

ПЕРВИЧНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ОСНОВНЫХ МЕТЕОРОЛОГИЧЕСКИХ И РАДИАЦИОННЫХ ДАННЫХ ПО ТОМСКОЙ ОБЛАСТИ

К.А. Лемешонок, И.А. Ботыгин
Томский политехнический университет
kal9@tpu.ru

Введение

В настоящее время прогнозирование и анализ метеорологических данных является актуальной проблемой, так как человек активно взаимодействует с окружающей средой [1, 2]. Существует множество различных ресурсов, предоставляющих хронологическую последовательность измерений различных метеорологических характеристик [3]. Одним из таких ресурсов является портал Томского Института мониторинга климатических и экологических систем СО РАН, предоставляющий различные метеорологические и радиационные данные по Томской области.

Одним из наиболее востребованных направлений в обработке измерений является их статистический анализ [4].

В настоящей работе осуществлена разработка программного обеспечения для визуализации, первичной статистической обработки и частотно-временного анализа метеорологических и радиационных данных различных метеостанций Томской области.

Описание алгоритма

Для мониторинга основных метеорологических и радиационных данных по Томской области спроектирована и реализована сеть аппаратно-программных пунктов наблюдения. Данная сеть включает в себя множество метеорологических станций, расположенных в Томской области. Этими пунктами наблюдения производится измерение следующих характеристик: уровень осадков, температура окружающей среды, атмосферное давление, скорость и направление ветра, влажность воздуха, поток фотоактивной радиации, влажность почвы и многие другие метеорологические параметры. Измерения происходят в автоматическом режиме и передаются на сервер хранения данных.

Обобщенная функциональная структура разработанного программного обеспечения включает в себя: пользовательский интерфейс, блок взаимодействия с сервером данных, блок обработки данных, блок экспорта/импорта данных.

Для реализации данного программного обеспечения был выбран язык Python. Выбор языка обусловлен широким выбором библиотек для статистической обработки данных, а также их визуализации.

Пользовательский интерфейс позволяет осуществить выборку необходимых данных для дальнейшей обработки и визуализации. В частности, предоставляется возможность указать дополнительные параметры выборки, такие как: станция,

серийный номер датчика, дата начала и окончания измерений, строить графики по различным характеристикам набора данных, масштабировать их, сдвигать по осям, а также сохранять на жестком диске в графическом формате PNG (рис. 1).

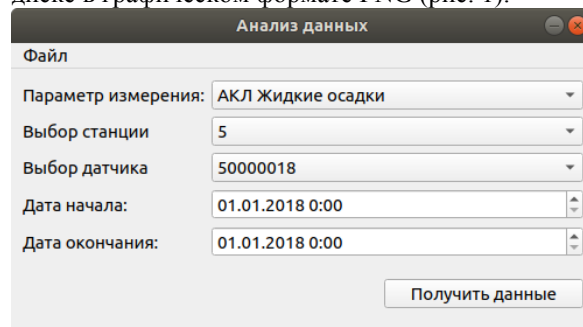


Рис. 3. Окно выбора параметров

При выборке необходимых данных, на основе указанных пользователем параметров, формируется запрос к серверу данных. Ответ сервера представляет собой набор данных в формате CSV, где строки – это текущие измерения метеорологических и радиационных данных.

Полученные от сервера данные преобразовываются в структуру, удобную для хранения и обработки.

После обработки данных перед пользователем появляется окно с выбором необходимого параметра для отображения. При выборе параметра происходит построение графика зависимости выбранной величины от времени, указанного в начальном окне выбора параметров (рис. 2).

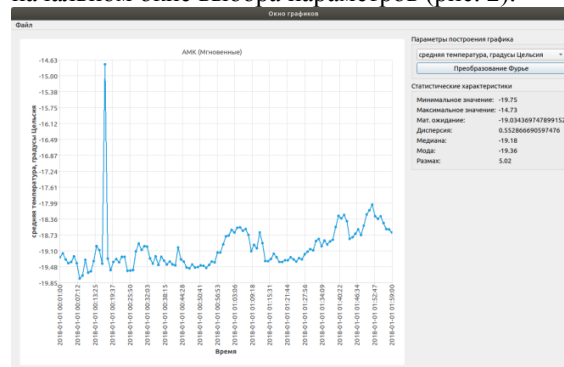


Рис. 4. Окно отображения характеристик

Первичная статистическая обработка включает в себя вычисление следующих характеристик: минимальное, максимальное, среднее значение, дисперсия, мода, медиана и размах исследуемого ряда наблюдений.

