів вимірювань і спостережень та складністю моделей. Однією із основних таких величин у моделях та синоптичних методах прогнозу погоди, кліматології, агрометеорології тощо є хмарність, тобто кількість хмар, яка характеризує ступінь покриття небосхилу хмарами і визначається спостерігачем у десятих частках одиниці (балах). Вона є одним із найважливіших факторів для перебігу багатьох фізичних процесів в атмосфері та біля земної поверхні. Змінюючись у часі та просторі, хмарність суттєво впливає на радіаційний та тепловий режими Землі, а отже, має безпосередній вплив на кліматичну систему планети, її глобальні та регіональні зміни. З іншого боку інформація про загальну хмарність, купчастодощові хмари та пов'язані із ними небезпечні та стихійні явища (наприклад, гроза, град, смерч, шквал і т. п.) є важливою як для цивільної, так і військової авіації, багатьох галузей економіки держави.

Протягом десятиліть в Україні кліматологічні дослідження хмарності (географічні та сезонні особливості, добовий хід тощо), розробка синоптичних та інших

методів прогнозування, верифікація моделей прогнозу погоди проводяться із використанням даних лише наземних спостережень [1–5]. Подібного характеру дослідження виконуються і в інших країнах у такий же спосіб [6, 7]. Водночас наявність супутникових даних відкриває широкий спектр можливостей покращення безпосереднього моніторингу хмарного покриву, пов'язаних із ним атмосферних опадів та інших явищ, отримання нових даних (шляхом обробки та аналізу різнорідної інформації) та для започаткування цілого вектору досліджень в Україні. Використання супутникової інформації дає можливість більш тонко здійснювати аналіз поля хмарності у регіонах зі складним рельєфом, оскільки орографічні особливості можуть змінювати хмарні системи на відносно коротких просторово-часових масштабах, а частота отримання даних дозволяє виокремити не тільки великомасштабне перенесення, але і особливості місцевої циркуляції у гірських районах, кластеризацію опадів тощо [8, 9].

На даний момент Україна не має своїх власних метеорологічних супутників, для потреб оперативного прогнозування та наукових досліджень використовуються дані EUMETSAT (Європейська організація з експлуатації метеорологічних супутників) та NOAA (Національна адміністрація з океанічних та

атмосферних досліджень, США), частина яких зберігається у вигляді знімків. Зокрема, дані щодо хмарного покриву та яскравісної температури в каналі 10,8 мкм радіометра SEVIRI геостационарного супутника MSG (EUMETSAT), які і було використано з метою створення системи аналізу хмарності для території України.

У даній роботі для вирішення задачі обробки супутникових знімків було вирішено використовувати розподілену обчислювальну платформу Apache Hadoop [10]. За допомогою системи для автоматизації проектування робочих процесів на основі алгеброалгоритмічного та онтологічного інструментарію ОКРП [11] було спроектовано робочий процес для системи Apache Oozie, призначеної для управління роботами Apache Hadoop. Система ОКРП є подальшим розвитком методологій та інструментів, що ґрунтуються на засобах високорівневої алгеброалгоритмічної формалізації і автоматизації перетворень програм [12–16], зокрема експериментальної інструментальної системи ІПС для конструювання та оптимізації паралельних програм, а також її онлайн версії системи ОДСП. З метою підтримки масштабованості та оптимізації обробки великого обсягу супутникових знімків для реалізації аналізу зображень було використано фреймворк Apache Spark. Проведено експерименти, які довели ефективність обраного підходу.

## 1. Програмна реалізація обробки супутникових знімків

У рамках даної роботи була поставлена задача обробки архіву супутникових знімків для обчислення та аналізу показників хмарності на території України. Знімки зберігаються у форматі збереження графічної інформації JPEG. На супутникових знімках в залежності від значень яскравісної температури хмарність зображено різними кольорами спектру. Конвективна хмарність ідентифікувалася згідно [17] з урахуванням географічного розташування України.

Для вирішення задачі необхідно обчислити загальну хмарність, конвективну

та неконвективну. При аналізі знімків виникло декілька проблем. Перша проблема пов'язана з тим, що JPEG є форматом, що використовує стиснення з втратами і в зв'язку з великим ступенем стиснення вихідних зображень втрата та спотворення даних є дуже великими. Цю проблему було вирішено за рахунок введення похибок та попередньої обробки зображення з метою вилучення шумів. Наступною проблемою були нанесені на зображення обриси кордонів областей та назв населених пунктів. Цю проблему було вирішено шляхом використання маски для їх вилучення з розрахунків з подальшим обчисленням значення кольорів точок зображення, що потрапили під маску, за найближчими точками поза маскою. В зв'язку з тим, що всі ці заходи не в змозі повністю відновити зображення, в обчисленнях присутні похибки.

Для реалізації аналізу супутникових знімків було вирішено використати систему для автоматизації проектування робочих процесів ОКРП [11]. В даній системі поєднання інструментів онтологій та алгеброалгоритмічних інструментів забезпечує значний потенціал для адаптації, оптимізації, інтеграції та модифікації розроблюваних робочих процесів. Робочі процеси [18] забезпечують систематичний спосіб опису необхідних методів та інтерфейс між фахівцями та обчислювальним середовищем, спрощуючи задачу обробки та аналізу даних з різних джерел на широкому спектрі обчислювальних платформ.

