

## Статья

# Краудсенсинговая архитектура IoT для всеобщего качества воздуха и Мониторинг экспозиции: проектирование, разработка, калибровка и Долгосрочная проверка

Саверио Де Вито 1,\*<sup>†</sup>, Елена Эспозито 1,\*<sup>†</sup>, Этторе Массера 1, Фабрицио Формисано 1, Грация Фатторузо 1, Серджо Ферлито 1, Антонио Дель Джудиче 1, Херардо Д'Элия 1, Мария Сальвато 1,<sup>†</sup>, Тициана Поликетти 1, Паоло Д'Аурия 2, Адриан М. Ионеску 3 и Джироламо Ди Франсия 1

1 ENEA CR-Portici, TERIN-FSD Division, P. le E. Fermi 1, 80055 Portici, Италия; ettore.massera@enea.it (EM); fabrizio.formisano@enea.it (FF); grazia.fattoruso@enea.it (GF); sergio.ferlito@enea.it (SF); antonio.delgiudice@enea.it (ADG); gerardo.delia@enea.it (GD); maria.salvato@enea.it (MS); tiziana.polichetti@enea.it (TP); girolamo.diffrancia@enea.it (GDF)

2 ARPA Campania, Via Vicinale Santa Maria del Pianto Centro Polifunzionale, Torre 1, 80143 Неаполь, Италия; p.dauria@arpacampania.it

3 EPFL-Ecole Polytechnique Federal de Lausanne, 1015 Лозанна, Швейцария; adrian.ionescu@epfl.ch \* Переписка: saverio.devito@enea.it (SDV); elena.esposito@enea.it (EE) <sup>†</sup> MS работает в ENEA DUEE-SIST Div.

Образец цитирования: Де Вито, С.; Эспозито, Э.; Массера, Э.; Формисано, Ф.; Фатторузо, Г.; Ферлито, С.; Дел Джудиче, А.; Д'Элия, Г.; Сальвато, М.; Поликетти, Т.; Д'Аурия, П.; Ионеску, ЯВЛЯЮСЬ; Ди Франсия, Г. Краудсенсинг Архитектура Интернета вещей для всепроникающего воздуха Мониторинг качества и экспозиции: дизайн, разработка, калибровка и долгосрочная проверка. Датчики 2021, 21, 5219. <https://doi.org/10.3390/c21155219>

Академический редактор: Пьетро Мандзони

Получено: 12 июля 2021 г.

Принято: 20 июля 2021 г.

Опубликовано: 31 июля 2021 г.

Примечание издателя: MDPI остается нейтральным в отношении юрисдикционных претензий в опубликованных картах и институтах. иональная принадлежность.



Авторское право: © 2021 автора. Лицензия MDPI, Базель, Швейцария.

Эта статья находится в открытом доступе распространяется на условиях и против положения Creative Commons Attribution (CC BY) (<http://creativecommons.org/licenses/by/4.0/>).

**Резюме:** всесторонняя оценка качества воздуха в городских или мобильных условиях имеет первостепенное значение для разработки и реализации действий по снижению индивидуального или общегородского воздействия. Возможность развертывания гибридной сети с высоким разрешением нормативного уровня и недорогих стационарных и мобильных устройств является основным фактором развития таких знаний как в качестве основного источника информации, так и для проверки прогнозирования качества воздуха с высоким разрешением. модели. Возможность мониторинга индивидуального воздействия в режиме реального времени и кумулятивного мониторинга также считается основной движущей силой для мониторинга экспозом и будущих подходов к прогностической медицине. Используя опыт химического обнаружения, машинного обучения и Интернета вещей (IoT), мы разработали интегрированную архитектуру, способную удовлетворить жесткие требования этой сложной проблемы. Здесь приводится подробный отчет о процедурах проектирования, разработки и проверки, а также результаты двухлетней проверки в полевых условиях.

**Ключевые слова:** узлы IoT AQ; сенсорная сеть; калибровка; мониторинг качества воздуха; машинное обучение

## 1. Введение

Качество воздуха (КВ) является одним из основных факторов, влияющих на качество жизни в городах [1]. Сельские районы также могут быть затронуты из-за специфических географических условий и связанной с ними местной динамики климата, которые могут негативно сказаться на рассеивании. К сожалению, загрязнители атмосферного воздуха все чаще и достоверно ассоциируются с рядом заболеваний, начиная от легко поддающихся лечению и временных заболеваний и заканчивая тяжелыми и серьезно опасными заболеваниями [2]. Как следствие, количество преждевременных смертей, которые, по оценкам, были непосредственно вызваны плохими условиями AQ, вызывает серьезную озабоченность [3,4]. Таким образом, огромные технологические усилия и политические действия, такие как Закон о чистом воздухе, разрабатываются для улучшения качества воздуха, иногда оказывая значительное влияние на региональном уровне. Несмотря на то, что концентрации некоторых загрязняющих веществ улучшаются благодаря технологическим достижениям, концентрации двуоксида азота (NO<sub>2</sub>) и твердых частиц по-прежнему вызывают озабоченность из-за их потенциальной заболеваемости и тенденции к их накоплению в некоторых условиях (например, городские каньоны, Пьянура-Падана, и др.), создающие угрозу для сообществ регионального или квартального масштаба и определяющие экологическую несправедливость для граждан [4,5].

Источники загрязнения воздуха варьируются от централизованных (например, электростанции, промышленные установки) до слабо распределенных (дорожное движение) и широко распространенных (дорожное движение, отопительные приборы и т. д.). Для адекватного планирования и проверки действий по выявлению и устранению последствий, включая, например, интеллектуальное управление трафиком и/или повышение осведомленности общественности, крайне важно получить всеобъемлющие, всеобъемлющие и количественные знания о AQ [1]. Сети высокоточных регулирующих станций мониторинга качества воздуха (AQMS) развертываются по всему миру уполномоченными агентствами по охране окружающей среды (EPA) в соответствии с четкими нормативными рамками, которые направляют их для выбора мест развертывания, целевых газов и технологий контрольно-измерительных приборов. Полученные в результате приборы эталонного класса обычно громоздки и дороги, что требует адекватного и постоянного обслуживания для соответствия нормативным требованиям. Соображения стоимости и требования к местоположению очень часто приводят к разреженным сетям, которые не в состоянии справиться с высокой пространственной и временной изменчивостью явлений, наблюдаемых в городской среде [6]. Мониторинг AQ с высоким разрешением и недавно появившиеся приложения, такие как экспозомный мониторинг (личное воздействие) и распределение источников, следовательно, требуют сквозных сетей AQMS, которые можно развернуть, только прибегая к недорогим, а иногда и портативным устройствам на основе микросенсоров [5]. Эти устройства обычно подключаются через сотовые сети разных поколений или специальные сетевые структуры к серверным системам, в которых происходит обработка данных. Их небольшие размеры, низкая стоимость и низкое энергопотребление позволяют по-настоящему широко использовать мобильные или автономные развертывания.

К сожалению, микросенсоры, далекие от идеальных газовых сенсоров, обычно подвержены нескольким источникам ошибок, среди которых наихудшими являются помехи от нецелевых газов, влияние окружающей среды и старение/загрязнение сенсора [5]. В долгосрочной перспективе они в конечном итоге вызывают снижение точности любого алгоритма калибровки, когда вынуждены делать выводы о концентрациях в условиях, отличных от условий калибровки [7]. Новые недорогие датчики твердых частиц (ТЧ), хотя, как правило, показывают высокие показатели точности, также подвержены влиянию окружающей среды. В результате их необходимо тщательно оценивать с точки зрения точности при длительном полевом развертывании, а процедура калибровки играет первостепенную роль в определении общей производительности и удовлетворении жестких требований к качеству данных [8]. С другой стороны, большинство устройств, доступных на рынке, особенно разработанных для потребительского рынка, продаются без каких-либо гарантий точности. Таким образом, их данные можно считать ограниченно пригодными для большинства наших приложений. Оптимальные процедуры калибровки активно изучаются, при этом полевые данные становятся основным источником информации для получения достаточно точной функции калибровки, способной оценивать концентрации загрязняющих веществ по необработанному сигналу датчика с поправкой на многочисленные факторы, отмеченные выше [9]. Тем не менее, лабораторная калибровка является основным источником информации о поведении датчиков и позволяет проводить контролируемую оценку различных источников помех и характеристик линейности отдельных датчиков. Исследователи все еще пытаются снизить затраты на полевые или лабораторные процедуры калибровки, прибегая к так называемым сетевым процедурам, уменьшая количество регистрируемых в полевых условиях образцов для достижения надлежащего качества калибровки. Машинное обучение на самом деле широко используется для достижения наилучших результатов; было предложено несколько алгоритмов для нескольких датчиков и мягкой калибровки (см. [10]) со смешанными результатами и без появления выигрышной парадигмы [11]. В то же время проводятся среднесрочные и долгосрочные эксперименты по развертыванию для оценки мультисенсорных устройств с течением времени. В этих условиях полилинейная регрессия, неглубокие нейронные сети и случайные леса кажутся наиболее убедительными подходами, которые применялись с положительными результатами [11–13].

Участие граждан и неквалифицированного персонала в настоящее время является обязательным для повышения осведомленности об AQ и достижения требований повсеместного распространения. Многие исследовательские проекты были профинансированы и начаты с заявленной целью повышения осведомленности и обеспечения возможности мониторинга личного воздействия в рамках кампаний гражданской науки [14]. Кампании краудфандинга и краудсенсинга могут на самом деле помочь вовлечь граждан с самого начала проекта всеобъемлющего зондирования, сохраняя при этом обязательства на достаточном уровне во время кампаний мониторинга [15–17].

Напротив, объем данных, генерируемых такой широко распространенной и гибридной сетью мобильных и стационарных устройств, может быть значительным, в то время как число граждан и различных операторов, заинтересованных в различных продуктах, которые могут быть разработаны путем обработки данных, постоянно растет. На самом деле необходимость интеграции таких интеллектуальных мультисенсорных устройств в инфраструктуру IoT воспринимается как очень насущная. Чтобы ответить на эти возникающие потребности, начиная с 2005 г. наша группа присоединилась к усилиям по разработке технологических процедурных и алгоритмических технологий для поддержания развития всепроникающих сетей мониторинга качества воздуха [18].

В этой работе мы сообщаем о результатах проектирования, разработки и проверки архитектуры IoT AQMS под названием MONICA (MONItoraggio Cooperativo della qualità dell'Aria — аббревиатура, которую можно перевести как «Совместный мониторинг качества воздуха»). Архитектура MONICA, направленная на разработку всеобъемлющей (стационарной, мобильной, регулирующей) и коллективной сети мониторинга качества воздуха (AQMN), основана на гибридной сети, включающей портативные недорогие устройства, основанные на массивах электрохимических датчиков и калиброванных счетчиках частиц. Мы начали проект с краудфандинговой кампании, которая позволила нам спроектировать, построить и функционально проверить прототипы недорогого устройства мониторинга качества воздуха (AQMD). Дальнейшие усовершенствования в тонкой настройке устройства и разработка специальных процедур калибровки и мониторинга воздействия были выполнены в ходе проекта CONVERGENCE, в результате чего текущая версия стала отправной точкой и основным технологическим фактором финансируемого ЕС городского инновационного проекта под названием AirHeritage, предназначенный для мониторинга и улучшения качества воздуха в малых и средних, но густонаселенных городах [15]. Самые последние результаты включают проверку полевой точности в полуконтролируемом экспериментальном развертывании в совместном расположении и функциональную проверку в кампании краудсенсинга, проведенной во время фазы 2 блокировки COVID-19 в Италии. Раздел 2 содержит подробную информацию об архитектуре MONICA, методологических и подготовительных аспектах кампаний по валидации. Затем в Раздел 3 включены результаты с особым акцентом на характеристику и долгосрочную оценку точности устройства, а также на объединение записанных в полевых условиях оппортунистических данных, полученных в ходе кампании по сбору информации.

## 2. Материалы и методы

### 2.1. Моника Архитектура

MONICA воплощает в своем названии свое основное предназначение. Целью всей системы является обеспечение возможности оценки качества воздуха с помощью нескольких взаимодействующих устройств, подвижных или стационарных, распределенных в определенной географической области. Такая система требует трех основных направлений развития.

Первый домен касается сенсорных узлов; предназначены для проведения точных измерений концентрации загрязняющих газов в воздухе в окрестностях объектов выбросов загрязняющих веществ (электростанции, дороги, отопительные приборы, промышленные предприятия). В следующем разделе подробно представлено оборудование, из которого построен узел; здесь стоит отметить, что важной ролью узла является передача измерений на второй уровень, который представляет собой сеть, отвечающую за сбор всех данных, поступающих от каждого отдельного узла.

Второй домен разработки отвечает за сеть, отвечающую за сбор данных, отправляемых узлами, и за серверную часть, которая преобразует грубые данные в информацию, понятную пользователю. На этом этапе можно принять несколько решений. В следующих разделах техника, использованная в этом проекте, представлена подробно.

Третья область развития касается представления результатов в форме, которая легко воспринимается человеком и позволяет гражданам (пользователям) делать осознанный выбор в отношении своего поведения. Смартфон был выбран из-за его широкого распространения и простоты использования в качестве средства представления пользователям результатов измерений.

#### 2.1.1. Узел МОНИКА

The diagram illustrates the Main Unit's architecture. It features a central BUS/ GPIO/ IRQline/ block. This block is connected to an MCU Board (via a bidirectional data bus), a Bluetooth wireless module (via a bidirectional data bus), and a Battery/ Charge regulator/ Step Up (via a control/power bus). The BUS/ GPIO/ IRQline/ block also interfaces with a Sensor Conditioning interface (AFE) and a Fan Driver (via bidirectional data buses). The Fan Driver is connected to a Coil Fan (via a control/power bus). An Air probe is connected to the AFE (via a control/power bus) and the Multi Sensor Array (via a data bus). The Multi Sensor Array is connected to the BUS/ GPIO/ IRQline/ block (via a bidirectional data bus) and receives input from an Air Inlet. The output of the Multi Sensor Array is labeled as Exhaust outlet. A DC IN is connected to the Battery/ Charge regulator/ Step Up (via a control/power bus).

Узел был разработан с нуля с целью поддержания энергопотребления на разумном уровне, преодоления ограничений по времени работы от батареи и достижения времени перезарядки, совместимого с аналогом для смартфона. Все компоненты платы имеют маломощные характеристики, включая электрохимические датчики в ее основе. Когда это возможно, все неиспользуемые компоненты переводятся в спящий режим.

