

УДК 528:004.528

DOI:10.30533/0536-101X-2022-66-6-37-49



ОСОБЕННОСТИ КОМПЛЕКСИРОВАНИЯ ГИДРОМЕТЕОРОЛОГИЧЕСКИХ И ЭКОЛОГИЧЕСКИХ ДАННЫХ С ЦЕЛЬЮ УТОЧНЕНИЯ РЕЗУЛЬТАТОВ АНАЛИЗА ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ НИЗКОГО ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ ДЛЯ ОБЕСПЕЧЕНИЯ ЗАДАЧ МОНИТОРИНГА ТЕРРИТОРИЙ

А.Ю. БЫСТРОВ¹✉, А.А. МАЙОРОВ²

¹ ООО «СберОбразование», Москва, Россия

² Московский государственный университет геодезии
и картографии, Москва, Россия

✉ bystrovaur@gmail.com

ЦИТИРОВАНИЕ Быстров А.Ю., Майоров А.А. Особенности комплексирования гидрометеорологических и экологических данных с целью уточнения результатов анализа данных дистанционного зондирования

Земли низкого пространственного разрешения для обеспечения задач мониторинга территорий // Известия вузов «Геодезия и аэрофотосъемка». 2022. Т. 66. № 6. С. 37–49. DOI:10.30533/0536-101X-2022-66-6-37-49

КЛЮЧЕВЫЕ СЛОВА геоинформационные системы, дистанционное зондирование Земли, водоохранные зоны, мониторинг территорий, анализ геоданных, краудсорсинг, пространственные данные, гидрометеорология, прогнозирование погоды

АННОТАЦИЯ

В статье проведен обзор современных гидрометеорологических мониторинговых геоинформационных систем, использующих данные дистанционного зондирования Земли (ДЗЗ) низкого пространственного разрешения. Описаны особенности данных ДЗЗ, получаемых с космических аппаратов Terra, Aqua, Meteosat, NOAA, принимаемых наземными антенными комплексами X и L-диапазона и используемых в таких системах. В качестве основного примера наиболее комплексной системы рассмотрена технология Meteum, основанная на принципе комплексирования разнородных пространственных данных и применения технологии искусственного интеллекта (ИИ) для анализа данных и прогнозирования погоды. Рассмотрены отдельные решения, позволяющие обеспечивать метеорологический мониторинг, такие как Экспериментальная национальная ГИС (NWS), MADIS, AccuWeather, а также представлен пример применения мобильных динамических геосенсоров, используемых в городской автобусной сети для получения дополнительных данных при прогнозировании погоды. На основе проведенного обзора произведена доработка концептуальной модели системы геоинформационного мониторинга, ключевыми особенностями, которой является применение подхода по уточнению данных на основе краудсорсинга и использования ИИ как для обработки пространственных данных, так и для анализа состояния объекта мониторинга на основе результатов обработки.

1 ВВЕДЕНИЕ

Гидрометеорологический мониторинг и прогнозирование является одним из самых массовых с точки зрения охвата пользователей направлений по анализу пространственных данных. При этом, как и любое основанное на анализе данных направление, сейчас оно претерпевает существенные изменения под влиянием 2 факторов: экспоненциальный рост данных за счет появления большого объема различных источников, предоставляющих разнородные данные (различная точность и достоверность), и рост технологических

решений, позволяющих значительно увеличить как аналитические возможности, так и вычислительные [1].

Системы, реализующие такого рода направление, относятся к мониторинговым ГИС, так как обеспечивают регулярный анализ состояния определенной территории. Однако такие системы имеют ряд важных особенностей, которые важно учитывать при их построении. К таким особенностям можно отнести:

- 1) особенности приема и обработки данных низкого пространственного разрешения,
- 2) учет данных, получаемых наземными станциями,
- 3) использование прогностических моделей.

Таким образом, для метеорологических ГИС становится наиболее важным подход, предполагающий комплексирование и высокую скорость обработки данных и внедрение технологий искусственного интеллекта, направленных на эффективный анализ и последующую предиктивную аналитику.

В работе будет проведен анализ существующих систем, используемых для метеорологического и экологического прогнозирования, рассмотрены особенности данных дистанционного зондирования (ДЗЗ), используемых такими системами.