Робочий процес для реалізації поставленої задачі було розроблено для системи Apache Oozie [19]. Apache Oozie є системою керування робочими процесами для управління роботами розподіленої платформи Apache Hadoop. Apache Hadoop призначена для розподіленого зберігання й обробки великих об'ємів даних на великих комп'ютерних кластерах.

З метою підтримки масштабованості та оптимізації обробки великого обсягу супутникових знімків для реалізації кроку аналізу знімків по отриманню даних про хмарність використовується

фреймворк Apache Spark [20, 21]. Apache Spark – це кластерний обчислювальна фреймворк з відкритим вихідним кодом для розподіленої обробки великих об'ємів даних. Spark не використовує MapReduce [22] як механізм виконання, замість цього він використовує власне розподілене середовище. Тим не менш, Spark має багато спільного з MapReduce з точки зору API та середовища виконання. Spark тісно інтегрований з Apache Hadoop і він може запускатися як на Hadoop YARN так і у відокремленому режимі. Spark може працювати з різноманітними форматами файлів та сховищами даних Hadoop такими як розподілена файлова система HDFS, розподілена база даних HBase і т. д.

На відміну від MapReduce, де дані завжди завантажуються з диска, Spark здатен між робочими задачами зберігати їх у оперативній пам'яті, що дозволяє значно збільшити швидкість виконання для деяких класів задач. Ця можливість дозволяє Spark перевершити за продуктивністю еквівалентний робочий процес MapReduce (на порядок або більше в деяких випадках).

Spark підтримує такі мови програмування як Scala, Java і Python.

Для більш детального статистичного аналізу отриманих з супутникових знімків даних про хмарність застосовується

програмне середовище R [23], що призначене для статистичних обчислень та аналізу даних.

На рис. 1 показано схему розробленого робочого процесу для аналізу хмарності за супутниковими знімками.

Робочий процес складається з наступних кроків:

- **start** – запуск робочого процесу;
- **spark-job** – крок робочого процесу, який запускає програму Apache Spark, що виконує обробку масиву супутникових знімків для отримання показників хмарності. Деталі реалізації даного кроку описані нижче;
- **fork** – розбиває шлях виконання робочого процесу на кілька паралельних. Це дозволяє паралельно провести статистичний аналіз отриманих на попередньому кроці даних про хмарність;
- **r-mean-by-month** – shell-скрипт, який запускає на виконання програму R для статистичного аналізу даних та виводу результатів. Програма обчислює середні показники хмарності за місяцями та графічно представляє результат зі зберіганням графіків у форматі PDF;
- **r-mean-by-day** – shell-скрипт, який запускає на виконання програму R для статистичного аналізу даних та виводу

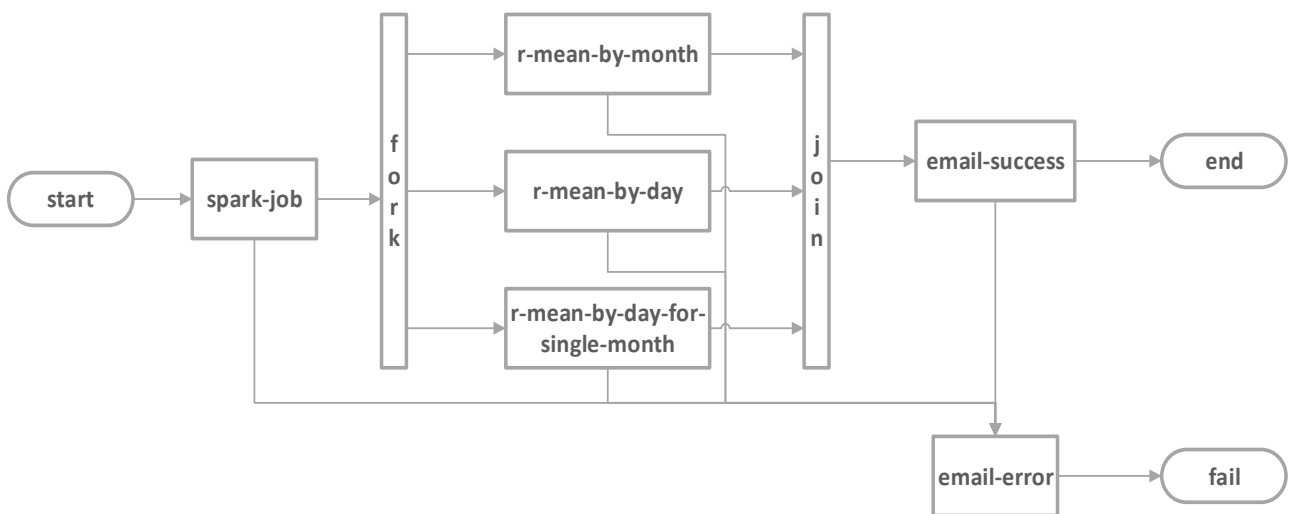


Рис. 1. Схема робочого процесу аналізу супутникових знімків

результатів. Програма обчислює середні показники хмарності за днями та графічно представляє результат зі зберіганням графіків у форматі PDF;

- **r-mean-by-day-for-single-month** – shell-скрипт, який запускає на виконання програму R для статистичного аналізу даних та виводу результатів. Програма обчислює середні показники хмарності за днями для обраного місяця та графічно представляє результат зі зберіганням графіків у форматі PDF;

- **join** – очікує, поки закінчаться всі паралельні шляхи виконання робочого процесу;

- **email-success** – сповіщає про вдале виконання робочого процесу та надсилає результати;

- **email-error** – сповіщення у разі виникнення помилок та збою робочого процесу.