Основная плата оснащена тремя датчиками Alphasense: CO (окись углерода) (CO-A4), NO2 (диоксид азота) (NO2-A43F) и O3 (озон) (OX- A431). измерения летучих органических соединений (ЛОС), когда установлен специальный датчик фотоионизационного устройства (ФИД). Эти датчики выдают сигнал напряжения, связанный с концентрацией определенного газа. Датчики установлены на аналоговом интерфейсе (АФЭ 810-0020-00), который выдает некоторые сигналы, связанные с концентрацией газов в воздухе.

Узел оснащен двумя вентиляторами, которые могут работать на пяти различных скоростях. Принудительная вентиляция гарантирует минимальный приток воздуха к датчикам для отслеживания динамики концентрации и способствует реактивности системы. Экспериментальные испытания с вентиляторами MC36304 проводились на разных скоростях. Все тесты проводились с использованием датчика

камера кондиционирования с фиксированной концентрацией 150 частей на миллиард NO<sub>2</sub> (которая часто встречается в полевых условиях), изменяя скорость вращения вентиляторов в диапазоне пяти значений. В качестве газа для испытаний был выбран NO<sub>2</sub>, так как он наиболее чувствителен к проблемам, связанным с потоком воздуха. Датчики, используемые для обнаружения, относятся к тому же типу, что и установленные на узлах. Результаты показаны на рисунке 2.

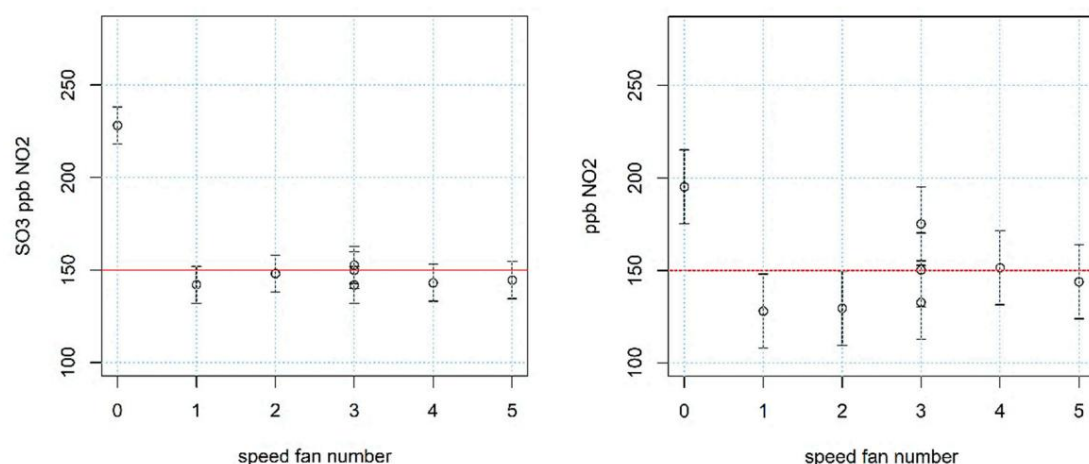


Рисунок 2. Оценка концентрации и доверительных интервалов для NO<sub>2</sub> с использованием датчика O<sub>3</sub>+NO<sub>2</sub> и датчика NO<sub>2</sub>, оба подвергались воздействию фиксированной концентрации NO<sub>2</sub> 150 частей на миллиард.

Эти испытания показали, что используемые датчики нуждаются в минимальном расходе воздуха для правильной работы в номинальных условиях эксплуатации, и позволили настроить скорость вращения вентиляторов на минимум для экономии энергии, гарантируя хорошую работу датчиков. Рабочие характеристики датчиков различаются при различных условиях воздушного потока. Некоторые газы (например, NO<sub>2</sub> и O<sub>3</sub>) трудно смешиваются с воздухом и не очень стабильны в случае датчиков, заключенных в частично открытые камеры; вращение вентилятора не только способствует смешиванию, но и компенсирует проблему нестабильности за счет нагнетания большего количества воздуха.

После обнаружения, кондиционирования, отбора проб и цифровой фильтрации микроконтроллер сохраняет в своих регистрах набор значений, которые кодируют информацию, необходимую для оценки концентрации CO, NO<sub>2</sub>, O<sub>3</sub> и, возможно, ЛОС в воздухе. Эти значения являются входными данными второго этапа пути данных MONICA, где алгоритм калибровки определяет фактическую концентрацию.

На борту реализован простой алгоритм, основанный на заводских данных, основанный на типичной функции линейной калибровки для этого типа датчика EC:

[illegible]

с использованием рассчитанной на заводе или в лаборатории чувствительности  $S$  (мВ/млрд) и смещения сигналов датчика  $V_{we\_zero}$  и  $V_{ae\_zero}$  (смещения электронного датчика и датчика нулевого чистого воздуха в мВ).

Необработанные и заводские калиброванные данные необходимо передать на второй этап системы. Стратегия, выбранная для этой цели, заключается в использовании возможностей подключения смартфона (SP) для передачи данных. Измеренные значения передаются от узла к SP через Bluetooth (BT). На плату был установлен трансивер HC06 BT; он связывается с микроконтроллером посредством последовательного интерфейса, а с другой стороны устанавливает BT-соединение с SP. Связь должна быть надежной, безопасной и работать таким образом, чтобы потери информации были минимальными. Таким образом, узел используется вместе с SP, на котором размещается приложение, выполняющее три основные функции: во-первых,

он связывает полезную информацию с приблизительными данными (например, GPS-координаты); во-вторых, предлагает удобный интерфейс; в-третьих, он служит шлюзом к серверной части.

При проектировании МОНИКИ требовалось особое внимание. Надежная конструкция, которая предотвращает отказы из-за ударов разумной силы или неправильных измерений из-за неблагоприятных погодных условий, необходима для сбора надежной информации о концентрации газа. Кроме того, часть корпуса состоит из металлического материала, который действует как экран, защищающий плату и датчики ЕС от электромагнитных помех.

### 2.1.2. Будущие работы

С помощью проведенных испытаний найдено минимальное значение скорости вращения, гарантирующее достижение достаточного расхода воздуха. Так, в следующей версии МОНИКИ (в настоящее время тестируется) вентиляторы имеют фиксированную скорость и приводятся в действие транзистором, который, в свою очередь, управляется выводом общего назначения микроконтроллера; это позволяет отключать вентиляторы, когда они не нужны, тем самым обеспечивая режим пониженного энергопотребления всей системы.

Аналоговые сигналы, поступающие от датчиков, усиливались для соответствия полосе обзора АЦП и фильтровались для уменьшения электрических помех. Частота дискретизации была увеличена, чтобы включить цифровую фильтрацию. Печатная плата (PCB) была переработана, чтобы свести к минимуму электрические помехи, которым неизбежно подвергается плата.

Основным улучшением нового узла по сравнению с текущей версией является возможность измерения концентрации твердых частиц. В частности, используется датчик Plantower PMS7003. Он поставляется в компактном корпусе, который хорошо помещается в корпус узла; связывается с микроконтроллером посредством интерфейса UART; эффективный диапазон датчика составляет от 0 до 500 мкг/м<sup>3</sup> и имеет разрешение 1 мкг/м<sup>3</sup>. Последнее улучшение касается приемопередатчика Bluetooth; новым является Bluetooth с низким энергопотреблением (BLE), который позволяет еще больше снизить энергопотребление и обеспечивает связь с современными смартфонами, использующими эту технологию в качестве интерфейса для других устройств.

Это обновление будет в центре внимания следующей версии MONICA (3.0).

## 2.2. Стратегии калибровки и проверки устройств

Здесь мы описываем детали некоторых из наиболее важных кампаний по валидации, проведенных за годы разработки второй версии архитектуры MONICA, в основном в рамках проекта CONVERGENCE. На самом деле кампания краудфандинга послужила основой для проверки концепции и первой функциональной проверки в условиях эксплуатации. Примерно через год после окончания этой кампании начался полевой эксперимент, длившийся два с лишним года, который только недавно закончился, с целью разработки соответствующих стратегий калибровки и оценки производительности самого устройства в рабочей среде. Наконец, кампания оперативных измерений, проведенная с помощью четырех устройств MONICA, была проведена в Портичи (город, расположенный в 7 км к югу от Неаполя), в то время как Италия переживала вторую фазу блокировки, вызванной пандемией COVID-19, в которой принимали непосредственное участие граждане. В следующем разделе представлены результаты.

### 2.2.1. Краудфандинговая кампания

Устройство MONICA 2.0 было целью краудфандинговой кампании, которая длилась 3 месяца и закончилась 17 декабря 2016 года [19]. Кампания была сочтена оптимальным финансовым инструментом для проекта, финансируемого из внутренних источников, и помогла преодолеть разрыв между нами, исследователями и гражданами, а также привлекла внимание регулирующих органов.

Собранные средства были использованы для разработки парка из 10 мультисенсорных блоков и их лабораторной калибровки. Граждане были разделены на категории в соответствии с их вкладом и получили взамен соответствующее вознаграждение (см. Таблицу 1).

Все спонсоры получили доступ к анонимным данным, зарегистрированным премиальными краудфандерами, и участвовали в рассылке новостей, чтобы быть в курсе развития проекта. Вместо этого «Охотники за смогом» приняли участие в функциональном тестировании, получив устройство MONICA и соответствующее приложение для смартфона сроком на один месяц в своих помещениях.

Таблица 1. Доступные ставки и связанные с ними премии во время краудфандинговой кампании.

Ставки (евро)	Квалификация/вознаграждение
5	Smog Enemy: периодически обновляется информационным бюллетенем о развитии проекта.
10	Smog Mapper: информирован о разработанных тестах
20	Smog Tracer: доступ к картам качества воздуха, подготовленным Smog. Охотники
45	Охотник за смогом: получил MONICA 2.0 дома на 1 месяц Исследователь
100	смога: посетитель лаборатории ВДНХ на 1 день Покровитель смога:
200	сторонник проекта Мастер смога: посол проекта
300	

Кампания рекламировалась в национальной прессе, национальном телевидении, радиопрограммах и социальных сетях. В частности, интерес к проекту проявили научные теле- и радиопрограммы, что помогло кампании добиться ожидаемого успеха. Кроме того, вовлечение граждан в кампанию поддерживалось на высоком уровне с помощью анкет обратной связи и предложений. Личные благодарности каждому основателю были опубликованы на веб-странице Eppela [19].

2.2.2. Лабораторная характеристика и настройка калибровки

В лаборатории определения характеристик газовых датчиков ENEA была разработана специальная установка для проведения кампании по определению характеристик и первой оценке эффективности.

Испытательная камера большого объема объемом 15 л (LVTC, рис. 3) была установлена в современной системе определения характеристик газовых датчиков (GSCS). Вкратце, GSCS состоит из герметичной испытательной камеры из нержавеющей стали, закрытой регулируемым термобоксом.



Рисунок 3. Фотография испытательной камеры большого объема 15 л (LVTC).

В LVTC состав воздуха (влажность и концентрация химических соединений) настраивается с помощью входного потока газа, который точно контролируется сертифицированными регуляторами массового расхода (серия MKS 1179). Точность химического состава газа обеспечивается смешиванием сертифицированных баллонов (Rivoira SpA). Для точности концентрации диоксида азота необходима дальнейшая проверка путем соединения выхода газа из камеры с датчиком Teledyne.

Хемилюминесцентный анализатор общего оксида азота T200. Температура и влажность регистрируются промышленными датчиками (LSI Pt100). LVTC может поддерживать калибровку нескольких комплектов сенсорных систем одновременно. Метод калибровки заключается во вдувании во входную трубку ЛВТК постоянного потока целевого газа, должным образом разбавленного до максимальной концентрации (C0) влажным синтетическим воздухом. Нарастающая во времени концентрация C(t) целевого газа точно предсказывается следующим экспоненциальным законом, который в общем случае описывает переход между двумя стационарными состояниями физического параметра при возмущении постоянной времени:

$$C(t) = C_0 + (C_1 - C_0) e^{-t/\tau} \quad (1)$$

Характерное время ( $\tau$ ) можно точно оценить с помощью калиброванного датчика; этот параметр пропорционален свободному пространству внутри камеры и должен быть соответствующим образом скорректирован, когда внутри камеры находится несколько датчиков.

Процедура калибровки (прогон) состоит из трех временных шагов: сначала вводится синтетический воздух для записи выходного сигнала датчика в невозмущенном состоянии (базовый уровень); на втором этапе в газ-носитель вводят должным образом разбавленную газовую мишень и регистрируют адсорбционную фазу отклика сенсора; наконец, на третьем этапе испытательную камеру промывают в постоянном потоке синтетического воздуха, регистрируя фазу десорбции выходного сигнала датчика. С помощью этой процедуры можно проверить поведение выходного сигнала сенсора во время фазы адсорбции и десорбции химического соединения на поверхности сенсоров. Можно обнаружить и измерить гистерезис чувствительности или отравление.

При расходе газа 1 л/мин  $\tau$  LVTC оценивается в 1100 с  $\pm$  50 с. Это означает, что за 3 часа несколько датчиков с временным откликом быстрее 2 минут могут быть откалиброваны от 0 до C0 с максимальной допустимой точностью.

### 2.2.3. Настройка полуконтролируемых полевых условий

4 апреля 2018 года устройство MONICA, полностью оснащенное датчиками NO2, O3 и CO, было развернуто совместно с регулирующей AQMS, расположенной в Неаполе (Via Argine, под кодовым названием NA09 в региональном реестре). AQMS эксплуатировалась и обслуживалась в соответствии с контролем регионального агентства по охране окружающей среды (ARPA-Campania)

[20]. Устройство было заключено в коробку, на несколько см больше самого устройства, а воздух, поступающий из обогреваемого воздушного коллектора самой станции, пропусклся через входное отверстие в результате действия расположенного ниже по потоку роторного насоса (см. рис. 4). В результате устройство MONICA анализировало ту же воздушную матрицу, которая подавалась на прибор нормативного уровня, которым оснащается AQMS. Устройство MONICA было подключено с помощью подключения BT к Raspberry Pi ver. 3, на котором работала ОС RASPBIAN и скрипт Python, предназначенный для получения данных с устройства и обеспечения локального хранения и удаленной передачи через специальный канал Wi-Fi на серверные системы MONICA. И Raspberry Pi, и помпа питались через интеллектуальную розетку, которую можно было включать и выключать удаленно.

MONICA была развернута и эксплуатировалась в блоке здания AQMS, поэтому рабочая температура поддерживалась в широком диапазоне 15 ° C системой HVAC блока, которая оставалась полностью функциональной в течение большей части времени развертывания. При анализе этих результатов следует учитывать, что температурные помехи, таким образом, были ограничены по отношению к полным колебаниям наружного воздуха и связанным с ними потерям производительности. Эта установка, однако, может выявить потенциальные эффекты старения датчика или отравления, упрощая сравнение откликов датчика в аналогичных условиях окружающей среды и концентрации целевого газа. Изменение температуры приточного воздуха поддерживалось системой ОВКВ на минимальном уровне, несмотря на внешние условия с относительной влажностью, зависящей от абсолютной влажности снаружи.