2 МАТЕРИАЛЫ И МЕТОДЫ

2.1 Данные ДЗЗ, используемые в метеорологических ГИС

Важной особенностью данных ДЗЗ, используемых для гидрометеорологического и экологического мониторинга, является минимизация временного интервала с момента выполнения съемки космическим аппаратом (КА) до начала анализа данных. Именно поэтому для таких систем, как правило, используются приемные комплексы данных ДЗЗ.

Одним из наиболее распространенных в РФ является приемный комплекс X-диапазона «Унискан-36». Он позволяет вести прием и обработку информации, передаваемой с низкоорбитальных искусственных спутников Земли, по радиоканалам диапазона 8 ГГц с темпом до 120 Мбит/с в радиусе 3500 км. «Унискан-36» позволяет вести ежедневный прием широкого спектра данных дистанционного зондирования от высокого до низкого пространственного разрешения: как оптические, так и радиолокационные данные. Их можно принимать с искусственных спутников Земли, таких как Terra (сканер Modis), Aqua (сканер Modis), SuomiNPP, Метеор-M2, SPOT, LANDSAT, FORMOSAT, IRS, CARTOSAT, Resourcesat, EROS, THEOS, RADARSAT-1, RADARSAT-2, ENVISAT-1, TerraSAR-X и COSMO-SkyMed и др.

2.2 Характеристики спутников Terra и Aqua

В то же время, говоря о гидрометеорологическом и экологическом мониторинге, необходимо использование орбитальных систем, позволяющих получать космические снимки, которые имеют большую полосу захвата, покрывающую весь регион, а также имеют широкий спектральный диапазон и высокую частоту обновления данных. На сегодняшний день наиболее популярными космическими аппаратами, обеспечивающими эту потребность, являются Terra и Aqua с аппаратурой MODIS.

Мультиспектральные данные, получаемые этими космическими аппаратами, используются во многих сферах, в том числе для:

- исследования облачного покрова, концентрации взвешенных частиц, распределения водяного пара в атмосфере, мониторинга опасных атмосферных явлений;
- мониторинга природных и антропогенно-спровоцированных катастроф на региональном и глобальном уровнях (наводнения, цунами, извержения вулканов и т.п).

Для тематической обработки данных ДЗЗ MODIS, как правило, используются инструменты тематической обработки данных ДЗЗ. При этом для этих данных уже реализованы инструменты автоматического расчета масок – слоя раstra с заранее просчитанными значениями яркости для различного пространственного разрешения по определенному алгоритму:

- маска LST, расчет температуры земной поверхности, разрешение 1000 м/пиксель;
- маска Clouds, детектирование облачности, разрешение 1000 м/пиксель;
- маска Full Clouds, детектирование облачности, разрешение 1000 м/пиксель;
- маска Fires, детектирование пожаров, разрешение 1000 м/пиксель;
- маска Land.Snow, детектирование снежного покрова суши, разрешение 1000 м/пиксель;
- маска Sea.Ice, детектирование ледового покрова акваторий, разрешение 1000 м/пиксель.

Для большинства операций, связанных с тематической обработкой для целей метеорологического мониторинга, используются каналы с разрешением 1 километр на пиксель [2]. Помимо приемных комплексов X-диапазона для подобных систем активно используются более доступные комплексы L-диапазона, например, позволяющие принимать данные с КА NOAA, Himawari и Meteosat.

2.3 Характеристики спутника Meteosat

Meteosat – серия геостационарных спутников, на высоте 35800 километров. КА получает изображения с сенсоров VIS,

IR и VA каждые полчаса, что позволяет эффективно решать задачи метеорологического мониторинга.

Размер изображения с данного КА составляет 3712×3712 пикселей, что позволяет проводить анализ значительных территорий, так как пространственное разрешение в кадре составляет 3 км и уменьшается до 12 км на краях сцены снимка¹.