- **end** – успішне завершення робочого процесу;

- **fail** – примусове завершення робочого процесу, наприклад, у разі виникнення помилок.

Програму для обробки та аналізу супутникових знімків на платформі Apache Spark, що виконується у рамках спроектованого робочого процесу, було реалізовано мовою Java. Вона складається з таких основних кроків:

- завантаження зображень та налаштувань;

- перевірки потрапляння точок зображення, що аналізується, у вибрану область (вся територія України, Київська область і т. д.) з використанням масиву координат досліджуваної області за допомогою методу трасування променем;

- фільтрація точок зображення за маскою, для вилучення нанесених на карту кордонів регіонів та назв населених пунктів. Обчислення значення кольорів точок, що потрапили під маску виконується за допомогою обходу масиву точок за спіраллю на предмет пошуку найближчої точки, що не потрапила під маску;

- обхід масиву точок та визначення наявності в них хмарності (див. рис. 2) та її типу (конвективна, неконвективна) на підставі попадання кольору точки у відповідний діапазон кольорового спектру;

- обробка проміжних результатів для вилучення шумів;

- обрахунок площі хмарності різних типів та збереження результатів.

Далі наведено фрагмент згенерованого коду специфікації робочого процесу для цільової системи управління робочими процесами, у даному випадку Apache Oozie.

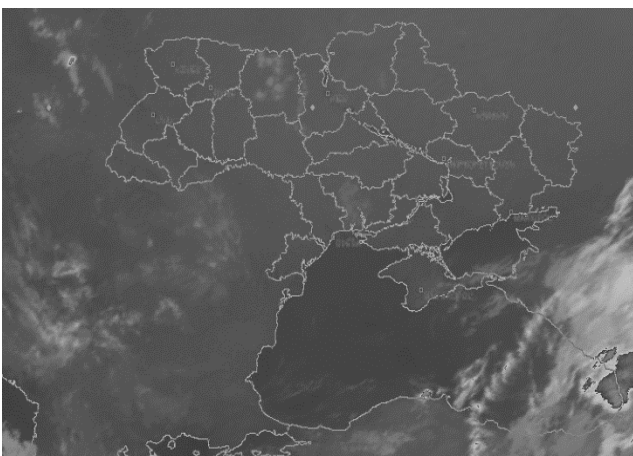


Рис. 2. Фрагмент супутникового знімку та результат фільтрації точок за маскою з ідентифікованою хмарністю

```
<workflow-app xmlns="uri:oozie:workflow:1.0" name="forecast-wf">
<startto="spark-job"/>
<actionname="spark-job">
<sparkxmlns="uri:oozie:spark-action:1.0">
...
</spark>
<okto="r-analyse"/>
<errorto="email-error"/>
</action>
<forkname="r-analyse">
<pathstart="r-mean-by-month" />
<pathstart="r-mean-by-day" />
<pathstart="r-mean-by-day-for-single-month" />
</fork>
<actionname="r-mean-by-month">
<shellxmlns="uri:oozie:shell-action:1.0">
...
</shell>
<okto="email-error"/>
<errorto="email-error"/>
</action>
...
<joinname="join-r-analyse" to="email-success"/>
<actionname="email-success">
...
</action>
<actionname="email-error">
...
</action>
<killname="fail">
...
</kill>
<endname="end"/>
</workflow-app>
```

В рамках роботи було обчислено показники хмарності для усієї території України та окремо для Київської області.

## 2. Результати експерименту

Розроблений в рамках даної роботи робочий процес було застосовано для обробки архіву супутникових знімків, збережених у форматі JPEG. Архів містить знімки за кожен день з інтервалом 15 хвилин. Для прикладу було обрано вибірку за один рік з 01.03.2018 по 28.02.2019 об'ємом 8 Гб.

Для виконання робочого процесу використано фізичну обчислювальну машину з такими характеристиками:

- 4-ядерний процесор IntelCore i7-6700HQ з частотою 2.6 ГГц,
- об'єм оперативної пам'яті 32 Гб.

Для проведення експерименту було розгорнуто інфраструктуру Apache Hadoop у псевдорозподіленому режимі. Архів зні-

мків було перенесено у розподілену файловою системою HDFS.

При випробуваннях було досліджено швидкість виконання робочого процесу при однаковому об'ємі вихідних даних для різної кількості потоків виконання програми Apache Spark.

Графік на рис. 3 показує залежність часу виконання робочого процесу для обчислення показників хмарності по Київській області для різної кількості потоків виконання Apache Spark. Значення прискорення при подвоєнні кількості потоків у середньому складає 1,8.