Рис. 4. Устройство МОНИКА, расположенное внутри установки АКМС, со схемой подачи воздуха от коллективного коллектора.

#### 2.2.4. Набор данных и процедуры долгосрочной полуконтролируемой полевой калибровки

Опираясь на эту настройку, была проведена кампания по совместному размещению, общая продолжительность которой превысила 2 года. Для оценки производительности, включая проверку более новой, но дорогостоящей схемы адаптивной калибровки, записанный набор данных позволил провести различные исследования, в которых было реализовано несколько стратегий калибровки [21].

Зарегистрированный набор данных состоит из 13 600 почасовых проб, зарегистрированных с апреля 2018 г. по ноябрь 2020 г. В результате пандемии COVID-19 в течение 2020 г., особенно с марта по июнь и в течение ноября 2020 г., можно было наблюдать значительное снижение концентраций загрязняющих веществ по сравнению с соответствующим месяцам прошлых лет.

В течение периода совместного размещения, несмотря на наличие кондиционера (AC), внутри эталонной станции происходили значительные колебания температуры, которые иногда достигали максимума (рис. 5) до более чем 40 °C. Частично это было связано с неправильной настройкой системы ОВКВ и ее неисправностями. В частности, набор данных содержит почасовые усредненные данные с устройства, т. е. необработанные показания датчиков рабочего электрода (WE) и вспомогательного электрода (AE) (мВ) для NO<sub>2</sub>, CO, O<sub>3</sub>, а также температуры (°C) и влажности (%).), соединенные с почасовыми усредненными данными эталонного анализатора ARPAC для CO (ppm) и NO<sub>2</sub> (ppb) (см. рис. 6 и 7). Набор данных был предварительно обработан, удалены все записи, содержащие пропущенные значения, а также обнаружены и удалены возможные выбросы с помощью обычных процедур на основе порога 6 сигм.

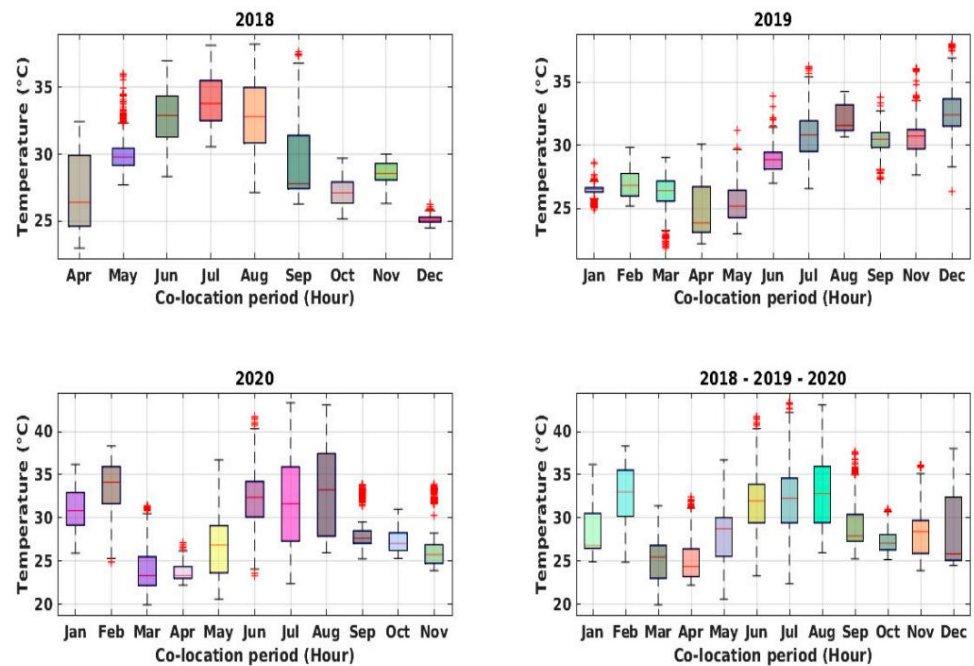


Рисунок 5. Графики почасовой регистрации температуры в течение всего периода совместного размещения и за каждый год.

Затем была реализована процедура калибровки с использованием двух разных методологий многомерного машинного обучения: множественной линейной регрессии (MLR) и трехслойной мелкой нейронной сети (SNN) с использованием трех сигмоидальных касательных нейронов в скрытом слое и линейного выходного нейрона. Алгоритмы используют в качестве входных необработанные данные датчика вместе с переменными окружающей среды для устранения помех и возвращают оценки концентрации в качестве выходных.

знак равно (2)

с участием

(3)

как соответствующий вектор концентрации загрязняющих веществ и

(4)

в качестве массива входных переменных используются необработанные данные с датчиков (напряжение на рабочем и вспомогательном электродах) для целевого газа и мешающих веществ, а также мешающих воздействий окружающей среды. В нашем случае несколько алгоритмов были оптимизированы для оценки концентрации одного загрязняющего вещества, поэтому  $C$  представлял собой концентрацию либо  $NO_2$ , либо  $CO$ .

На самом деле, для  $NO_2$  использовались показания температуры вместе с данными датчиков, нацеленных на  $NO_2$ :

$= WE \cdot T$

(5)

То же самое относится к концентрации газа CO, которая оценивалась с помощью датчиков CO. WE и AE, плюс показания температуры:

$$T_{CO} = T_{CO} + T_{CO} + T_{CO}$$

(6)

Были реализованы различные варианты длины обучающей и тестовой выборки, чтобы определить наилучшие размеры обучающей выборки для оптимальной процедуры калибровки. Результаты представлены в разделе 3.2 и сосредоточены на NO2 и CO, в то время как справочные данные по O3 отсутствуют.

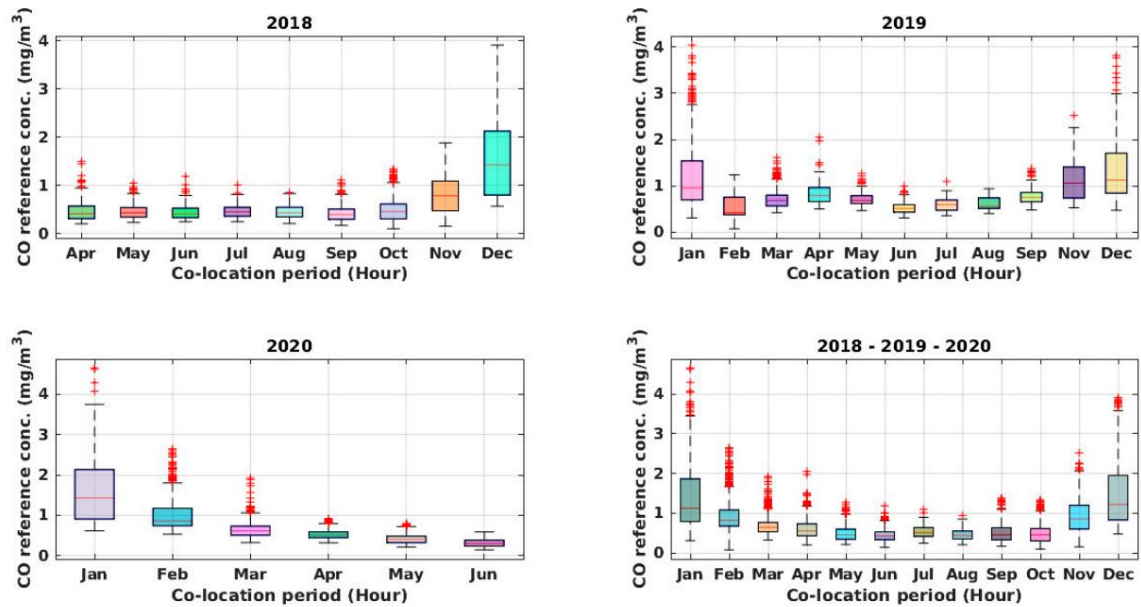


Рисунок 6. Графики почасовой регистрации эталонной концентрации CO в течение всего периода совместного размещения и за каждый год.

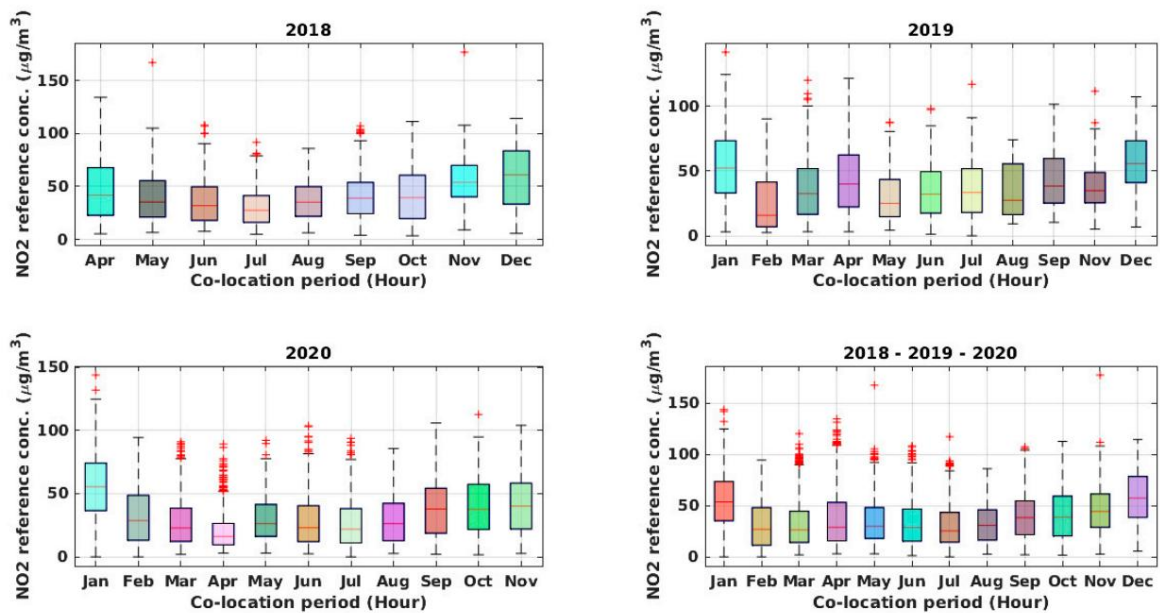


Рис. 7. Графики почасовой регистрации эталонной концентрации NO2 в течение всего периода совместного размещения и за каждый год.

### 2.2.5. Заключительная кампания по валидации краудсенсинга и процедуры калибровки

Чтобы проверить возможности устройства по краудсенсингу, во время фазы 2 первой блокировки из-за COVID-19 (конец мая 2020 г.) четыре устройства MONICA, предварительно откалиброванные в полевых условиях, были назначены четырем ассоциациям граждан для проведения 15-дневной кампании по мониторингу.

С 1 января 2020 г. по 1 марта 2020 г. эти четыре устройства MONICA были размещены на 2 месяца вместе с мобильной лабораторией ARPAC для регистрации данных как узлов, так и эталонных измерений (см. рис. 8).



Рис. 8. Четыре откалиброванных устройства MONICA в период совместного размещения (а) с мобильной опорной станцией ARPAC (б) в районе города Портичи.

Записанные наборы данных состоят из 1440 часов, снятых в режиме непрерывной выборки. В частности, для каждого узла были построены два набора данных с образцами, усредненными по минутам и часам. Данные от каждого из датчиков MONICA, т. е. необработанные показания датчиков WE и AE (мВ) для целевых датчиков NO<sub>2</sub>, CO, O<sub>3</sub> плюс T (°C) и RH (%), были объединены с усредненными данными того же временного масштаба с мобильного устройства. Эталонный анализатор ARPAC для NO<sub>2</sub> (мкг/м<sup>3</sup>), CO (мг/м<sup>3</sup>) и O<sub>3</sub> (мкг/м<sup>3</sup>). На рис. 9 недельные средние концентрации NO<sub>2</sub> показаны в период совместного размещения.

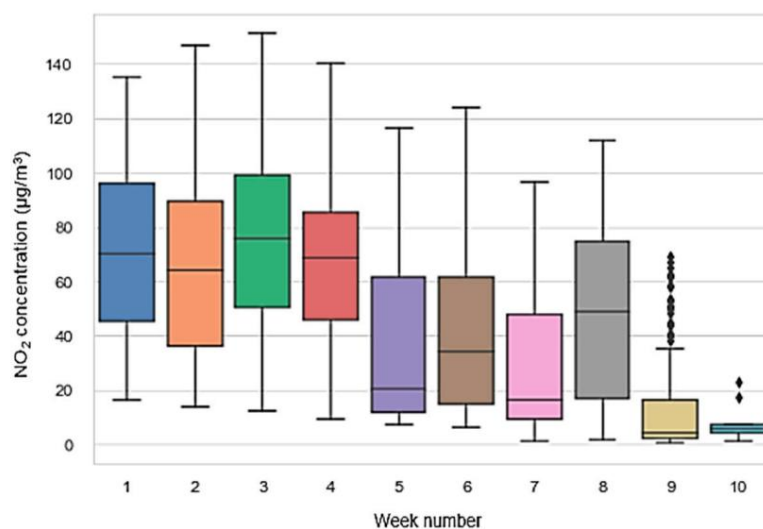


Рисунок 9. Средненедельные концентрации NO<sub>2</sub> во время совместного размещения.

Эти данные использовались для обучения линейных и нелинейных (мелкая нейронная сеть) моделей, характеристики которых сравнивались для выбора оптимальной стратегии калибровки. Предполагая, что  $X$  — вектор входных признаков, включая WE и AE для каждого из соответствующих датчиков, а  $y$  — прогнозируемое значение, используемая модель MLR может быть математически выражена как

$$y = X\beta + c, \quad (7)$$

где  $c$  — точка пересечения, а  $\beta$  — оптимальные коэффициенты наименьших квадратов.

Кроме того, в качестве нелинейного алгоритма была выбрана неглубокая нейронная сеть с трехслойной архитектурой, эмпирически оснащенная тремя стандартными сигмоидальными касательными нейронами в скрытом слое и линейным выходным слоем. В частности, в качестве алгоритма обучения использовалась автоматическая байесовская регуляризация (АБС). Поскольку целью кампании является слияние датчиков оппортунистических данных, в этой главе мы также сообщаем о результатах, полученных для задачи оценки среднечасовой концентрации NO<sub>2</sub> с использованием усредненных за час данных датчиков рабочего электрода (WE) и вспомогательного электрода (AE) для NO<sub>2</sub>, O<sub>3</sub> и датчики CO, а также данные о температуре и влажности в качестве входных данных для двух алгоритмов калибровки. Таким образом, для обоих алгоритмов входная матрица  $X$  включала восемь признаков (WE\_NO<sub>2</sub>, AE\_NO<sub>2</sub>, WE\_CO, AE\_CO, WE\_O<sub>3</sub>, AE\_O<sub>3</sub>, T и RH). Два алгоритма калибровки сравнивались с использованием разных длин обучения; остальные данные использовались в целях тестирования для имитации реальных условий, в которых узлы должны были работать после проведения калибровки.