2.4 Характеристика спутника NOAA

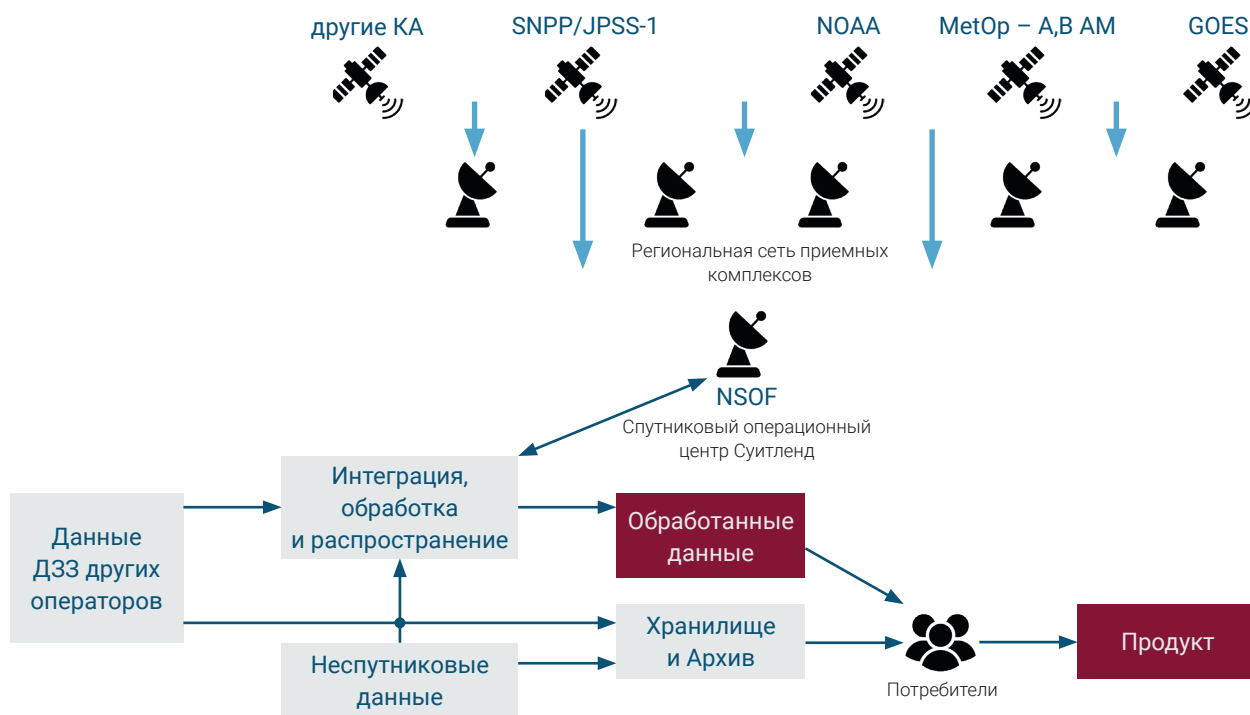
К орбитальной группировке NOAA относится 11 спутников, из них для гидрометеорологического мониторинга используются:

- 5 геостационарных (GOES-14, -15, -16, -17 и -18);
- 5 полярно-орбитальных (NOAA-15, -18, -19, -20, и -21).

Так, например, пространственное разрешение КА NOAA-19 составляет 1 км/пиксель при ширине полосы съемки 2900 км, период повторной съемки 1–2 раза в сутки и 6 спектральными каналами. Для GOES-17 пространственное разрешение составляет от 500 м/пиксель до 2 км/пиксель (в зависимости от канала) с периодом съемки 5–15 минут и 16 спектральными каналами².

Так данные NOAA являются основой для сервиса поставки метеорологических данных Madis³.

Рис. 1
Наземная сеть,
объединяющая орбитальную
группировку КА.



1 Спутниковый метеосат [Электронный ресурс]. Режим доступа: <https://www.meteorologiaenred.com/ru/satellite-meteosat.html> (дата обращения: 12.12.2022)

2 National Environmental Satellite Data and Information Service [Электронный ресурс]. Режим доступа: <https://www.nesdis.noaa.gov/about/our-offices/office-of-satellite-ground-services> (дата обращения: 12.12.2022)

3 MADIS – Meteorological assimilation data ingest system [Электронный ресурс]. Режим доступа: <https://madis.noaa.gov/> (дата обращения: 12.12.2022)

3 ОСОБЕННОСТИ СОВРЕМЕННЫХ СИСТЕМ, ОБЕСПЕЧИВАЮЩИХ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ И ЭКОЛОГИЧЕСКИЙ МОНИТОРИНГ

3.1 Meteum



Рис. 2

Источники для уточнения и комплексирования данных технологии Meteum.

Технология Meteum является основной частью проекта «Яндекс.Погода», ключевой особенностью которого является синтез машинного обучения, данных погодного прогнозирования нескольких независимых поставщиков и данных метеостанций.