У результаті проведених обчислень було отримано часові розподіли площі, вкритої хмарним покривом, у межах України різного ступеня деталізації. На рис. 4 та рис. 5 показано річний хід на території України та для березня 2018 р. середні добові значення у межах Київської області площ загальної (позначено суцільною лінією), конвективної (позна-

чено штриховою лінією) та неконвективної хмарності (позначено штрихпунктирною лінією). Отриманий річний хід (див. рис. 4) на якісному рівні добре узгоджується із даними багаторічних спостережень, приведеними у [2]. Так, найбільше хмар спостерігається у зимові місяці, весною хмарність зменшується, влітку тенденція зменшення продовжується (проте просторовий розподіл хмар більш неоднорідний, ніж зимою та весною; з'являються сприятливі умови для розвитку потужних конвективних хмар), восени хмарність починає збільшуватися. Ясно, що у межах окремо взятого місяця (див. рис. 5), протягом року та від одного року до іншого значення загальної хмарності та окремих її типів будуть суттєво відрізнятися. Проте відповідно до отриманих результатів, можна висловити припущення, що використання супутникових даних дає вищі значення хмарності, ніж значення отримані на основі наземних спостережень. Згідно [2] у зимові місяці майже на всій території країни близько 70–75 % небосхилу закрито хмарами, у нашому випадку за даними супутникових спостережень для періоду в один рік це – близько 90 %. Навесні за даними багаторічних наземних спостережень це – 60 %, у бере-

зні – травні 2018 р. за даними супутникових спостережень мало місце зменшення від 83 % до 28 % відповідно. Аналогічна схожість має місце і для інших сезонів. На рис. 6 приведено дані спостережень за загальною хмарністю метеостанції Київ для березня 2018 р. Розриви у кривій (лакуни) пояснюються відсутністю даних в архіві Українського гідрометеорологічного інституту ДСНС України та НАН України. Порівнюючи рис. 5 та 6 і беручи до уваги, що площа Київської області становить 28131 км<sup>2</sup>, можемо бачити добре узгодження ходу загальної хмарності як за величиною площі, так і за ступенем покриття небосхилу. Проте є відмінність щодо настання мінімумів. У Києві вони спостерігались 11, 19 та 24 березня, а по області 13, 20 та 23 березня 2018 р.

Розбіжності, які мають місце, можна пояснити одночасним і рівномірним охопленням великої площі, відсутністю місцевих перешкод (закритості частини горизонту) для спостереження, а також проходженням атмосферних фронтів (коли над однією частиною точок наземних спостережень ще не спостерігається, а над іншою вже спостерігається суцільний хмарний покрив).

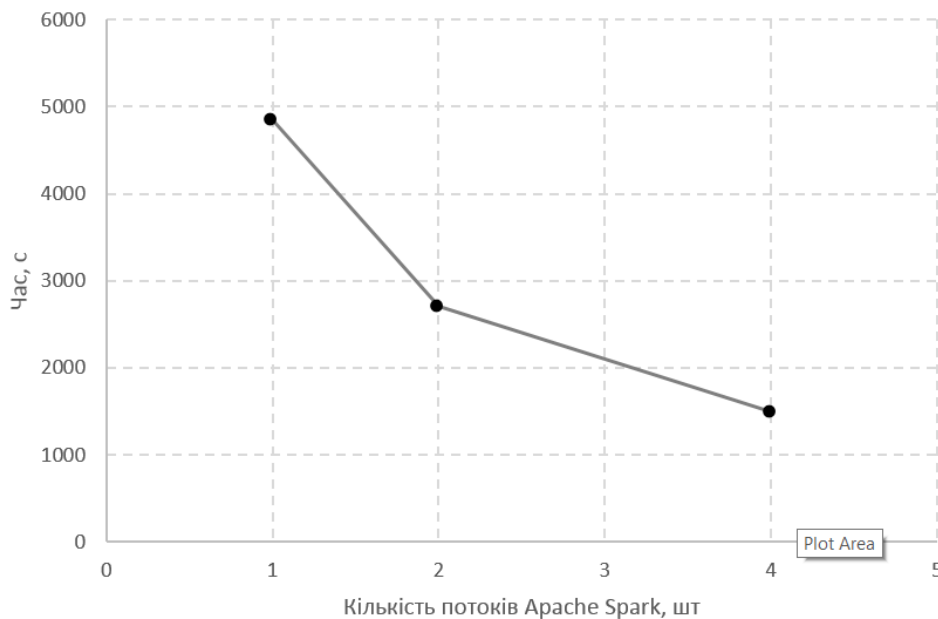


Рис. 3. Залежність часу виконання робочого процесу від кількості потоків виконання Apache Spark

