

Статистический анализ указал на свалку мусора

Алгоритм, позволивший вычислить источник неприятного запаха на основе только жалоб местных жителей, разработала группа российских исследователей, в числе которых сотрудники Института космических исследований РАН. Авторы алгоритма опробовали его в действии, проанализировав выборку жалоб жителей востока Московской области в сентябре 2013 — июне 2014 года. В результате обработки данных с учетом розы ветров и других особенностей местного ландшафта исследователи пришли к выводу, что источником загрязнения с высокой степенью вероятности стала местная свалка.

Интересен при этом не только сам результат, но и тот факт, что статистические методы обработки неформализованных данных способны давать хорошие результаты. А значит, можно использовать подобные алгоритмы для быстрой оценки экологической ситуации, и лишь после этого прибегать к более сложным методам с использованием данных спутников или специальных полевых исследований.

Задачу найти неизвестный источник неприятного запаха, на который жаловались жители восточных районов Московской области, поставила группа исследователей из Российской медицинской академии последиplomного образования, Института космических исследований РАН и Института народнохозяйственного прогнозирования РАН. При этом, фактически, единственным источником информации о происходящем стали жалобы местных жителей, с уточнениями, где и когда эти жалобы были зафиксированы. Но ни природа загрязнителя, ни его концентрация не были известны.

«Хотя подобные задачи — поиски источников загрязнения — сейчас очень актуальны, насколько нам известно, это первая работа, где предпринята попытка найти источник загрязнения с помощью только математических методов, используя жалобы местных жителей. Эта информация гораздо более «неопределенна», чем данные точных измерений, — говорит **Борис Балтер**, сотрудник Института космических исследований РАН, один из авторов статьи, принятой к публикации в журнале «Air Quality, Atmosphere & Health».

Этот процесс авторы сравнили с поиском источника света с помощью линзы, которая фокусирует идущие от него лучи. В данном случае «лучи» — выборка жалоб людей, проживающих на востоке Московской области (общая площадь территории, с которой собиралась информация, — приблизительно 20х30 км, около 30 км от Москвы) вдоль Горьковского шоссе с сентября 2013 по июнь 2014 года. А собирающей «линзой» стал довольно сложный процесс обработки и анализа этих данных с учетом того, как мог бы распространяться неизвестный загрязнитель.

Источником запаха, как предположили авторы, должен быть довольно низко расположенный объект с температурой окружающей среды (а не хорошо заметная высокая труба, выпускающая горячий дым). Это означает, в том числе, что количество жалоб примерно соответствует концентрации вещества в воздухе, которое, в свою очередь, зависит от ветра, и поэтому в исследовании использовались данные метеорологических станций вблизи исследуемого района, и учитывалось предположительное время, которое требуется для распространения от источника к людям. Чтобы восстановить предположительную концентрацию, использовался специальный алгоритм AERMOD, а также специальное программное обеспечение, разработанное в ИКИ РАН.

ИНСТИТУТ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ

РОССИЙСКОЙ АКАДЕМИИ НАУК

СООБЩЕНИЕ ПРЕСС-СЛУЖБЫ

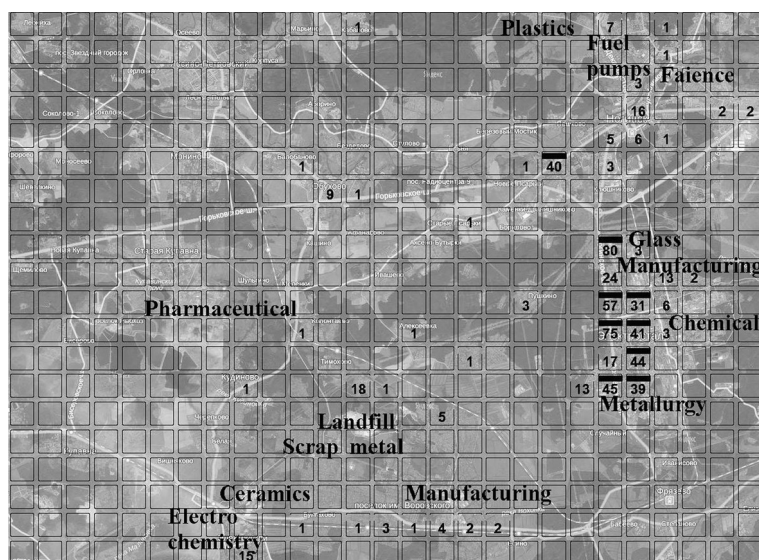
Территория исследуемого района была разделена на ячейки размером 1 квадратный километр, и каждой из них было присвоено некоторое значение, полученное с помощью статистических методов из количества жалоб и предположительной концентрации зловонного вещества. Можно сказать, что каждая ячейка рассматривалась как потенциальный источник запаха, а присвоенные ей значения определяли вероятность этого. Затем, опять-таки используя статистические алгоритмы обработки данных, были определены участки, где наиболее вероятно расположение источника. На диаграммах эти участки выглядят как концентрические овалы, и центральный овал очерчивает территорию, где, скорее всего, и находится загрязнитель. Если посмотреть на реальную карту, то именно здесь — на участке диаметром 1–2 км — находится мусорная свалка.