Практически все реальные физические сигналы, полученные с датчиков, изменяются во времени.

Наиболее популярным методом анализа частотных составляющих входного сигнала является преобразование Фурье [5]. В нашем случае для анализа ряда использовалось быстрое преобразование Фурье (FFT) – это алгоритм быстрого вычисления прямого дискретного преобразования Фурье.

Помимо этого, предоставляется возможность визуального отображения полученных коэффициентов для более детального анализа измерений (рис. 3).

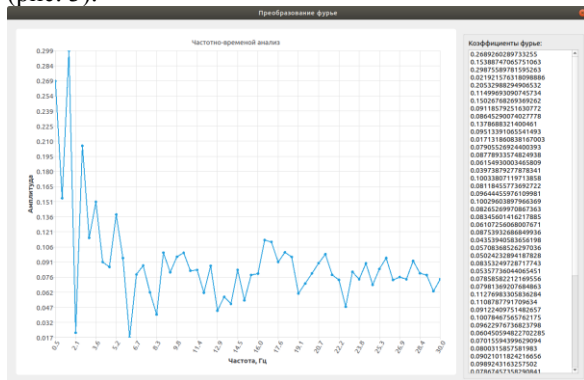


Рис. 5. Окно отображения результатов преобразования Фурье

Вычисление основных статистических характеристик осуществлялось с использованием пакетов numpy и SciPy. NumPy – это библиотека языка Python, добавляющая поддержку больших многомерных массивов и матриц, вместе с большой библиотекой высокоуровневых математических функций для операций с этими массивами. SciPy – библиотека для языка программирования Python с открытым исходным кодом, предназначенная для выполнения научных и инженерных расчетов. [6].

В целях обеспечения обработки измеряемых параметров наблюдений сторонними программными средствами реализован экспорт данных. Экспорт позволяет предоставить данные для использования и анализа с помощью сторонних приложений, а также использовать их в дальнейшем в разработанном программном обеспечении. Экспортируемый файл состоит из двух частей: метаданная, данные измерений. Метаданные содержат в себе название величины измерения, идентификатор станции и серийный номер датчика, с помощью которого производились измерения.

Также для обработки возможен ввод ранее экспортированных данных. Импорт данных является важной составляющей, так как производить запросы к серверу с целью загрузки данных за большой интервал времени задача трудоемкая и очень

затратная по времени. Поэтому за счет импорта ранее сохраненных данных указанный недостаток устраняется.

Для разработки и отладки программного обеспечения использовалась интегрированная среда разработки PyCharm.

Заключение

В результате проделанной работы было разработано программное обеспечение, позволяющее динамически выгружать метеорологические данные по Томской области, производить их первичную обработку, структурировать, визуализировать для более наглядного представления, а также производить их первичный статистический и частотно-временной анализ.

Список использованных источников

1. Воронин С.В., Ухабов С.С. Анализ эффективности существующей системы мониторинга и прогнозирования чрезвычайных ситуаций (на примере Тверской области) // Научно-аналитический журнал «Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России». 2015. №1. // URL: <https://cyberleninka.ru/article/n/analiz-effektivnosti-suschestvuyushey-sistemy-monitoringa-i-prognozirovaniya-chrezvychaynyh-situatsiy-na-primere-tverskoy-oblasti> (дата обращения: 12.11.2018).
2. Борисова А. С. Методы анализа и прогноза климатических временных рядов // Известия вузов. Северо-Кавказский регион. Серия: Естественные науки. 2008. №1. URL: <https://cyberleninka.ru/article/n/metody-analiza-i-prognoza-klimaticheskikh-vremennyh-ryadov> (дата обращения: 12.11.2018).
3. National Centers for Environmental Information // URL: <https://www.ncdc.noaa.gov/> (дата обращения 10.11.2018)
4. Шорохова И.С., Кисляк Н.В., Мариев О.С. Статистические методы анализа: [учеб. пособие]. – Екатеринбург: Изд-во Урал. ун-та, 2015. – 300 с.
5. Частотно-временной анализ нестационарных сигналов // Zetlab. URL: <https://zetlab.com/chastotno-vremennoy-analiz-nestatsionarnyih-signalov-v-programmnom-obespechenii-zetlav/> (дата обращения 13.11.2018)
6. SciPy.org // URL: <https://www.scipy.org/> (дата обращения 14.11.2018)