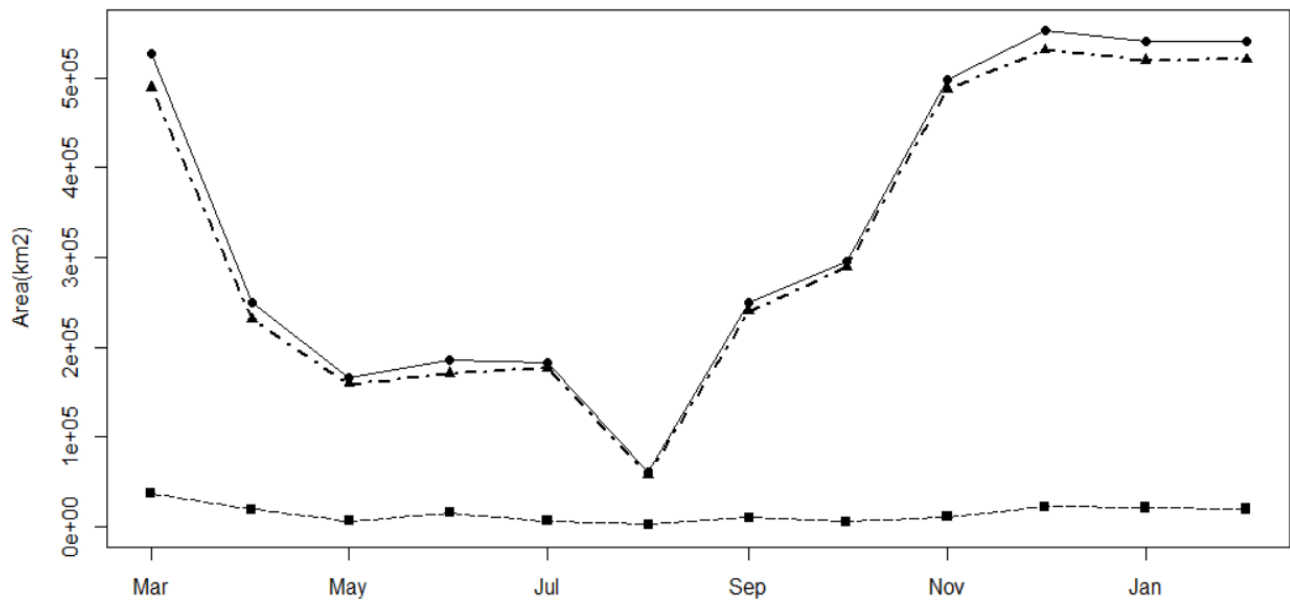


Рис. 4. Річний хід площ загальної, неконвективної та конвективної хмарності [ $10^5 \text{ км}^2$ ] на території України

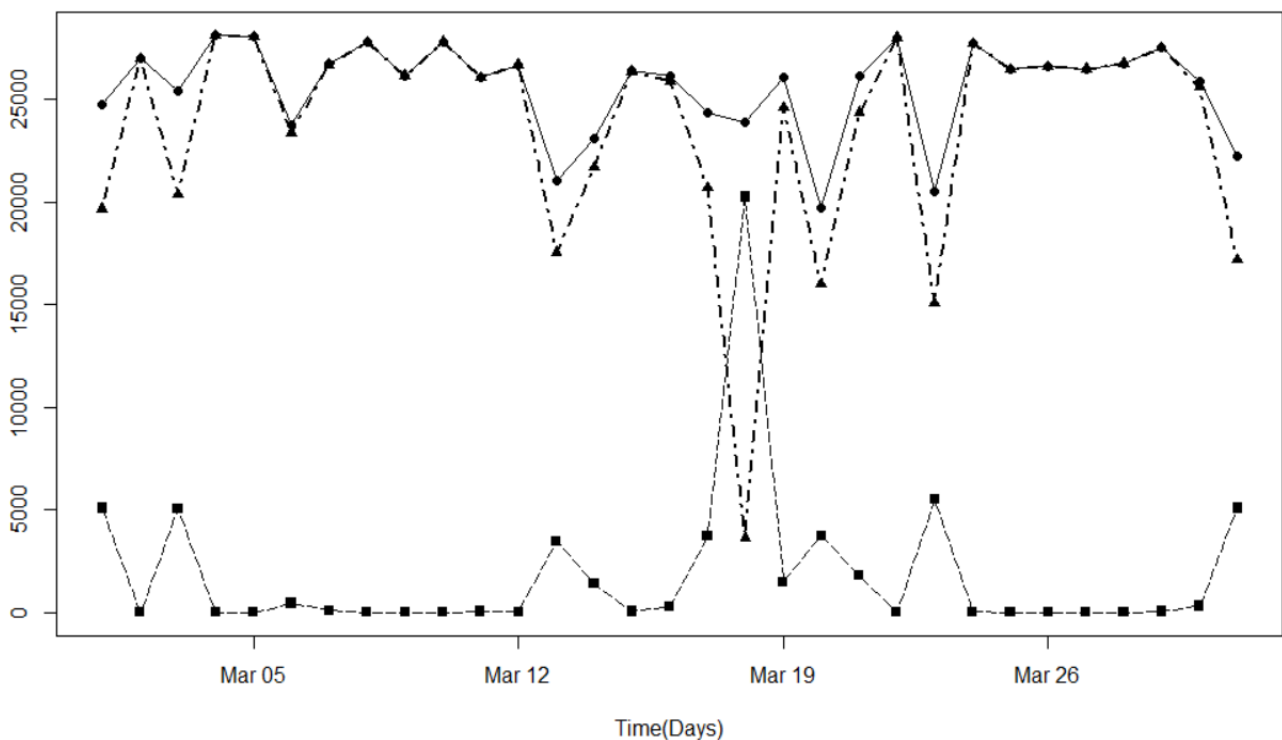


Рис. 5. Середні добові значення площ загальної, неконвективної та конвективної хмарності [ $\text{км}^2$ ] на території Київської області протягом березня 2018 р.