Технология основывается на пяти источниках прогноза из национальных центров США, Канады, Японии и ЕС, а также собственный прогноз, рассчитываемый на основе модели с открытым исходным кодом WRF (Weather Research and Forecasting Model).

В качестве первоначальной использовалась модель машинного обучения на базе MatrixNet (CatBoost). Модель учитывала не только погодные параметры из доступных источников, но и прописанные разработчиками дополнительные параметры, которые могут влиять на погоду, такие как высота солнца над горизонтом, удаленность от водоема и календарная дата.

Для обучения использовались архивные данные с профессиональных метеостанций. То есть модель обучалась на архивных данных, формируя зависимости, и на основе актуальных данных создавала прогноз.

Модель давала низкий процент точности для мест с нерегулярной сетью станций. Это было особенно значительно при формировании прогноза осадков, где очень важна частота измерений, а также технологическая особенность учета факта осадков самой станцией, при которой реальные осадки могут быть не учтены (то есть факт дождя может быть не зафиксирован).

Именно поэтому для прогноза осадков на ближайшие часы используются косвенные измерения посредством метеорологических радаров и данные ДЗЗ с геостационарных КА. Это позволяет получать данные на больших территориях раз в 10–15 минут, хотя анализ данных и их последующая интерпретация требуют значительных ресурсов.

Так в центральной части РФ используются данные с метеорологических радаров Росгидромета. Проводится обработка данных, которые позволяют получать классифицированные растры, отображающие зоны осадков в радиусе 250 километров от точки установки радара с пространственным разрешением 2 километра на пиксель, с частотой получения данных 10 минут⁴.

⁴ Хабр [Электронный ресурс]. Режим доступа: <https://habr.com/ru/company/yandex/blog/565238/> (дата обращения: 12.12.2022)

С целью повышения качества результатов авторы данного проекта внедрили нейронные сети. В качестве входных данных использовались мультиспектральные 11-канальные оптические данные ДЗЗ, а обучение проводилось на радарных данных, интерполированных на сетку 2×2 километра. Использовались две архитектуры, основанные на ResNet-подобных и U-Net-подобных моделях.

В качестве основной была выбрана U-Net — архитектура сверточной нейронной сети, использующаяся в задачах сегментации изображений. Ее важной особенностью является возможность применять обученную модель не поточечно, в отличие от архитектуры ResNet, за счет чего значительно вырастает скорость обработки данных ДЗЗ.

При работе нейросети, помимо выделения облаков на сцене снимка, важно определить тип облака. Для этого необходимо учитывать также угол солнца над горизонтом, рельеф, прогноз на ближайшие часы из метеорологической модели (влагозапас облаков, влажность в атмосфере на различных уровнях и т.п.).

При этом для мест с плотной сетью метеорологических радаров учитывались показатели, получаемые ими. Командой проекта была проделана трудоемкая работа, так как необходимо было провести согласование данных этих спутников и радаров по времени и отобразить их корректно на границе видимости радаров.

Чтобы привести 15-минутные спутниковые снимки к 10-минутным интервалам радаров, использовалась Optical Flow для создания промежуточных кадров между последовательными спутниковыми снимками. На основе двух снимков строилось поле векторов смещения изображения в каждой точке. С помощью векторов переноса создавались промежуточные кадры для приведения к единой временной шкале с 10-минутными интервалами.

При объединении разнородных данных возможно периодическое появление артефактов, например, на границе видимости радаров не детектируются осадки, и здесь лучше всего отсутствующие данные замещают оптические спутниковые снимки, по которым данное природное явление можно выявить.

Для уточнения данных у клиентов есть возможность указать в пользовательском приложении актуальное состояние погоды (наличие/отсутствие осадков). Это делается для того, чтобы создать измерительную сеть, основанную на краудсорсинге. Первоначально данный функционал использовался для накопления, статистики и использования в качестве дополнительного контроля, а теперь эти данные полноценно участвуют в аналитике. То есть с помощью библиотеки CatBoost берется прогноз погоды, спутниковые измерения, радары и другие влияющие на погоду параметры, а в качестве целевых значений — информация о дожде от пользователей, после чего проходит обучение CatBoost на таком датасете, а затем применяется обученную модель, чтобы получить карту осадков.