Анализируя показатели эффективности (таблица 2) для результатов, становится ясно, что ограниченная польза может быть получена при использовании данных более чем за 3 недели и что MLR и NN дали очень похожие результаты. Напротив, четыре разных устройства получили значительно разные результаты, причем станция AQ8 отличалась худшими результатами. В конце концов мы выбрали алгоритм MLR в качестве функции окончательной калибровки для всех устройств, внедрив полученный коэффициент в Android-приложение MONICA, управляющее устройствами. Фактически, мы использовали весь набор данных в учебных целях, ожидая, что MAE для оценки NO<sub>2</sub> будет находиться в диапазоне от 6 до 12 мкг/м<sup>3</sup> в зависимости от узла MONICA. На рисунке 10 показаны выходные данные оценки концентрации газа NO<sub>2</sub> с использованием алгоритма MLR, рассчитанные для каждого узла, на протяжении всего периода совместного местоположения в зависимости от целевой линии концентрации газа.

Таблица 2. Средние абсолютные ошибки: (а) средняя абсолютная ошибка, (б) коэффициент корреляции Пирсона и (в) коэффициент детерминации (R<sup>2</sup>) для оценок NO<sub>2</sub>, полученных с использованием двух калибровочных моделей с различным выбором длины обучения (L, в недели) для каждого узла. Жирным шрифтом выделен уровень производительности, который был достигнут лучше всего.

L	Средняя абсолютная ошибка (MAE) [мкг/м <sup>3</sup> ]							
	AQ6		AQ8		AQ11		AQ12	
	NN	MLP	NN	MLP	NN	MLP	NN	MLP
1	11,7	7,94	21,94	23,36	8,20	7,78	12,23	6,55
2	7,53	7,70	25,64	16,78	10,07	9,51	8,82	6,92
3	8,89	7,73	19,48	13,30	10,09	8,86	8,33	6,49
4	8,74	7,56	11,71	12,63	10,24	9,88	7,08	6,31
5	7,98	7,63	13,15	11,37	9,6	9,65	5,79	5,15
L	Коэффициент корреляции Пирсона r							
	AQ6		AQ8		AQ11		AQ12	
	NN	MLP	NN	MLP	NN	MLP	NN	MLP
1	0,93	0,97	0,94	0,93	0,97	0,97	0,93	0,98
2	0,97	0,97	0,92	0,94	0,97	0,97	0,98	0,98
3	0,97	0,98	0,93	0,94	0,97	0,97	0,98	0,98
4	0,97	0,98	0,95	0,95	0,98	0,98	0,98	0,98
5	0,98	0,96	0,96	0,96	0,98	0,98	0,98	0,98

л	Коэффициент детерминации R2							
	AQ6		AQ8		AQ11		AQ12	
	NN	МЛР	NN	МЛР	NN	МЛР	NN	МЛР
1	0,79	0,91	0,47	0,41	0,91	0,92	0,78	0,94
2	0,91	0,9	0,22	0,62	0,85	0,88	0,88	0,92
3	0,88	0,89	0,49	0,74	0,86	0,88	0,89	0,92
4	0,87	0,88	0,77	0,75	0,84	0,84	0,91	0,93
5	0,88	0,88	0,75	0,81	0,87	0,87	0,94	0,95

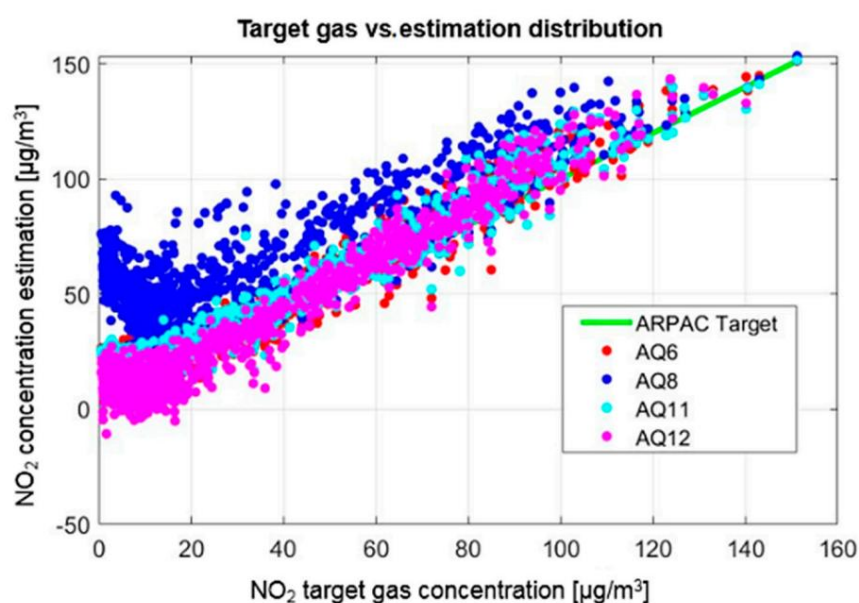


Рисунок 10. Выходные данные оценки концентрации газа NO<sub>2</sub> с использованием алгоритма MLR, рассчитанные для каждого узла на протяжении всего периода совместного размещения в зависимости от целевой линии концентрации газа. Разница в производительности четырех датчиков становится более очевидной, если принять во внимание низкую истинную концентрацию целевого загрязняющего вещества; AQ8 демонстрирует сильную предвзятость.

Четыре участвующие ассоциации отобрали шесть добровольцев, которые использовали четыре калиброванные устройства MONICA для мониторинга качества воздуха в соответствии с конкретной предложенной схемой мониторинга (рис. 11). Волонтеры обучались с помощью дистанционных сеансов в прямом эфире и коротких обучающих видеороликов. Это подразумевало минимум 1 час совокупной продолжительности сеанса мониторинга каждый рабочий день, проходя пешком и следуя по одному из четырех различных путей с использованием одного из четырех калиброванных устройств, которые еженедельно назначались одно. Каждое устройство использовалось на одном пути. Помимо технических трудностей, только четыре из 60 (15 × 4) предполагаемых интервалов наблюдения были пусты.



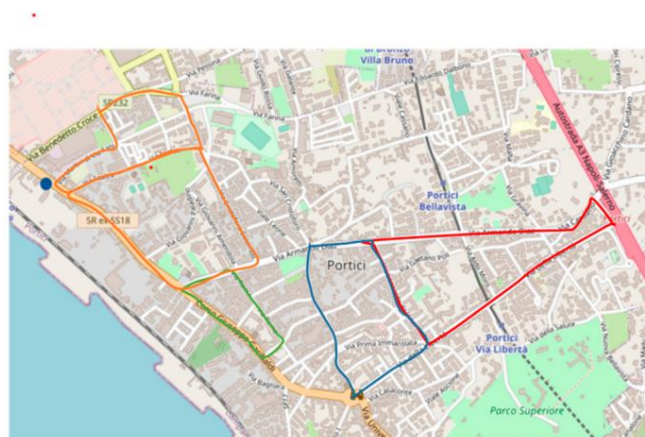


Рис. 11. Четыре предустановленных пути мониторинга (красный, синий, зеленый, оранжевый), показывающие местоположение мобильной лаборатории (синяя точка).

Данные были собраны и отправлены на серверную часть, где был подготовлен определенный набор данных, включающий все доступные измерения. Слияние датчиков проводилось с использованием геостатистической интерполяции, в частности, с использованием алгоритма обратного взвешивания по расстоянию (IDW) [22,23]. IDW фактически использовался для расчета средней интерполяции концентраций загрязняющих веществ в заданной сетке путем применения оппортунистических измерений, проведенных в определенный временной интервал (в данном случае, продолжительность кампании), используя:

$$c(x,y) = \frac{\sum_{i=1}^n \frac{c(x_i,y_i)}{d(x,y,x_i,y_i)^k}}{\sum_{i=1}^n \frac{1}{d(x,y,x_i,y_i)^k}}$$

где  $c(x,y)$  – концентрация в точке интерполяции  $p = (x,y)$ ,  $c(x_i,y_i)$  – концентрация в точках интерполяции (точках измерения)  $p_i = (x_i,y_i)$ , и

$$d = \frac{1}{\sqrt{k}} \sqrt{\sum_{i=1}^k (x_i - x)^2 + (y_i - y)^2}$$

где  $d$  — евклидово расстояние при  $k=2$ . В позиции интерполяции IDW использует фактические записи концентрации. На предварительном этапе проводится усреднение всех измерений, проведенных в пределах конкретной ячейки сетки. Шаг сетки был установлен в 5 м.

### 3. Результаты

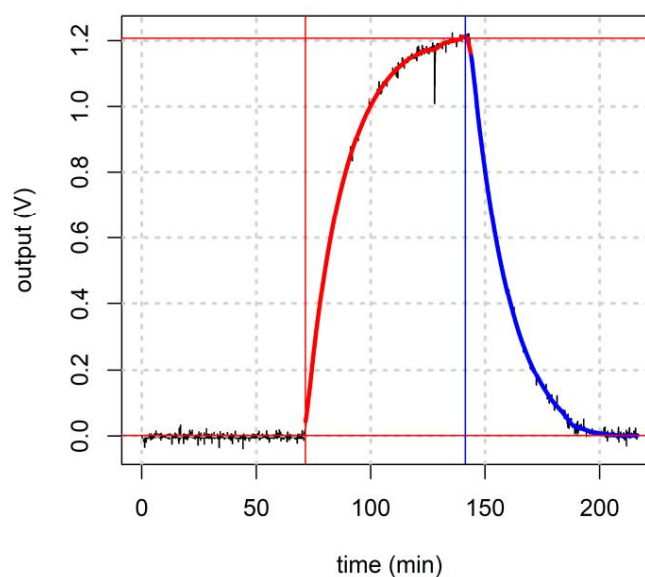
#### 3.1. Результаты лабораторной калибровки

Наш LVCT позволил нам разместить до восьми таких сенсорных систем (рис. 8a). Raspberry Pi 3 с Raspbian и скриптом Python собирал данные через синтаксические анализаторы MONICA в файлах журнала.

Калибровку сенсора проводили относительно 0:500 частей на миллиард NO<sub>2</sub> и 0:5 частей на миллион CO при контролируемых и постоянных температуре и влажности. На рис. 12b показан график зависимости выходного сигнала датчика во время калибровочного цикла с впрыском монооксида углерода в концентрации 5 частей на миллион. Легко различить три этапа калибровочного прогона; красная линия подчеркивает фазу адсорбции, а синяя линия показывает фазу десорбции. В результате прогона калибровки была оценена кривая чувствительности по журналу выходных данных датчика с использованием скрипта на языке R, который синхронизирует и коррелирует выходные данные датчика с концентрацией газа. После оценки чувствительность, показанная с помощью линейной регрессии данных (рис. 13a), может быть использована для изучения точности выходного сигнала датчика во всем диапазоне калибровки, как показано на рисунке 13б, где относительная погрешность датчика сообщается расчетная концентрация газа по сравнению с концентрацией газа. Таким образом, можно оценить полезные параметры датчика, такие как LOD (предел обнаружения), LOQ (предел количественного определения), линейность выходного сигнала, прецизионность и правильность.

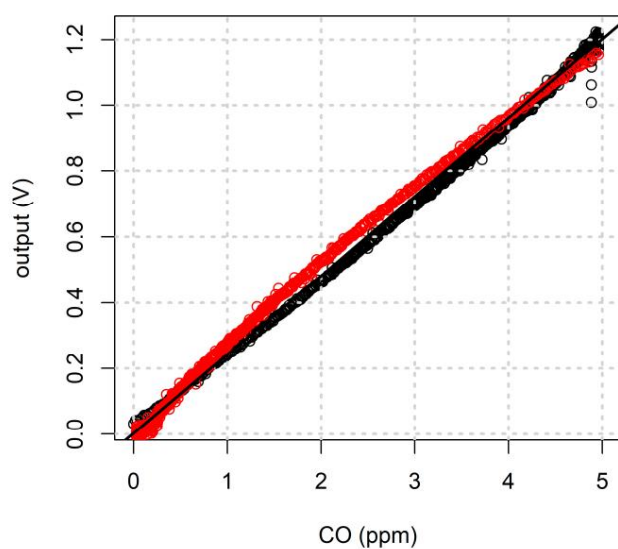


(a)

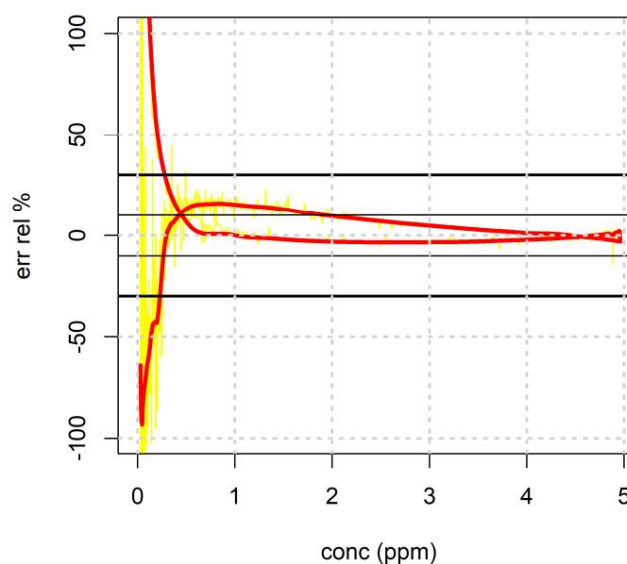


(б)

Рис. 12. (а) Восемь беспроводных сенсорных систем MONICA для мониторинга загрязнения воздуха во время калибровочного прогона. (б) График времени для одного выходного сигнала датчика во время впрыска 5 частей на миллион монооксида углерода. Красная линия показывает фазу адсорбции, а синяя линия показывает фазу десорбции выходного сигнала датчика.



(a)



(б)

Рис. 13. (а) Кривая чувствительности и линейная регрессия для выходного сигнала сенсора в диапазоне 0:5 ppm при адсорбции (черные точки) и десорбции (красные точки). (б) График относительной погрешности оценки концентрации газа по выходному сигналу датчика с рассчитанной чувствительностью в зависимости от концентрации газа.