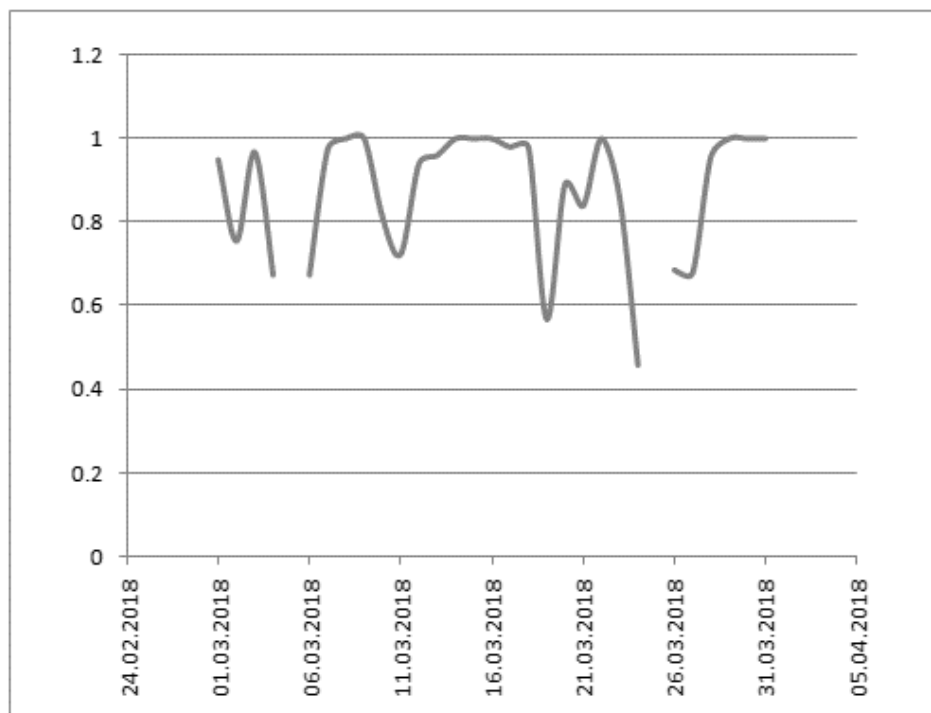


Рис. 6. Середня за добу загальна хмарність за даними спостережень у м. Київ у березні 2018 р.

## Висновки

В даній роботі за допомогою он-лайнного діалогового конструктору для автоматизації проектування робочих процесів ОКРП було спроектовано робочий процес для обробки архіву супутникових знімків з метою аналізу показників хмарності на території України. Робочий процес для реалізації поставленої задачі було розроблено для системи Apache Oozie, призначеної для управління роботами розподіленої платформи Apache Hadoop. З метою підтримки масштабованості та оптимізації обробки великого обсягу супутникових знімків для реалізації аналізу зображень було використано фреймворк Apache Spark. Проведено експеримент щодо виконання розробленого робочого процесу при різній кількості потоків Apache Spark, результати якого продемонстрували хороший показник ефективності розпаралелювання обчислень. Показано, що отримані у ході експерименту результати на якісному рівні добре узгоджуються із даними багаторічних наземних спостережень та фактичними даними спостережень. Висловлено припущення, що використання

супутникових даних дає вищі значення хмарності, що пояснюється одночасним і рівномірним охопленням великої території, відсутністю місцевих перешкод (закритості частини горизонту) та проходженням атмосферних фронтів.

## Література

1. Клімат України [Под редакцией Г.Ф. Приходько, А.В. Ткаченко, В.Н. Бабиченко]. Ленинград: Гидрометеиздат, 1967. 227 с.
2. Клімат України [За редакцією В.М. Ліпінського, В.А. Дячука, В.М. Бабіченко]. Київ: видавництво Раєвського, 2003. 344 с.
3. Заболоцька Т.М., Шпиг В.М. Трансформація баричного поля та хмарності у випадку тривалих і сильних опадів. *Наукові праці Українського науково-дослідного гідрометеорологічного інституту*. 2014. № 266. С. 12–19.
4. Заболоцька Т.М., Шпиг В.М. Кількісні зміни хмарності як показник тривалості періоду глобального потепління. *Наукові праці Українського науково-дослідного гід-*