На текущий момент технология Meteum позволяет:

- 1) интегрировать данные с радаров и спутников в общую систему прогноза наравне с метеомоделями;

2) реализовывать обучение модели, чтобы она предсказывала не показания метеостанций, а сообщения пользователей о погоде. При этом сформирован рейтинг данных сообщений, так как информация с высоким кредитом доверия используется для корректировки ошибочной карты осадков в реальном времени, что позволяет улучшить точность карты до 20 % в случае внезапного дождя.

Использование краудсорсинга позволяет получать более миллиона сообщений об осадках в день (в особо дождливые дни — более 2,5 миллионов). При этом метеостанции обеспечивают около 8 тысяч наблюдений в день⁴.

примечание | Далее существующие системы будут описаны кратко, так как приведенный выше фрагмент о Meteum является примером подробного описания устройства метеорологической ГИС и подхода по комплексированию.

3.2 Национальная метеорологическая служба NOAA (NWS)

С 1870 г. NWS (Национальная метеорологическая служба) обеспечивала метеорологические наблюдения за военными базами и побережьем. Более чем 150-летний опыт улучшил их способность прогнозировать погоду и регистрировать метеорологические явления.

Данные Национальной метеорологической службы в основном доступны в виде шейп-файлов и языка разметки keyhole (KML). Их данные о погоде включают:

- осадки, снегопады и температуру;
- наводнения, ураганы, пожары и связанные с ними засухи.

NWS также имеет большую базу исторических данных (начиная с 1950 г.) с информацией об осадках и погоде/климате.

В июле 2021 года NWS начала экспериментальную фазу веб-приложения Экспериментальная национальная ГИС. Приложение предоставляет возможности для отображения, опроса, доступа и обнаружения данных и легко предоставляет эту информацию, чтобы те, кто знаком с геопространственными данными и технологиями, могли беспрепятственно получать к ней доступ и анализировать ее. Слои данных включают в себя все, от быстро обновляемых продуктов (например, радаров) до более традиционных (например, прогнозов) и от часов и предупреждений до эталонных границ³.

3.3 Система приема данных метеорологического моделирования (MADIS)

MADIS расшифровывается как «система приема данных метеорологического моделирования». Это похоже на расширение NOAA, потому что оно объединяет данные с нескольких станций, включая федеральные, государственные, университеты, волонтеров и частный сектор.

Комбинируя данные из нескольких источников, они преобразуются в набор данных с большей плотностью и более высокой частотой. Данные о погоде MADIS GIS содержат данные в реальном времени и данные наблюдений за поверхностью земли (температура, ветер, относительная влажность, осадки) и радиозондов (метеозондов)³.

PRISM (ПРИЗМА) расшифровывается как «регрессия высоты параметра на модели независимых склонов». В этой программе используются точечные измерения осадков, температуры, других климатических факторов и цифровая модель рельефа. Кроме того, она использует экспертные знания для построения непрерывной сетки климатической модели⁵.