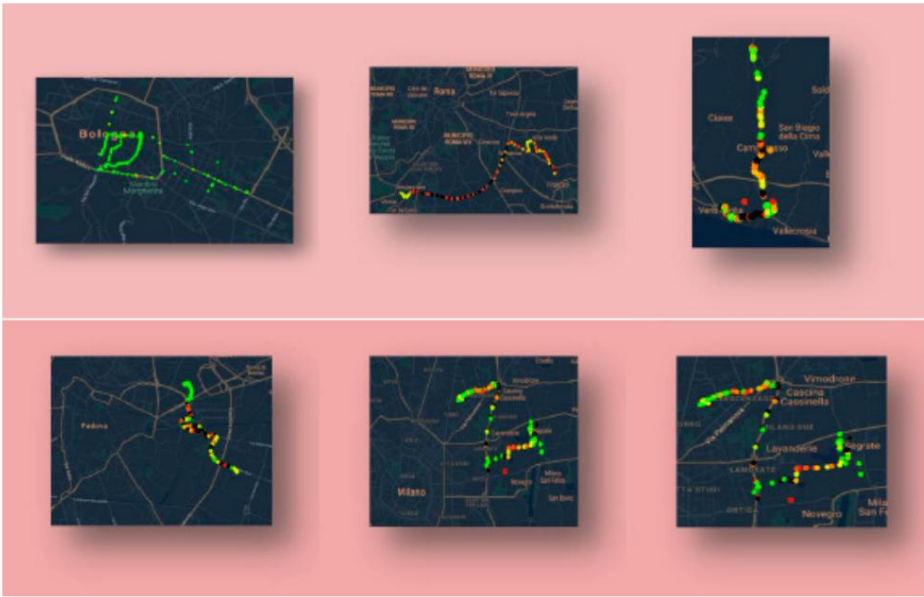
3.2. Результаты краудфандинга

В конце кампании проект MONICA был профинансирован на 8730 евро, что составляет 145% от ожидаемого потолка в 6000 евро. Кампания была полностью реализована и превысила целевые показатели финансирования, установленные для 102 сотрудников. 44 охотника за смогом использовали устройство MONICA в течение 30 дней вместе с его приложением для Android.

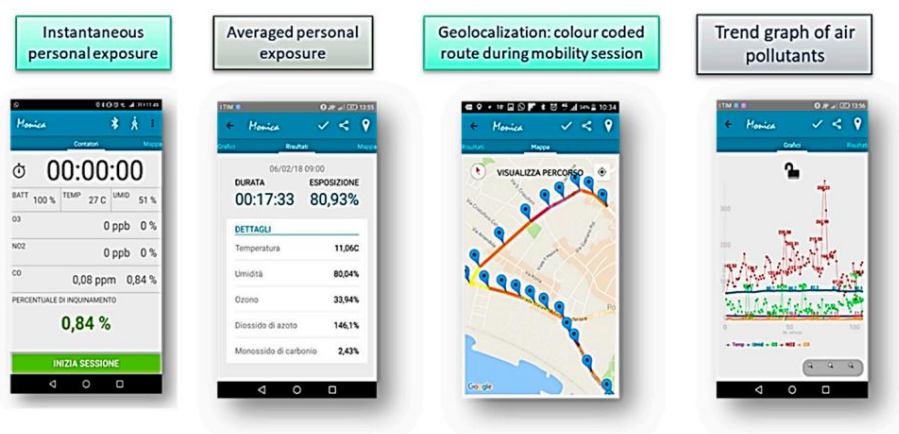
Устройства были отправлены итальянским краудфандерам, начиная с января 2018 года, и возвращены в лабораторию ENEA в конце 30-дневного периода для отправки оставшимся пользователям в циклическом режиме (см. Таблицу 3). Вместе с системой MONICA было поставлено подробное руководство по эксплуатации. Эта операционная система также позволила протестировать откалиброванную платформу в полевых условиях в качестве серверной части. Кроме того, их данные помогли создать крупную базу данных о загрязнении воздуха, на основе которой можно было строить карты, включая их обычные маршруты в своих городах (см. рис. 14).

Таблица 3. Особенности сеансов мобильности спонсоров.

Моника ID	Время Период	Место нахождения	Выборка Время	Нет. Сессии	Нет. Образцы
1	16 января – 22 февраля 2018 г.	Рима	30 с	31	2656
2	29 января – 26 февраля 2018 г.	Болонья	20 с	19	1129
3	19 февраля–16 3 марта 2018 г.	Падуа	20 с	29	1246
4	21 февраля – 16 марта 2018 г.	Сегрете (Мичиган)	30 с	20	774
5	19 февраля – 15 марта 2018 г., Новате Миланезе (Мичиган)		30 с	27	3836
6	14 3 марта–27 3 марта 2018 г.	Сан-Ремо (ИМ)	10 с	18	6150



(a)



(6)

Рис. 14. (a) Примеры карт маршрутов, составленных Smog Hunters (краудфандерами). Цвета кодируют нормализованную оценку концентрации загрязняющих веществ CO с заводской калибровкой, полученную приложением MONICA и (b) сохраненную в собственном бэкенде IoT. Обратите внимание, что на снимках экрана приложения для смартфона текст представлен на итальянском языке.

### 3.3. Два с лишним года результатов полуконтролируемого полевого развертывания

Для оценки процедур среднесрочной калибровки набор данных был первоначально разделен на два периода обучения, подчеркивающих как влияние сезонности на эмпирическое вероятностное распределение загрязняющих веществ, так и параметры окружающей среды.

Для оценки концентрации CO и NO<sub>2</sub> эксперименты проводились путем выбора разной длины тренировочной выборки и тестирования производительности в оставшиеся недели.

Результаты подвергались перекрестной проверке с использованием непересекающихся наборов обучающих данных, при этом первая кратность начиналась с первой недели периода совместного размещения.

Однако для результатов CO набор данных был дополнительно усечен в июне 2020 г. из-за неисправности датчика.

В табл. 4 и 5 представлены результаты линейной и нелинейной процедур калибровки для оценки концентрации CO и NO<sub>2</sub> соответственно, рассчитанные путем усреднения показателей производительности на недельных тестовых наборах. В частности, части a и b отображают результаты, полученные на двух разных временных отрезках набора данных: с апреля 2018 г. по июнь 2019 г. и с июля 2019 г. по ноябрь 2020 г.

Как мы видим, модель калибровки MLR почти всегда обеспечивает наилучший результат калибровки с точки зрения MAE, MRE и NRMSE, а также показывает повышенную эффективность обучения на данных датчика в этой конкретной конфигурации. Однако обе модели дают хорошие результаты калибровки при приемлемой длине обучающей выборки. В частности, результаты, полученные в двух разных наборах данных, основанные на разделении двух наборов данных, не показывают никаких признаков ухудшения характеристик датчика. На самом деле, при правильной повторной калибровке производительность, полученная в двух половинах, не показывает заметного ухудшения. Однако в долгосрочной перспективе производительность резко снижается, независимо от количества обучающих данных, как показывают данные, представленные в таблицах 4c, d и 5c, d. В частности, для части c в обеих таблицах показана перекрестная оценка эффективности с использованием самого большого размера рассматриваемой обучающей выборки, т. е. 4 недели. При рассмотрении всех оставшихся данных за более чем 2 года наблюдается значительный скачок производительности по индикаторам NRMSE и R<sup>2</sup>. В части d показана оценка эффективности использования данных всего набора данных за первые 4 недели в учебных целях.

Ухудшение производительности становится неприемлемым. Учитывая результаты, показанные в таблицах 4 и 5 и на рисунках 15–19, мы вынуждены винить датчики и эффекты дрейфа концепции, которые могут быть частично устранены с помощью соответствующих стратегий повторной калибровки.

Таблица 4. Показатели эффективности калибровки для оценок CO, полученные с использованием двух моделей калибровки с различными вариантами продолжительности обучения. Значения, выделенные жирным шрифтом, показывают наилучшую производительность, зарегистрированную в одном столбце, для медианы и среднего (курсив) показателя, при этом предпочтение отдается минимальному количеству требуемых выборок в случае равенства.

(а) Результаты перекрестной проверки промежуточных результатов калибровки CO (апрель 2018 г. – июль 2019 г.).

Длина тренировочного набора		МАЭ (мг/м3)		ЗППП		СКО (мг/м3)		NRMSE		R2		p	
		MLP	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
1 неделя (резюме)	(Иметь в виду)	0,27	0,28	0,24	0,30	0,37	0,42	0,84	0,96	-0,06	-0,22	0,79	0,63
	(медиана)	0,19	0,21	0,18	0,29	0,26	0,36	0,59	0,82	0,65	0,33	0,88	0,68
2 недели (резюме)	(Иметь в виду)	0,19	0,25	0,17	0,28	0,26	0,38	0,59	0,86	0,59	0,11	0,87	0,68
	(медиана)	0,17	0,21	0,16	0,28	0,24	0,36	0,54	0,80	0,70	0,35	0,89	0,69
3 недели (резюме)	(Иметь в виду)	0,17	0,21	0,16	0,26	0,23	0,35	0,52	0,77	0,72	0,37	0,89	0,72
	(медиана)	0,16	0,20	0,15	0,26	0,23	0,34	0,51	0,76	0,74	0,41	0,90	0,76
4 недели (резюме)	(Иметь в виду)	0,16	0,21	0,16	0,26	0,23	0,33	0,50	0,74	0,74	0,43	0,90	0,74
	(медиана)	0,15	0,20	0,15	0,26	0,22	0,32	0,48	0,71	0,77	0,47	0,91	0,77

(b) Результаты среднесрочной калибровки CO с перекрестной проверкой (июль 2019 г. – июнь 2020 г.).

Длина тренировочного набора		МАЭ (мг/м3)		ЗППП		СКО (мг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
1 неделя (резюме)	(Иметь в виду)	0,18	0,29	0,13	0,30	0,22	0,43	0,37	0,71	0,85	0,43	0,96	0,85
	(медиана)	0,17	0,29	0,12	0,32	0,20	0,41	0,34	0,69	0,89	0,49	0,96	0,87
2 недели (резюме)	(Иметь в виду)	0,16	0,24	0,12	0,26	0,21	0,36	0,34	0,60	0,87	0,61	0,96	0,89
	(медиана)	0,16	0,22	0,11	0,25	0,19	0,35	0,32	0,58	0,90	0,64	0,96	0,90
3 недели (резюме)	(Иметь в виду)	0,16	0,22	0,12	0,23	0,21	0,32	0,34	0,53	0,87	0,67	0,96	0,92
	(медиана)	0,15	0,20	0,11	0,17	0,19	0,27	0,31	0,44	0,90	0,79	0,97	0,94
4 недели (резюме)	(Иметь в виду)	0,16	0,21	0,12	0,21	0,20	0,29	0,33	0,49	0,89	0,71	0,96	0,92
	(медиана)	0,16	0,18	0,12	0,14	0,20	0,23	0,34	0,37	0,89	0,86	0,96	0,95

(с) Перекрёстная проверка результатов долгосрочной калибровки CO (апрель 2018 г. – июнь 2020 г.).

Длина тренировочного набора		МАЭ (мг/м3)		ЗППП		СКО (мг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
4 недели (резюме)	(Иметь в виду)	0,20	0,25	0,17	0,27	0,27	0,37	0,52	0,72	0,72	0,45	0,89	0,76
	(медиана)	0,18	0,23	0,17	0,26	0,26	0,35	0,50	0,70	0,75	0,53	0,91	0,77

(d) Результаты первоначальной калибровки CO (апрель 2018 г. – июнь 2020 г.).

Длина тренировочного набора		МАЭ (мг/м3)		ЗППП		СКО (мг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLP	SNN	MLR	SNN
4 недели	(Среднее)	0,22	0,24	0,21	0,22	0,30	0,33	1,16	0,99	-0,86	-0,17	0,76	0,66
	(Медиана)	0,21	0,17	0,21	0,14	0,31	0,23	0,72	0,95	0,47	0,04	0,85	0,68

Таблица 5. Показатели эффективности калибровки для оценок NO2, полученные с использованием двух моделей калибровки с различными вариантами продолжительности обучения.

(а) Калибровка NO2 с перекрестной проверкой (CV) (апрель 2018 г. – июль 2019 г.).

Длина тренировочного набора		МАЭ (мкг/м3)		ЗППП		СКО (мкг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
1 неделя	(Иметь в виду)	16,91	15,64	14,59	12,92	22,35	20,33	0,92	0,84	-0,15	0,20	0,72	0,69
	(медиана)	13,85	13,73	12,22	11,97	18,40	18,29	0,76	0,76	0,42	0,43	0,78	0,74
2 недели	(Иметь в виду)	13,90	14,89	12,08	13,00	18,43	19,79	0,76	0,81	0,40	0,25	0,76	0,69
	(медиана)	13,80	13,61	11,23	11,88	17,87	17,93	0,74	0,74	0,46	0,44	0,79	0,74
3 недели	(Иметь в виду)	14,42	13,85	12,98	12,30	19,42	18,55	0,80	0,76	0,23	0,39	0,76	0,72
	(медиана)	12,81	12,85	10,92	11,28	16,84	16,98	0,69	0,70	0,52	0,51	0,79	0,75
4 недели	(Иметь в виду)	13,02	13,34	11,40	11,92	17,33	17,91	0,71	0,73	0,49	0,42	0,78	0,74
	(медиана)	13,33	11,87	10,69	10,50	17,03	15,80	0,70	0,65	0,51	0,58	0,80	0,78

(b) Калибровка NO2 с перекрестной проверкой (CV) (июль 2019 г. – ноябрь 2020 г.).