- рометеорологічного інституту. 2015. № 267. С. 23–27.
5. Shpyg V. et al. The application of regional NWP models to operational weather forecasting in Ukraine. CAS Technical Conference (TECO) on “Responding to the Environmental Stressors of the 21st Century”: 18–19 November 2013: Conf. Materials. 2013. URL: <http://www.wmo.int/pages/prog/arep/cas/documents/Ukraine-NWPMODELS.pdf> (дата звернення: 27.06.2019).
  6. Хлебникова Е.И., Салль И.А. Особенности климатических изменений облачного покрова над территорией России. Метеорология и гидрология. 2009. № 7. С. 5–13.
  7. Warren S.G., Eastman R.M., Hahn C.J. A survey of Changes in Cloud Cover and Cloud Types over Land from Surface Observations, 1971-96. *Climate*. 2007. N 20. P. 717–738.
  8. Giovannetone J.P. and Barros A.P. Probing Regional Orographic Controls of Precipitation and Cloudiness in the Central Andes Using Satellite Data. *Journal of Hydrometeorology February*. 2009. Vol. 10, N 1. P. 167–182.
  9. Sumargo E. and Cayan D. R. Variability of Cloudiness over Mountain Terrain in the Western United States. *Journal of Hydrometeorology*. 2017. Vol. 18, N 5. P. 1227–1245.
  10. Apache Hadoop: сайт. URL: <http://hadoop.apache.org/> (дата звернення: 1.06.2019).
  11. Овдій О.М. До питання автоматизації проектування робочих процесів на основі алгебро-алгоритмічного та онтологічного інструментарію. *Проблеми програмування*. 2019. № 1. С. 37–47.
  12. Дорошенко Е.А., Яценко Е.А. О синтезе программ на языке Java по алгеброалгоритмическим спецификациям. *Проблеми програмування*. 2006. № 4. С. 58–70.
  13. Андон Ф.И., Дорошенко А.Е., Цейтлин Г.Е., Яценко Е.А. Алгеброалгоритмические модели и методы параллельного программирования. Киев: Академперіодика, 2007. 631 с.
  14. Андон Ф.И., Дорошенко А.Е., Бекетов А.Г., Иовчев В.А., Яценко Е.А. Инструментальные средства автоматизации параллельного программирования на основе алгебры алгоритмов. *Кибернетика и системный анализ*. 2015. № 1. С. 162–170.
  15. Дорошенко А.Ю., Иваненко П.А., Овдій О.М., Яценко О.А. Автоматизоване проектування програм для розв’язання задачі метеорологічного прогнозування. *Проблеми програмування*. 2016. № 1. С. 102–115.
  16. Дорошенко Е.А., Овдей О.М., Яценко Е.А. Онтологические и алгеброалгоритмические средства автоматизации проектирования параллельных программ для "облачных" платформ. *Кибернетика и системный анализ*. 2017. Т. 53, № 2. С. 181–192.
  17. Setvák M. et al. Blended "Sandwich" Image Products in Nowcasting. 2012 EUMETSAT Meteorological Satellite Conference. URL: [https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET\\_FILE&dDocName=PDF\\_CONF\\_P61\\_S7\\_11\\_SETVAK\\_V&RevisionSelectionMethod=LatestReleased&Renderition=Web](https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_CONF_P61_S7_11_SETVAK_V&RevisionSelectionMethod=LatestReleased&Renderition=Web) (дата звернення: 27.06.2019).
  18. Workflow Management Coalition: сайт. URL: <https://www.wfmc.org/> (дата звернення: 1.06.2019).
  19. Apache Oozie Workflow Scheduler for Hadoop: сайт. URL: <http://oozie.apache.org/> (дата звернення: 1.06.2019).
  20. Apache Spark: сайт. URL: <https://spark.apache.org/> (дата звернення: 1.06.2019).
  21. White T. Hadoop: The Definitive Guide, 4th Edition. O'Reilly Media, Inc. 2015. 756p.
  22. Dean J., Ghemawat S. Mapreduce: Simplified data processing on large clusters. In Proc. of the 6th USENIX OSDI. 2004. P. 137-150.
  23. The R Project for Statistical Computing: сайт. URL: <https://www.r-project.org/> (дата звернення: 1.06.2019).

## References

1. Climate of Ukraine [Ed. by G.F. Prikhodko, A.V. Tkachenko, V.N. Babichenko] (1967). Leningrad: Hydrometeoizdat. 227 p. (in Russian).
2. Climate of Ukraine [Ed. by V.M. Lipinsky, V.A. Diachuk, V.M. Babichenko] (2003). Kyiv: Publishing House of Raevsky. 344 p. (in Ukrainian).
3. Zabolotska, T.M., Shpyg V.M. (2014). Transformation of baric field and cloudiness in the case of long-term and heavy precipitation. *Scientific Proceedings of the UHMI*. 266. P. 12–19. (in Ukrainian).
4. Zabolotska, T.M., Shpyg V.M. (2015). Quantitative changes of cloud cover as indicator of global warming period. *Scientific Proceedings of the UHMI*. 267. P. 23–27. (in Ukrainian).
5. Shpyg V. et al. (2013). The application of regional NWP models to operational weather forecasting in Ukraine. CAS Technical Con-