Впрочем, поскольку мы имеем дело не с данными измерений, а довольно неформальной информацией и множеством допущений, то как проверить правильность выводов? Один из способов — «поиграть» разными параметрами: временем распространения запаха, способами обработки метеорологических данных и другими — и посмотреть, как это влияет на вывод. Авторы много раз прогнали модель, и оказалось, что результат довольно стабилен, а «фокус» такой математической «линзы» сдвигается всего на 1 км.

«Наша работа показала, что статистические методы обработки данных могут оказаться исключительно полезными для поиска источников загрязнения, о которых очень мало информации, — говорит Борис Балтер. — Это особенно важно для развивающихся стран с активной промышленностью. Но, конечно, делать окончательные выводы можно лишь на основе точных измерений».

Это исследование можно продолжать в разных направлениях. Например, оценить активность источника запаха в граммах в секунду, сопоставив результаты моделирования с прямыми измерениями концентрации вредных веществ. Другая возможность — попытаться изучить «порог чувствительности», после которого начинаются жалобы, и затем «превратить» его в показатели концентрации.

Работа выполнена по заданию Министерства образования и науки Российской Федерации.



ИНСТИТУТ КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ

РОССИЙСКОЙ АКАДЕМИИ НАУК

СООБЩЕНИЕ ПРЕСС-СЛУЖБЫ

Рис. 1. Карта изучаемой местности, с нанесенной «сеткой» (размер ячейки 1х1 км) и указанием основных производств и возможных источников запаха. Числа в ячейках — полное количество жалоб за 10 месяцев. Рис. из ст. [1]

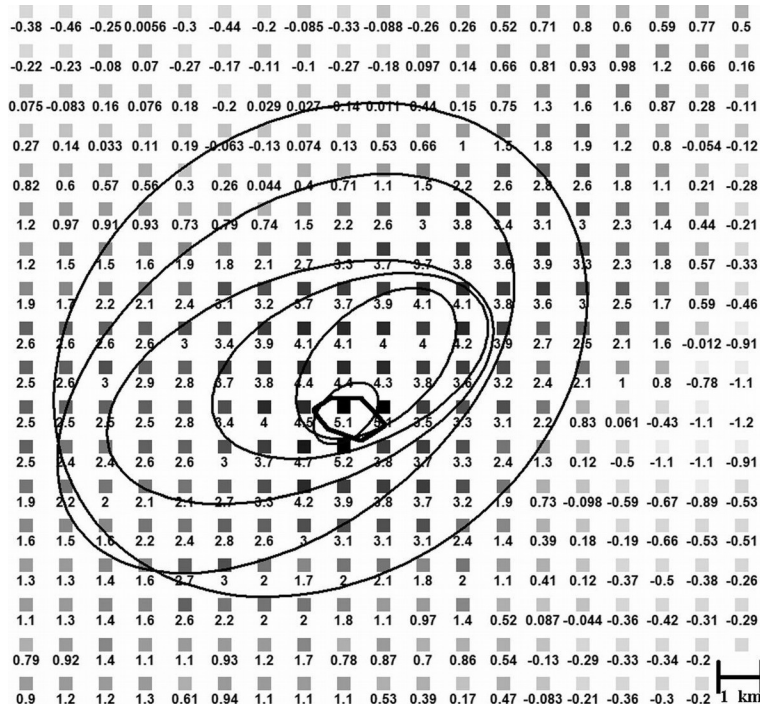


Рис. 2. Так выглядит карта, на которой овалами очерчены районы с одинаковой вероятностью нахождения источника загрязнения. Шестиугольником обозначено расположение свалки. Рис. из ст. [1]

1. S. L. Avaliani & B. M. Balter & D. B. Balter & M. V. Faminskaya & B. A. Revich & M. V. Stalnaya Air pollution source identification from odor complaint data, **Air Quality, Atmosphere & Health**, опубликована онлайн 15 января 2015
<http://link.springer.com/article/10.1007/s11869-015-0317-8>