3.4 Система прогнозирования погоды AccuWeather

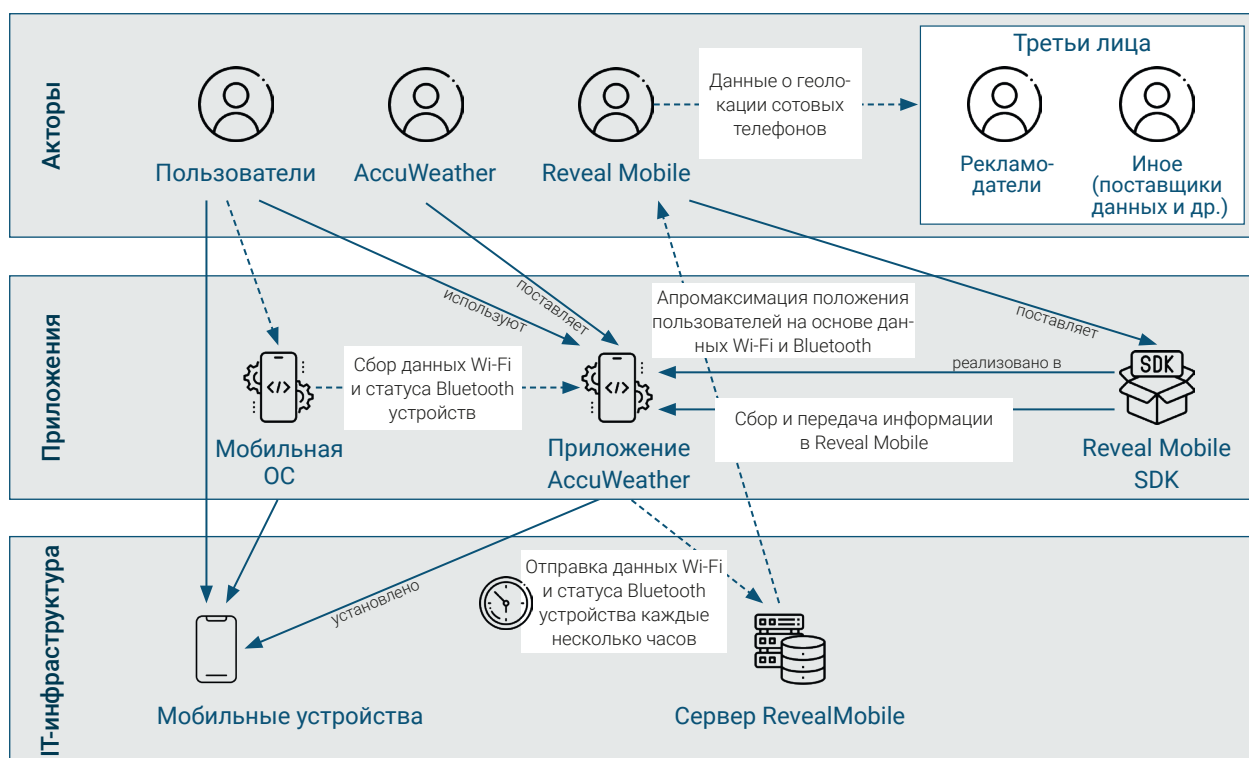


Рис. 3
 ▲
 Архитектурная карта AccuWeather, отправляющая данные устройства в Reveal Mobile.

Одним из самых популярных зарубежных решений является AccuWeather. Одной из его важных функциональных частей является аппроксимация местоположений пользователей путем агрегирования нескольких типов данных, даже если пользователи запретили доступ к своим данным о местоположении. В таком случае приложение AccuWeather получает данные об их устройствах на основе расположения сотовых вышек с помощью технологии

⁵ GIS Geography. 4 weather data sources for GIS [Электронный ресурс]. Режим доступа: <https://gisgeography.com/gis-weather-data-sources/> (дата обращения: 12.12.2022)

компании Reveal Mobile. Так SDK Reveal Mobile, реализованный в приложении AccuWeather, иницирует сбор и передачу данных устройства. Затем приложение AccuWeather непрерывно собирает статус Bluetooth-устройства, а также имя и уникальный BSSID (базовый идентификатор набора услуг) используемого маршрутизатора Wi-Fi из операционной системы мобильного устройства и отправляет эти данные в Reveal Mobile каждые несколько часов. Затем, обобщая данные, Reveal Mobile приблизительно определяет местонахождение пользователей и продает эту информацию третьим сторонам [3].

Еще одним нетривиальным функционалом в системах гидрометеорологического мониторинга является применение мобильных динамических геосенсоров. Мобильная метеостанция устанавливается на транспортное средство, которое постоянно ездит по определенной территории для сбора данных и их отправки на разные приемники по проводным или беспроводным технологиям, что приводит к значительно лучшему охвату, чем у стационарных станций. Городской автобус используется в качестве мобильной метеостанции для сбора данных о пути, пройденном транспортным средством. Система имеет три уровня взаимодействия: устройство интегрировано в шины, терминальный компьютер и системный компьютер.

Так в одном из городов используют данные об автомобилях для прогнозов погоды на дорогах, что направлено на повышение разрешения сети наблюдения за погодой и модели прогноза. С учетом влияния каналов связи на качество информации были проведены работы, в которых анализировалась корреляция между несколькими погодными переменными и поведением контрольных кадров в наружной беспроводной локальной сети, чтобы определить влияние погоды на производительность радиолинии. Основываясь на системе управления автобусной информацией, удалось совместить преимущества локальной обработки информации и автобусной мобильности, что обеспечило систему мониторинга погоды в реальном времени [4].

5 РЕЗУЛЬТАТЫ И ВЫВОДЫ

Как показывает обзор современных систем, весь процесс их выстраивания построен на подходе по комплексированию и уточнению исходных метеорологических данных ДЗЗ с использованием всех возможных данных: от данных ДЗЗ более высокого разрешения до использования разветвленной сети геосенсоров различных типов и в некоторых случаях использования инструментарию краудсорсинга.