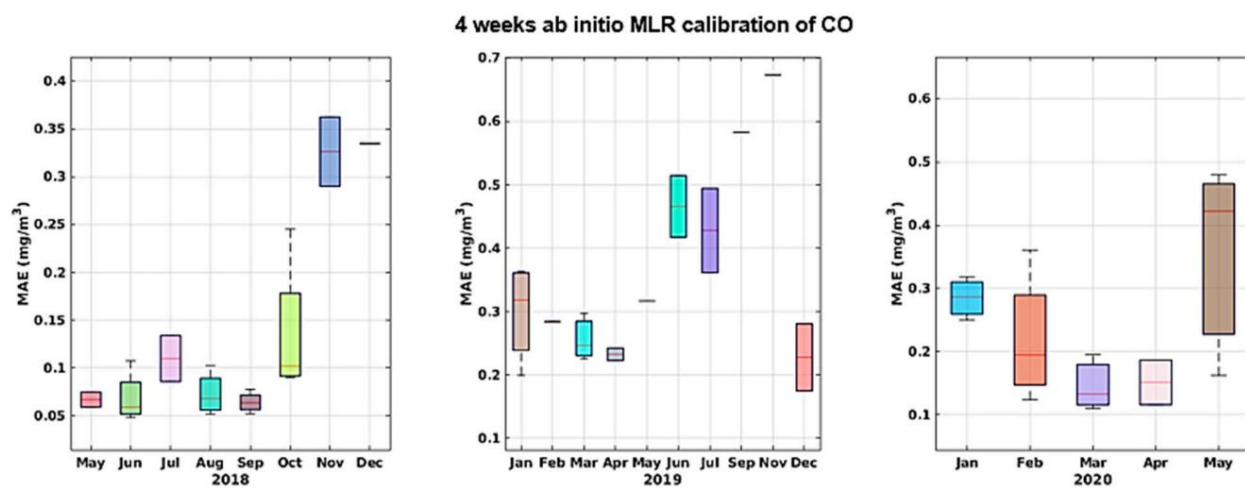
Длина тренировочного набора		МАЭ (мкг/м3)		ЗППП		СКО (мкг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
1 неделя	(Иметь в виду)	18.04	18.19	14.12	13,94	22,92	22,94	0,99	0,99	-0,04	-0,05	0,60	0,54
	(медиана)	16.30	16.80	12,78	12,96	20,96	21.20	0,90	0,92	0,18	0,15	0,65	0,58
2 недели	(Иметь в виду)	16.13	17,73	12,63	13,62	20.50	22.39	0,89	0,97	0,19	-0,01	0,63	0,56
	(медиана)	15.59	17.11	12.10	13.12	19,79	21,68	0,86	0,94	0,27	0,11	0,68	0,60
3 недели	(Иметь в виду)	15.20	17.06	12.05	13,78	19.41	21,95	0,84	0,95	0,27	0,04	0,66	0,56
	(медиана)	14.46	15,71	11.55	12,65	19.01	20.06	0,83	0,87	0,32	0,24	0,68	0,63
4 недели	(Иметь в виду)	13,76	14,73	10,99	11,83	17,62	18.90	0,76	0,82	0,41	0,31	0,71	0,65
	(медиана)	13,96	14.59	10,99	11.53	17,74	18.53	0,77	0,80	0,41	0,35	0,71	0,66

(c) Калибровка NO2 с перекрестной проверкой (CV) (апрель 2018 г. – ноябрь 2020 г.).

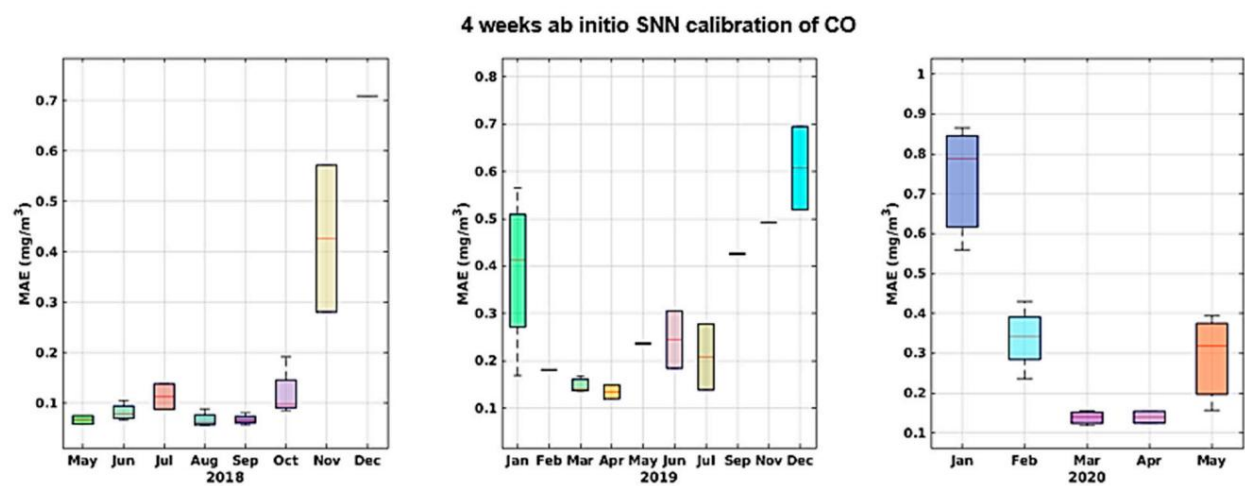
Обучающий набор	Набор тестов	МАЭ (мкг/м3)		ЗППП		СКО (мкг/м3)		NRMSE		R2		p	
Длина	Длина												
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
4 недели	CV за 4 недели (среднее)	15.09	16.55	12.40	13,96	19.54	21,67	0,82	0,90	0,32	0,12	0,70	0,61
	(медиана)	14,91	15.59	12.08	12,91	19.41	20,61	0,81	0,86	0,34	0,25	0,72	0,66

(г). Калибровка NO2 ab initio (апрель 2018 г. – ноябрь 2020 г.).

Длина тренировочного набора	Длина тестового набора	МАЭ (мкг/м3)		ЗППП		СКО (мкг/м3)		NRMSE		R2		p	
		MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN	MLR	SNN
4 недели	4 недели	(Среднее) 14,72	15,68	11.22	10,78	18.56	19.07	0,86	0,89	0,18	0,11	0,69	0,60
		(Медиана) 14,93	15,83	10,78	10,95	17.41	19.20	0,84	0,88	0,28	0,22	0,70	0,62



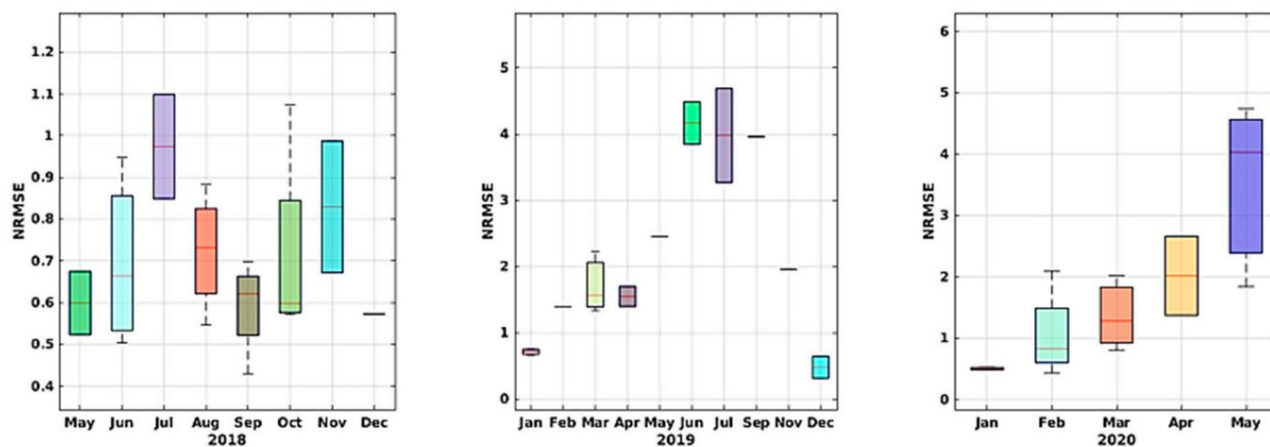
(a)



(b)

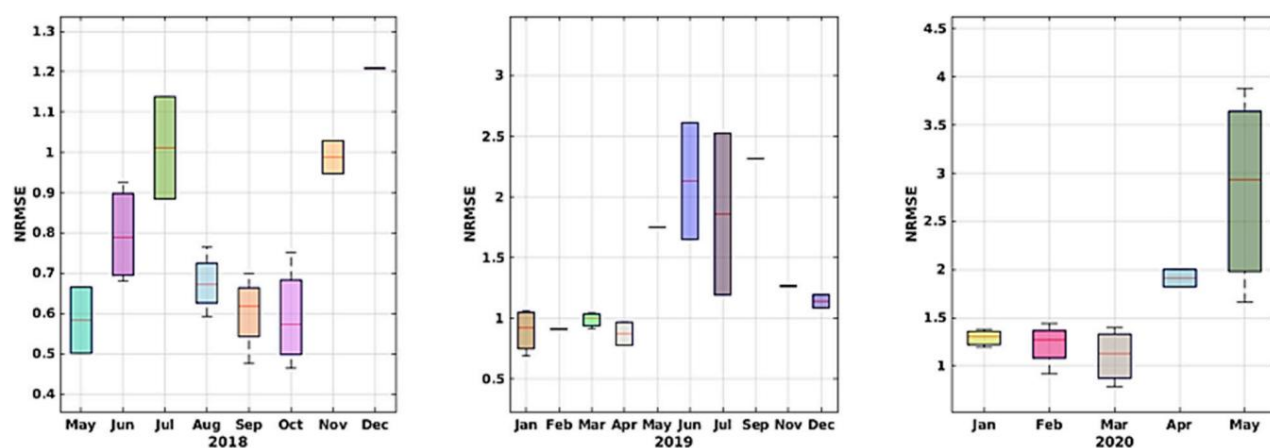
Рисунок 15. Тенденции MAE, показанные ежемесячной диаграммой для ab initio калибровки CO для MLR (a) и SNN (b); последний показывает несколько лучшие показатели в течение первого и последнего года.

## 4 weeks ab initio MLR calibration of CO



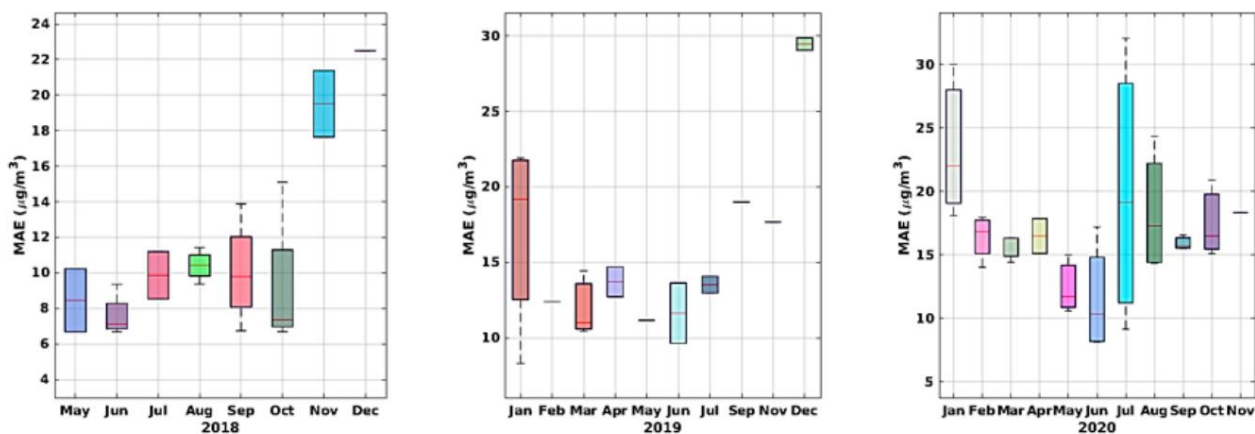
(a)

## 4 weeks ab initio SNN calibration of CO

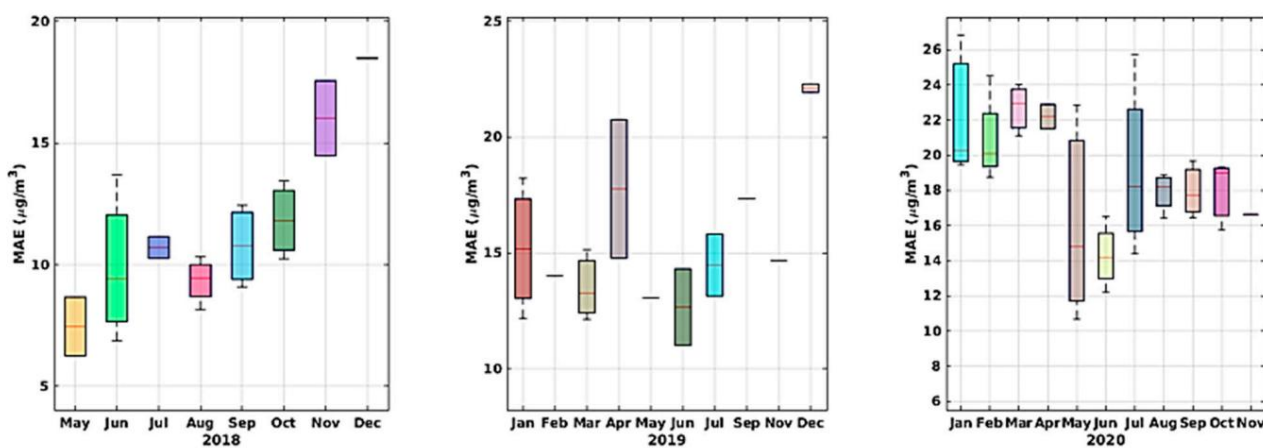


(b)

Рисунок 16. Тенденции NRMSE, показанные ежемесячной диаграммой для калибровки ab initio CO для MLR (a) и SNN (b), причем последний показывает несколько лучшие показатели в течение первого и последнего года.