- ference (TECO) on “Responding to the Environmental Stressors of the 21st Century”. [online] Available from: <http://www.wmo.int/pages/prog/arep/cas/documents/Ukraine-NWPMODELS.pdf> [Accessed 27 Jun. 2019].
6. Khlebnikova E.I., Sall I.A. (2009). Peculiarities of climatic changes in cloud cover over the Russian Federation. *Russian Meteorology and Hydrology*. 7. P. 5–13. (in Russian).
  7. Warren S.G., Eastman R.M., Hahn C.J. (2007). A survey of Changes in Cloud Cover and Cloud Types over Land from Surface Observations, 1971-96. *Climate*. N 20. P. 717–738.
  8. Giovannetone J.P. and Barros A.P. (2009). Probing Regional Orographic Controls of Precipitation and Cloudiness in the Central Andes Using Satellite Data. *Journal of Hydrometeorology* February. Vol. 10, N 1. P. 167–182.
  9. Sumargo E. and Cayan D.R. (2017). Variability of Cloudiness over Mountain Terrain in the Western United States. *Journal of Hydrometeorology*. Vol. 18, N 5. P. 1227–1245.
  10. Hadoop.apache.org. Apache Hadoop Official Website. [online] Available from: <http://hadoop.apache.org/> [Accessed 1 Jun. 2019].
  11. Ovdii, O.M. (2018) On the issue of automating the workflow design based on algebra-algorithmic and ontological tools. *Problems in programming*. (1). P. 37–47. (in Ukrainian).
  12. Doroshenko, A.Yu. & Yatsenko O.A. (2006) About the synthesis of Java programs by algebra-algorithmic specifications. *Problems in programming*. (4). P. 58–70. (in Russian).
  13. Andon, P.I. et al. (2007) Algebra-algorithmic models and methods of parallel programming. *Kiev: Academperiodika*. (in Russian).
  14. Andon, P.I., Doroshenko, A.Yu., Beketov, O.G., Iovchev, V.O. & Yatsenko O.A. (2015) Software tools for automation of parallel programming on the basis of algebra of algorithms. *Cybernetics and systems analysis*. (1). P. 162–170. (in Russian).
  15. Doroshenko, A.Yu., Ivanenko, P.A., Ovdii, O.M., & Yatsenko, O.A. (2016) Automated design of programs for solving the task of meteorological forecasting. *Problems in programming*. (1). P. 102–115. (in Ukrainian).
  16. Doroshenko, A.Yu., Ovdii, O.M. & Yatsenko O.A. (2017) Ontological and algebra-algorithmic tools for automated design of parallel programs for cloud platforms. *Cybernetics and Systems Analysis*. 53(2). P. 181–192. (in Russian).
  17. Setvák M. et al. (2012). Blended "Sandwich" Image Products in Nowcasting. *EUMETSAT Meteorological Satellite Conference*. [online] Available from: [https://www.eumetsat.int/website/wcm/idc/idc.plg?IdcService=GET\\_FILE&dDocName=PDF\\_CONF\\_P61\\_S7\\_11\\_SETVAK\\_V&RevisionSelectionMethod=LatestReleased&Renderition=Web](https://www.eumetsat.int/website/wcm/idc/idc.plg?IdcService=GET_FILE&dDocName=PDF_CONF_P61_S7_11_SETVAK_V&RevisionSelectionMethod=LatestReleased&Renderition=Web) [Accessed 27 Jun. 2019].
  18. Wfmc.org. Workflow Management Coalition. [online] Available from: <https://www.wfmc.org/> [Accessed 1 Jun. 2019].
  19. Oozie.apache.org. Apache Oozie Workflow Scheduler for Hadoop Official Website. [online] Available from: <http://oozie.apache.org/> [Accessed 1 Jun. 2019].
  20. Spark.apache.org Unified analytics engine for BigData. [online] Available from: <https://spark.apache.org/> [Accessed 1 Jun. 2019].
  21. White T. (2015) Hadoop: The Definitive Guide, 4th Edition. O'Reilly Media, Inc.
  22. Dean J. & Ghemawat S. (2004) Mapreduce: Simplified data processing on large clusters. In: 6th USENIX OSDI, pp. 137-150.
  23. R-project.org. The R Project for Statistical Computing Official Website. [online] Available from: <https://www.r-project.org/> [Accessed 1 Jun. 2019].

Одержано 01.07.2019

### Про авторів:

*Дорошенко Анатолій Юхимович*, доктор фізико-математичних наук, професор, завідувач відділу теорії комп'ютерних обчислень Інституту програмних систем НАН України, професор кафедри автоматизації та управління в технічних системах НТУУ “КПІ імені Ігоря Сікорського”. Кількість наукових публікацій в українських виданнях – понад 150. Кількість наукових публікацій в зарубіжних виданнях – понад 50. Індекс Хірша – 5. <http://orcid.org/0000-0002-8435-1451>,

*Шниг Віталій Михайлович*,  
кандидат географічних наук,  
завідувач відділу фізики атмосфери  
Українського гідрометеорологічного  
інституту ДСНС України та НАН України,  
експерт Комісії з атмосферних наук  
Всесвітньої метеорологічної організації.  
Кількість наукових публікацій в  
українських виданнях – понад 40.  
Кількість наукових публікацій в  
зарубіжних виданнях – понад 50.  
Індекс Хірша – 2.  
<https://orcid.org/0000-0003-1055-7120>,

*Овдій Ольга Михайлівна*,  
молодший науковий співробітник  
Інституту програмних систем  
НАН України.  
Кількість наукових публікацій в  
українських виданнях – 22.  
Кількість наукових публікацій в  
зарубіжних виданнях – 5.  
<http://orcid.org/0000-0002-8891-7002>.

***Місце роботи авторів:***

Інститут програмних систем  
НАН України,  
03187, м. Київ-187,  
проспект Академіка Глушкова, 40.  
Тел.: (38)(044) 526 6033.  
E-mail: doroshenkoanatoliy2@gmail.com,  
olga.ovdiy@gmail.com;

Український гідрометеорологічний  
Інститут ДСНС України та НАН України,  
03028, м. Київ,  
проспект Науки, 37.  
Тел.: (38)(044) 525 8630.  
E-mail: vitold82@i.ua.