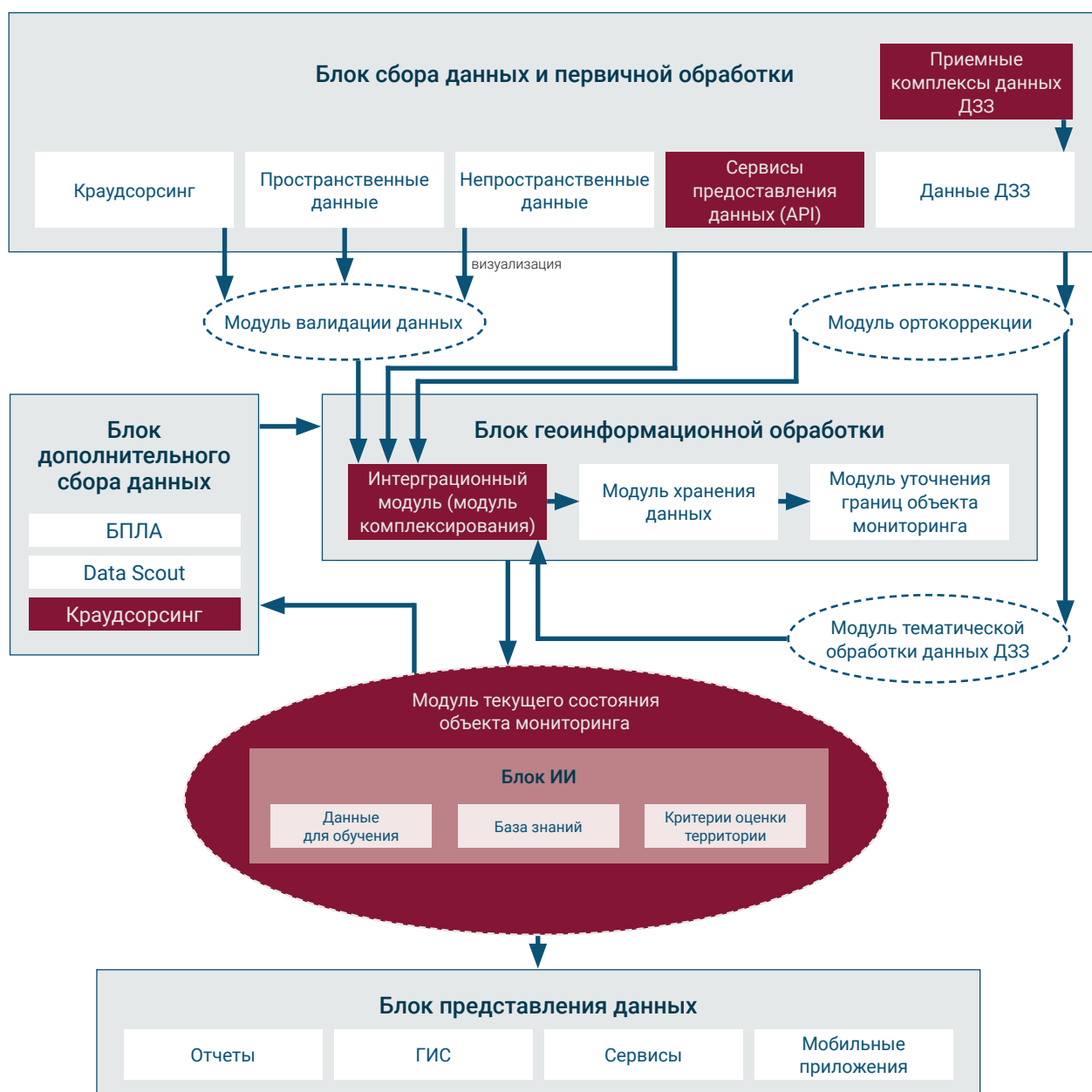
Принципиальную схему построения таких систем можно представить как мониторинговую ГИС (Рис. 4), предполагающую комплексирование пространственных данных и внедрение технологий, позволяющих значительно ускорить процесс обработки и анализа

данных [5], либо в целом использовать прогрессивные технологические решения, позволяющих значительно расширить вычислительные мощности, как например, Microsoft Azure.