4 weeks ab initio MLR calibration of NO<sub>2</sub>

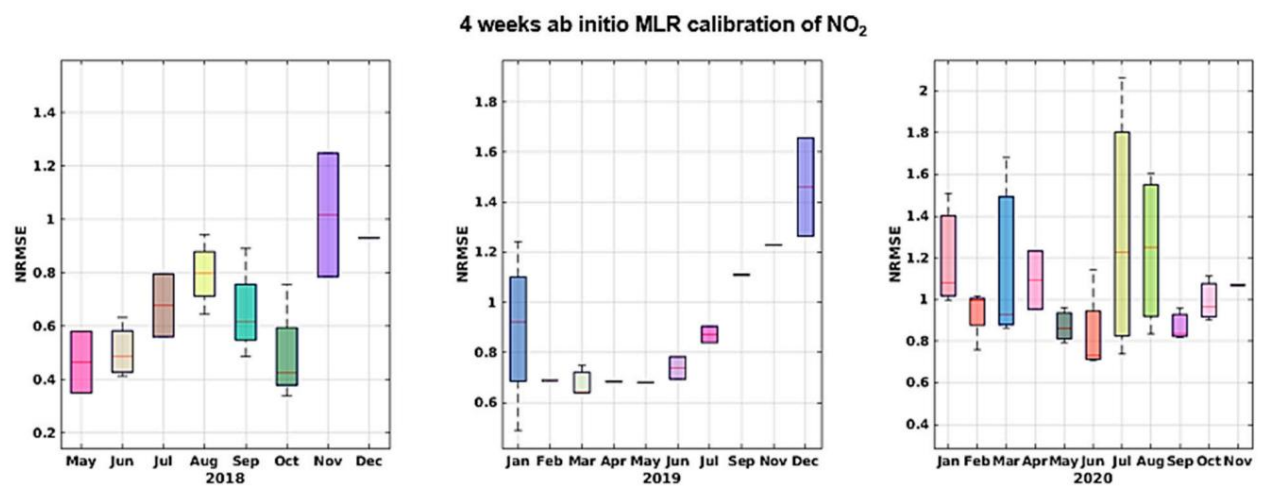
(a)

4 weeks ab initio SNN calibration of NO<sub>2</sub>

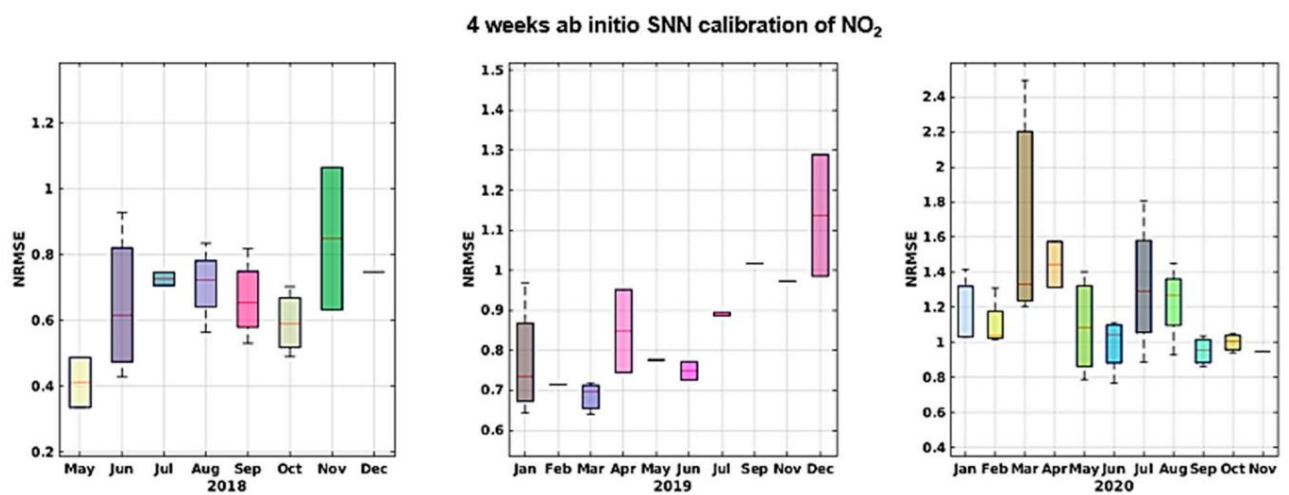
(b)

Рис. 17. Тенденции MAE, показанные на месячной ящичковой диаграмме для ab initio калибровки NO<sub>2</sub> для MLR (a) и SNN (b); последний показывает несколько лучшие показатели последовательно в течение 3 лет.





(a)



(b)

Рисунок 18. Тенденции NRMSE, показанные на месячной диаграмме для калибровки NO<sub>2</sub> ab initio для MLR (a) и SNN (b); последний показывает несколько лучшие показатели в течение первого и последнего года.

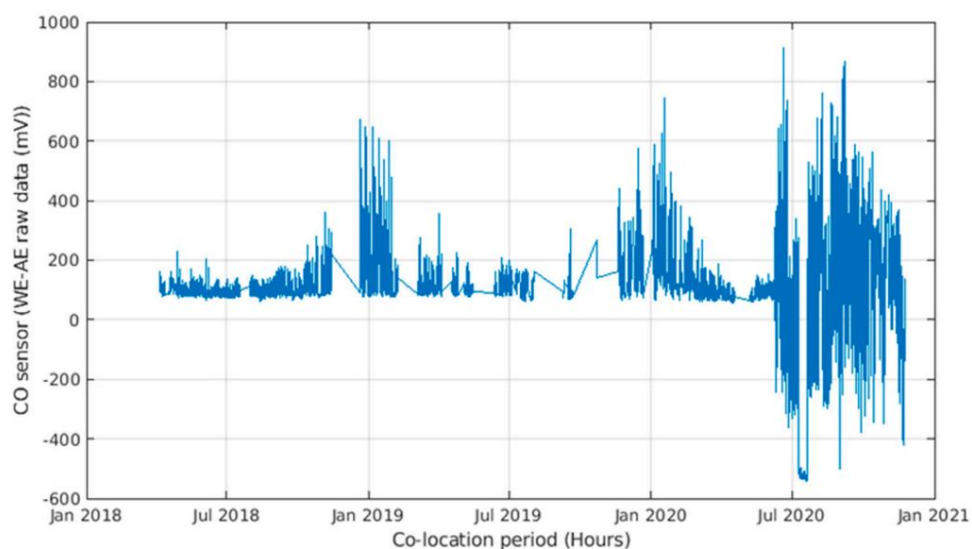


Рисунок 19. Разница между потенциалом электрода WE и потенциалом электрода AE (WE-AE) необработанных данных для датчика CO MONICA в течение всего периода совместного размещения; обратите внимание на внезапную потерю сенсорных свойств, которая произошла в июне 2020 года, более чем через 2 года после развертывания.

### 3.4. Результаты проверки методом

краудсенсинга В таблице 6 приведены усредненные концентрации, зарегистрированные всеми добровольцами на всей городской территории для всех четырех устройств. Для CO и NO<sub>2</sub> результаты совместимы с ожидаемым увеличением концентраций загрязняющих веществ по сравнению с измерениями этапа 1 (полная блокировка) из-за медленного возобновления производственной деятельности в районе из-за нормативной базы этапа 2. Озон сохраняет значения, аналогичные зарегистрированным в последние дни фазы 1.

Таблица 6. Характеристика записанных данных первого порядка.

	Статистика первого порядка	
	В среднем	Среднеквадратичное отклонение
CO (мг/м <sup>3</sup> )	0,44	0,64
NO <sub>2</sub> (мкг/м <sup>3</sup> )	40,0	37,1
O <sub>3</sub> (мкг/м <sup>3</sup> )	76,2	34,3

Внутренний оппортунистический характер деятельности по мониторингу гражданской науки и разница в длине четырех путей были отражены слегка неравномерной плотностью измерений, рассчитанной на основе записанных данных позиционирования (рис. 20). В частности, некоторые области кажутся чрезмерно представленными из-за нескольких записей, сделанных во время нескольких кругов по одному и тому же маршруту. Следует проявлять осторожность при оценке недопредставленных областей (самые темные цвета), которые будут страдать от зависимости от временной дисперсии, что может привести к нерепрезентативным результатам в пространственных шаблонах, усредненных IDW. Откалиброванные данные с измеренными концентрациями были объединены для построения карт взвешивания обратных расстояний. На рисунках 21–23 показаны результирующие модели загрязнения. На рис. 21 показаны средние концентрации CO, отслеживаемые во время кампании у всех добровольцев, независимо от времени суток. Для них характерны локальные горячие точки, расположенные вблизи основных перекрестков и в районах с интенсивным автомобильным движением. Однако возникла непредвиденная точка доступа, подтвердившая беспрецедентную разрешающую способность совместного мобильного мониторинга.

Анализ структуры NO<sub>2</sub> (рис. 22) в основном подтверждает наличие горячих точек, выявленных с помощью анализа структуры CO; однако некоторые наиболее загрязненные районы характеризуются значениями, приближающимися к нормативным порогам относительно измеренных средних значений концентрации CO. Наконец, на рис. 23 показаны графики концентраций O<sub>3</sub>.

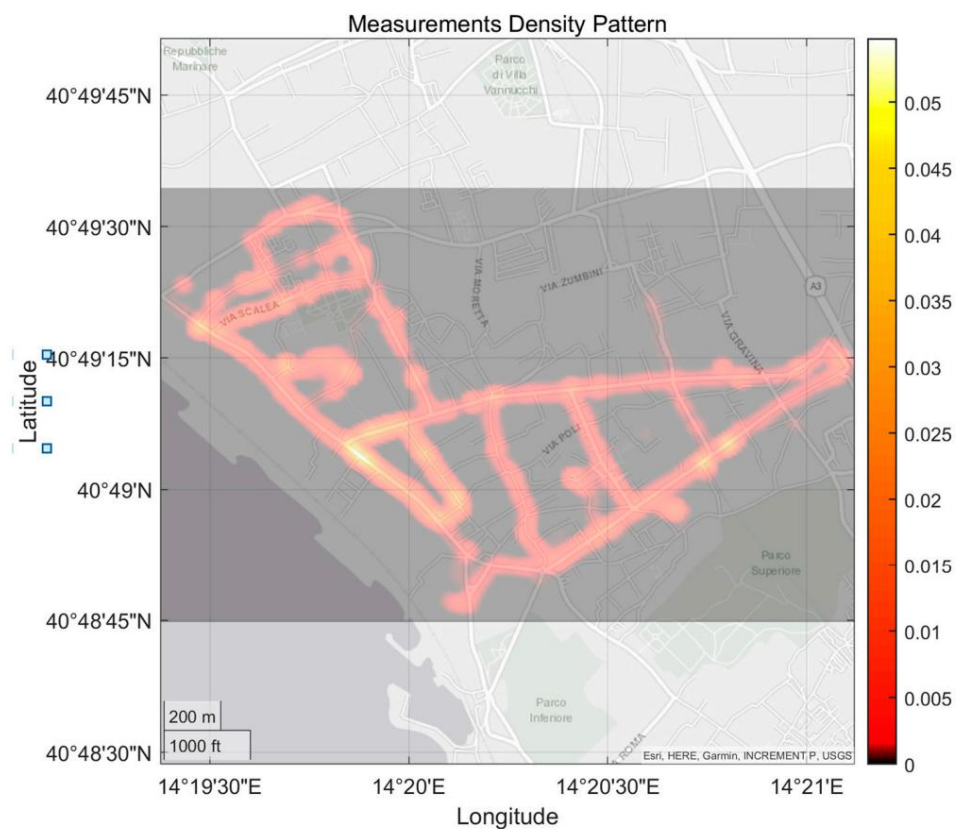


Рис. 20. Пространственная плотность конъюнктурных измерений, отобранных кампаниями с участием граждан. Неравномерная плотность паттернов возникает из-за разной длины предустановленных путей. Непредвиденные места измерения также показаны из-за путей, которые граждане решили выбрать для достижения предусмотренных мест измерения.

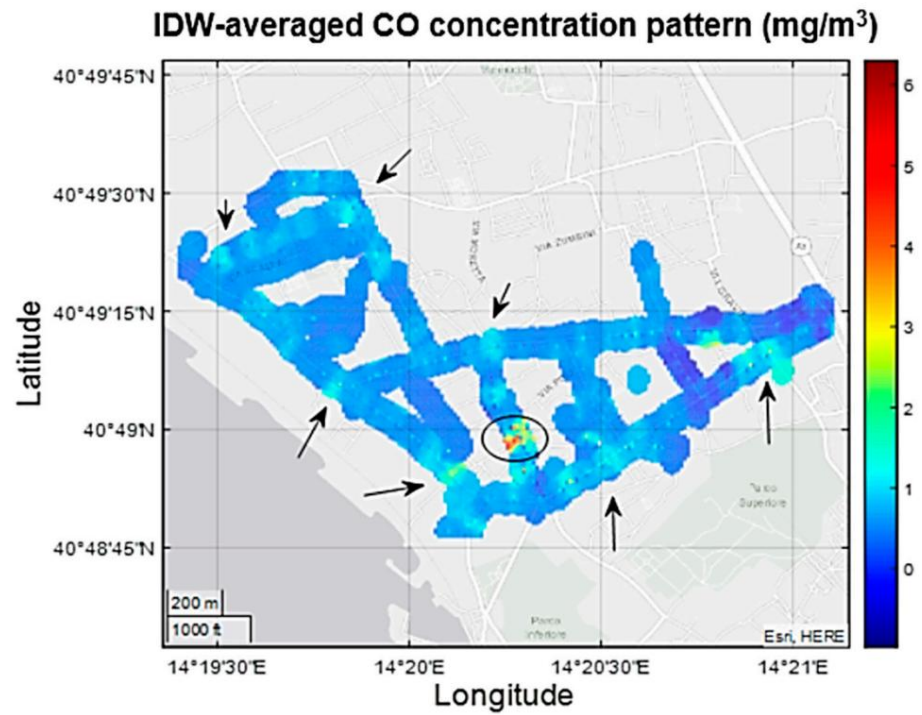


Рисунок 21. Усредненная по IDW картина концентрации CO характеризуется локализованными горячими точками вблизи основных перекрестков или улиц с интенсивным движением (стрелки). Также возникли непредвиденные горячие точки, что побудило к проведению специальных кампаний по измерению (эллипс).

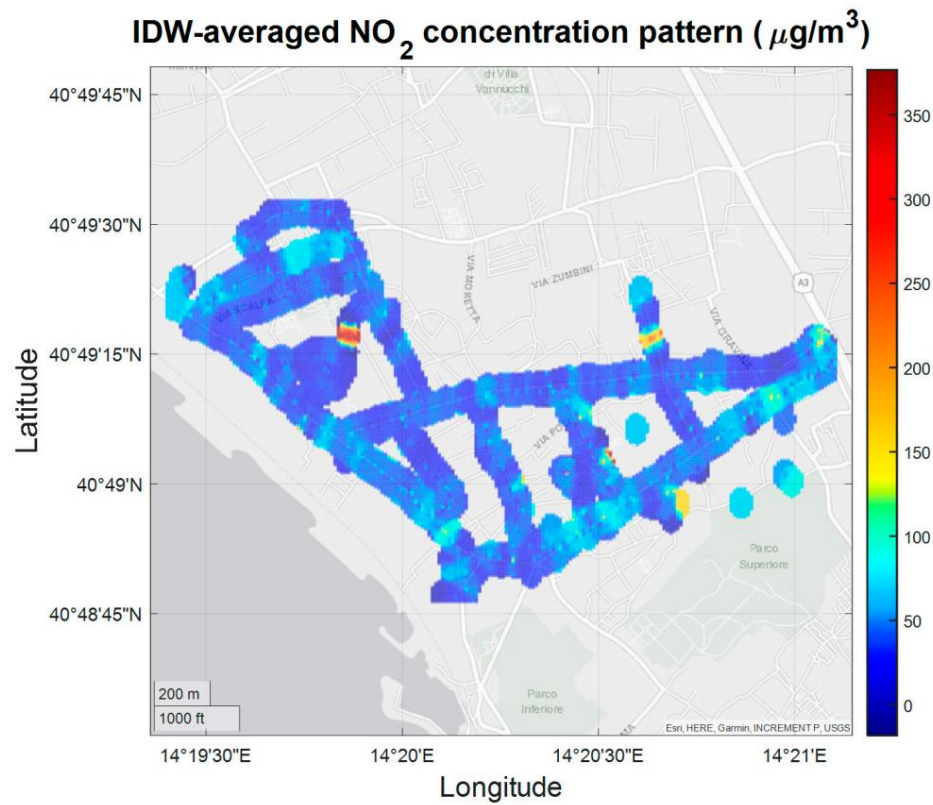


Рис. 22. График концентрации NO<sub>2</sub>, усредненный по IDW , подтвержденный областями горячих точек, выявленными с помощью мониторинга CO.