При этом важной особенностью подобных систем становится использование различных API, позволяющих получать доступ к текущим, прогностическим и архивным погодным данным (температура, влажность, скорость ветра, осадки, облачный покров, видимость и проч.). К наиболее популярным API относят:

- Tomorrow.io API;
- OpenWeatherMap;
- Meteogroup;
- Weatherstack;
- Weatherbit;
- Weather2020;

Рис. 4
Концептуальная модель системы геоинформационного мониторинга.



- AerisWeather;
- Accuweather;
- Visual Crossing⁶.

Также важной особенностью является переход от экспертной системы к системе анализа и последующей поддержке принятия решений на основе искусственного интеллекта (ИИ), а также появление в блоке дополнительного сбора данных элемента краудсорсинга как инструмента дополнительной верификации результатов анализа пространственных данных [1, 6–8].

БЛАГОДАРНОСТИ

Исследования выполнены в рамках госзадания Минобрнауки РФ № FSFE-2022-0002.

БИБЛИОГРАФИЯ

1. Быстров А.Ю. Обзор современных теорий и принципов построения мультифункциональных динамических мониторинговых геоинформационных систем / А.Ю. Быстров, А.А. Майоров // Известия высших учебных заведений. Геодезия и аэрофотосъемка. 2021. Т. 65. № 1. С. 108–116. DOI:10.30533/0536-101X-2020-65-1-108-116 EDN: OCXEF.
2. Быстров А.Ю. Возможности применения данных MODIS, принимаемых в МИИГАиК, для регионального мониторинга водоохранных территорий / А.Ю. Быстров, А.В. Гречищев // Экология. Экономика. Информатика: Сборник статей в 3 томах / Российский фонд фундаментальных исследований, Южный федеральный университет, Институт математики, механики и компьютерных наук имени Воровича И.И., Институт аридных зон, Южный научный центр Российской академии наук. Том 3. Ростов-на-Дону: Южный федеральный университет, 2015. С. 219–224. EDN VNOVVJ.
3. Burmeister F., Drews P., Schirmer I. Modeling the C(ourse) of Privacy-critical Location-based Services - Exposing Dark Side Archetypes of Location Tracking // Conference: HICSS. 2021. DOI:10.24251/HICSS.2021.799.
4. Huang Z.Q., Chen Y.C., Wen C.Y. Real-Time Weather Monitoring and Prediction Using City Buses and Machine Learning // Sensors (Basel). 2020. Vol. 20(18). 21 p. DOI:10.3390/s20185173.
5. Гвоздев О.Г., Майоров А.А., Матерухин А.В., Бондарев И.Н. Проблема вычисления пространственных предикатов вида «находится ли точка в полигоне» над потоками пространственно-временных данных высокой интенсивности // Известия вузов «Геодезия и аэрофотосъемка». 2022. Т. 66. № 2. С. 34–47. DOI:10.30533/0536-101X-2022-66-2-34-47.
6. Wilkinson S., Dunn S., Adams R., et al. Consequence forecasting: A rational framework for predicting the consequences of approaching storms // Climate Risk Management. 2022. Vol. 35. DOI:10.1016/j.crm.2022.100412.

⁶ Tomorrow.io [Электронный ресурс]. Режим доступа: <https://www.tomorrow.io/blog/top-8-weather-apis-for-2022/#ven8> (дата обращения: 12.12.2022)

7. Zar Chi Aye, Sprague T., Cortes V.J., et al. A collaborative (web-GIS) framework based on empirical data collected from three case studies in Europe for risk management of hydro-meteorological hazards // International Journal of Disaster Risk Reduction. 2016. Vol.15. P. 10–23. DOI:10.1016/j.ijdrr.2015.12.001.
8. Touati N., Gardes T., Hidalgo J. A GIS plugin to model the near surface air temperature from urban meteorological networks // Urban Climate. 2020. Vol. 34. DOI:10.1016/j.uclim.2020.100692.

АВТОРЫ

Быстров Антон Юрьевич

ООО «СберОбразование»
кандидат технических наук

 0000-0001-6892-1584

Майоров Андрей Александрович

Московский государственный университет геодезии и картографии,
Москва, Россия

Кафедра информационно-измерительных систем
Факультет геоинформатики и информационной безопасности
доктор технических наук
профессор

 0000-0001-5539-3719

Поступила 25.11.2022. Прошла рецензирование 10.12.2022. Принята к публикации 29.12.2022.