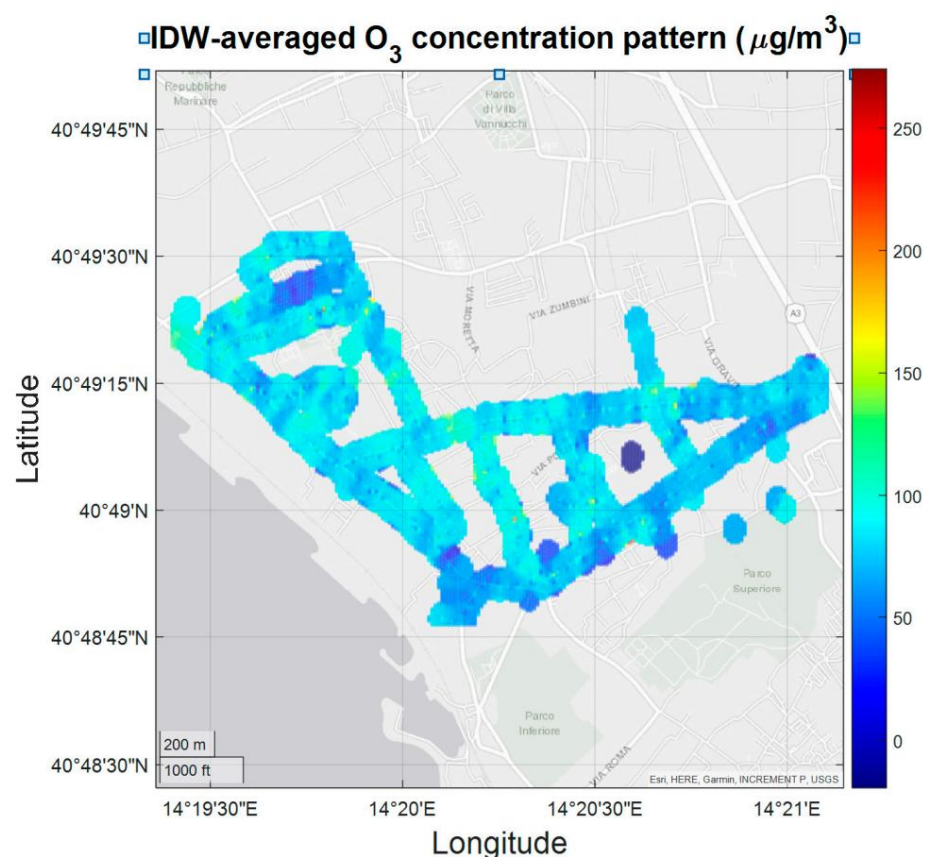


Рис. 23. Характеристика концентрации O<sub>3</sub>, усредненная по IDW, в целом демонстрирует более низкую пространственную дисперсию со средними значениями, которые достигают или превышают нормативный порог.

Усредненные значения IDW по озону демонстрируют меньшую пространственную дисперсию, но относительно ближе к нормативным пороговым значениям и превышают их локально. Хотя такое поведение характерно для летнего сезона в районе мониторинга, эти результаты требуют более тщательного анализа основных факторов.

#### 4. Обсуждение и выводы

В этой работе была разработана и описана архитектура IoT для пространственного и временного мониторинга качества воздуха с высоким разрешением. Разработанная архитектура основана на химической системе с несколькими датчиками, основанной на электрохимических датчиках и алгоритмах калибровки на основе полевых данных, полученных с помощью подходов машинного обучения. Приложение для смартфона было разработано для обеспечения обратной связи с пользователями в режиме реального времени или с задержкой, в то время как веб-сайт на основе нереляционной базы данных (NoSQL) обеспечивает интеграцию данных и объединение их с сообществом пользователей. Большинство этих достижений произошло во время реализации проекта FlagEra CONVER GENCE. Многогранные кампании по валидации, включая функциональную валидацию, финансируемую за счет краудфандинга, лабораторную характеристику, долгосрочное стационарное совместное размещение и краудсенсинг, продемонстрировали жизнеспособность проекта для персональных, мобильных или стационарных приложений.

В частности, совместное размещение в полевых условиях длилось два с лишним года и дало полезные сведения о долгосрочном рабочем поведении массивов электрохимических датчиков. В частности, мы не смогли обнаружить значительного ухудшения потенциальной точности датчиков в течение более чем 2 лет после развертывания, когда один из датчиков, в частности датчик CO, в конечном итоге сломался. Несмотря на изменение датчиков и концепции, ежегодные процедуры повторной калибровки позволили восстановить большую часть начальных годовых уровней производительности. Три недели

данных о совместном размещении с высокоточной системой мониторинга нормативного уровня оказалось достаточно, чтобы гарантировать хорошую работу в течение более 6 месяцев при использовании подходов полевой калибровки. Долгосрочная оценка производительности с помощью процедур перекрестного тестирования показала, что, используя в качестве обучающей выборки данные о совместном размещении как минимум за 4 недели независимо от даты начала калибровки, можно было получить в среднем достаточно хорошие характеристики. Производительность, полученная с помощью многолинейной калибровки и мелкой нейронной сети, была очень похожа на первую, что обеспечивало несколько лучшие свойства об

Увеличивая масштаб с помощью одной калибровки *ex ante* с данными за 4 недели, мы также показали, как начало сезонной и антропогенной изменчивости условий окружающей среды и, соответственно, концентраций загрязняющих веществ вызывало периодическое ухудшение точности, которое лишь частично восстанавливалось при восстановлении ситуации. до состояния, аналогичного периоду калибровки. Однако по прошествии первого года производительность стала неприемлемой, и ежегодная повторная калибровка является минимальным требованием, гарантирующим уровень производительности.

Наконец, краудсенсинговая кампания продемонстрировала жизнеспособность платформы в качестве персонального экспозиционного устройства, в то время как совместный сбор оппортунистического обмена данными обеспечил возможности картирования с высоким разрешением. Будущая работа будет включать масштабирование массива датчиков для включения возможностей мониторинга твердых частиц с последующей модификацией интерфейсной и серверной подсистем IoT, а также фактическое использование полученной архитектуры для долгосрочной полномасштабной краудсенсинговой кампании в рамках проекта UIA AirHeritage. проект.

Вклад авторов: Концептуализация, SDV и EM; Курирование данных, EE, SF и PD; Привлечение финансирования, SDV, TP, AMI и GDF; Методология, CDB, ЭЭ, ЭМ, ГФ, ТП.

Расследование, SDV, EE, EM, FF, SF, ADG и MS; Программное обеспечение, EE, FF и GD; Валидация, SDV, EE, EM, FF, GF, ADG, GD и PD; Надзор, CDB и ГДФ; Написание – исходный вариант, SDV, EE, EM, FF Все авторы прочитали и согласились с опубликованной версией рукописи.

Финансирование: Эта работа получила финансирование от программы Европейской комиссии H2020 FLAG-ERA в рамках проекта Convergence – Energy Efficient Convergent Wearables for Healthcare and Lifestyle Applications (номер ссылки: JTC-2016\_003) и от ERDF в рамках UIA . -Проект AirHeritage .

Заявление Институционального наблюдательного совета: Этическая проверка и одобрение этого исследования были отклонены, поскольку конфиденциальные данные не хранились на постоянной основе, что привело к сбору анонимных данных.

Заявление об информированном согласии: Информированное согласие было получено от всех участников исследования.

Заявление о доступности данных: подтверждающие данные в настоящее время доступны по запросу и, как ожидается, будут опубликованы в течение 2022 года.

Благодарности: Авторы выражают благодарность Анне Саварезе и Антонио Галлоцци (Legambiente Campania) за поддержку координации волонтеров во время финальной кампании по краудсенсингу и Пино Онорати из ARPA Campania за поддержку кампании по совместному размещению.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

#### использованная литература

1. Штейнле, С.; Рейс, С.; Sabel, CE Количественная оценка воздействия загрязненного воздуха на человека — переход от статического мониторинга к оценке индивидуального воздействия с пространственно-временным разрешением. науч. Общая окружающая среда. 2013, 443, 184–193.
2. Цзян, Х.-Q.; Мэй, Х.-D.; Фэн, Д. Загрязнение воздуха и хронические заболевания дыхательных путей: что люди должны знать и делать? Дж. Торак. Дис. 2016, 8, E31–E40, doi:10.3978/j.issn.2072-1439.2015.11.50.
3. Ландриган, П.Дж.; Фуллер, Р.; Акоста, Нью-Джерси; Адейи, О.; Арнольд Р.; Бальде, АВ; Чжун, М. Комиссия Lancet по загрязнению и здоровью. Ланцет 2018, 391, 462–512.
4. Европейское региональное бюро ВОЗ. Обзор данных о аспектах загрязнения воздуха для здоровья – проект Revihaar: технический отчет; Европа Региональное бюро ВОЗ: Женева, Швейцария, 2013 г.
5. Williams, DE Недорогие сенсорные сети: откуда мы знаем, что данные надежны? Acs Sens. 2019, 4, 2558–2565, doi: 10.1021/acssensors.9b01455.

6. Фатторузо Г.; Ночерино, М.; Тоскано, Д.; Париота, Л.; Соррентино, Г.; Манна, В.; Ди Франсия, Г. Анализ пригодности участка для низких Сети датчиков стоимости для мониторинга загрязнения воздуха в городах с пространственной плотностью. *Атмосфера* 2020, 11, 1215.
7. Де Вито, С.; Эспозито, Э.; Кастелл, Н.; Шнайдер, П.; Бартонова, А. Надежность полевой калибровки для интеллектуального качества воздуха мониторов. *Сенсорные приводы В Chem.* 2020, 310, 127869, doi:10.1016/j.snb.2020.127869.
8. Кастелл, Н.; Дауге, Франция; Шнайдер, П.; Фогт, М.; Лернер, У.; Фишбеин, Б.; Бродей, Д.; Бартонова, А. Кан коммерческий по низким ценам сенсорные платформы способствуют мониторингу качества воздуха и оценке воздействия? *Окружающая среда. Междунар.* 2017, 99, 293–302, doi:10.1016/j.environ.2016.12.007.
9. Эспозито, Э.; Де Вито, С.; Сальвато, М.; Фатторузо, Г.; Ди Франсия, Г. Вычислительный интеллект для интеллектуальных мониторов качества воздуха Калибровка. в компьютерном зрении; Springer Science and Business Media LLC: Берлин, Германия, 2017 г.; Том 10406, стр. 443–454.
10. Балрам, Д.; Лиан, К.-Ю.; Себастьян, Н. Новая система предупреждения об опасном приземном озоне на основе мягких датчиков с использованием усовершенствованной нейронной сети наименьших квадратов с демпфированием. *Экотоксикол. Окружающая среда. Саф.* 2020, 205, 111168, doi:10.1016/j.ecoenv.2020.111168.
11. Боррего, К.; Джинджа, Дж.; Коутиньо, М.; Рибейро, К.; Карацас, К.; Сиумис, Т.; Кацифаракис, Н.; Константиноидис, К.; Де Вито, С.; Эспозито, Э.; и другие. Оценка микродатчиков качества воздуха по сравнению с эталонными методами: Совместное упражнение EuNetAir – часть II. *Атмос. Окружающая среда.* 2018, 193, 127–142.
12. Конкас Ф.; Минеро, Дж.; Лагерспец, Э.; Варйонен, С.; Лю, Х.; Пуоламяки, К.; Нурми, П.; Таркома, С. Недорогой наружный воздух Мониторинг качества и калибровка датчиков: обзор и критический анализ. *ArXiv* 2021, ArXiv: 1912.06384.
13. Циммерман, Н.; Престо, АА; Кумар, СпН; Гу, Дж.; Хаурилюк, А.; Робинсон, Э.С.; Робинсон, Алабама; Субраманиан, Р. Машина обучающая модель калибровки с использованием случайных лесов для повышения производительности датчиков для более дешевого мониторинга качества воздуха. *Атмос. Изм. Тех.* 2018, 11, 291–313, doi:10.5194/amt-11-291-2018.
14. Кастелл, Н.; Лю, НУ; Шнайдер, П.; Коул-Хантер, Т.; Лахос, В.; Бартонова А. На пути к персонализированному экологическому здоровью информационный сервис с использованием недорогих датчиков и краудсорсинга. Доступно в Интернете: <https://meetingorganizer.coperni.cus.org/EGU2015/EGU2015-9058.pdf?pdf> (27 февраля 2021 г.)
15. Сайт проекта MAU AirHeritage. Доступно онлайн: <https://www.uia-initiative.eu/en/uia-cities/portici> (по состоянию на февраль 2021).
16. Проект MAU «Надежда». Доступно в Интернете: <https://www.projecthope.org/> (по состоянию на 27 февраля 2021 г.).
17. Марсельский проект MAU. Доступно онлайн: <https://www.uia-initiative.eu/en/uia-cities/aixmarseille-provence-metropole> (доступ 27 февраля 2021 г.).
18. Де Вито, С.; Массера, Э.; Пига, М.; Мартиното, Л.; Ди Франсия, Г. О полевой калибровке электронного носа для оценки бензола. в сценарии мониторинга городского загрязнения. *Сенсорные приводы В Chem.* 2008, 129, 750–757, doi:10.1016/j.snb.2007.09.060.
19. Сайт EPPELA «МОНИКА» - <https://www.eppela.com/it/projects/9652-monica-il-tuo-navigatore-personale-antismog> (ac прекращено 31 марта 2021 г.).
20. Региональное агентство по охране окружающей среды Кампани. Доступно онлайн: <https://www.arpacampania.it/> (по состоянию на 30 март 2021 г.).
21. Де Вито, С.; Ди Франсия, Г.; Эспозито, Э.; Ферлино, С.; Формисано, Ф.; Массера, Э. Стратегии адаптивного машинного обучения для сети калибровка интеллектуальных устройств мониторинга качества воздуха IoT. *Распознавание образов. лат.* 2020, 136, 264–271, doi:10.1016/j.patrec.2020.04.032.
22. Шепард, Д. Функция двумерной интерполяции для неравномерно расположенных данных. В материалах 23-го ACM На 1968 г. Национальная конференция, Нью-Йорк, США, 27–29 августа 1968 г.; стр. 517–524.
23. Ифань, Д.; Франсуаза, С.; Валери, И. IAM – Интерполяция и агрегация в движении: совместное краудсенсинг для пространственно-временных явлений. В материалах Международной конференции MobiQuitous 2020–EAI по мобильным и универсальным системам: вычисления, сети и услуги, Virtual, Германия, 31 декабря 2020 г.