Geospatial Health 8(3), 2014, pp. S631-S646

GeoMedStat: an integrated spatial surveillance system to track air pollution and associated healthcare events

Fazlay S. Faruque[1], Hui Li[1,a], Worth B. Williams[1,b], Lance A. Waller[2], Bruce T. Brackin[3], Lei Zhang[4], Kim A. Grimes[5], Richard W. Finley[6]

[1]GIS and Remote Sensing, University of Mississippi Medical Center, Jackson, Mississippi, United States of America; [2]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America; [3]Epidemiology Division, Mississippi State Department of Health, Jackson, Mississippi, United States of America; [4]Office of Health Data and Research, Mississippi State Department of Health, Jackson, Mississippi, United States of America; [5]Division of Information Systems, University of Mississippi Medical Center, Jackson, Mississippi, United States of America; [6]Department of Emergency Medicine, University of Mississippi Medical Center, Jackson, Mississippi, United States of America; [a]Moved to Reformed Seminary School, Jackson, Mississippi, United States of America; [b]Moved to Division of Information Systems, Baptist Medical Health Center, Jackson, Mississippi, United States of America

Abstract. Air pollutants, such as particulate matter with a diameter ≤2.5 microns ($PM_{2.5}$) and ozone (O_3), are known to exacerbate asthma and other respiratory diseases. An integrated surveillance system that tracks such air pollutants and associated disease incidence can assist in risk assessment, healthcare preparedness and public awareness. However, the implementation of such an integrated environmental health surveillance system is a challenge due to the disparate sources of many types of data and the implementation becomes even more complicated for a spatial and real-time system due to lack of standardised technological components and data incompatibility. In addition, accessing and utilising health data that are considered as Protected Health Information (PHI) require maintaining stringent protocols, which have to be supported by the system. This paper aims to illustrate the development of a spatial surveillance system (GeoMedStat) that is capable of tracking daily environmental pollutants along with both daily and historical patient encounter data. It utilises satellite data and the ground-monitor data from the US National Aeronautics and Space Administration (NASA) and the US Environemental Protection Agenecy (EPA), rspectively as inputs estimating air pollutants and is linked to hospital information systems for accessing chief complaints and disease classification codes. The components, developmental methods, functionality of GeoMedStat and its use as a real-time environmental health surveillance system for asthma and other respiratory syndromes in connection with with $PM_{2.5}$ and ozone are described. It is expected that the framework presented will serve as an example to others developing real-time spatial surveillance systems for pollutants and hospital visits.

Keywords: air pollution, asthma, environmental health tracking, GeoMedStat, hospital information system, ozone, $PM_{2.5}$, remote sensing, spatial.

Introduction

The quality of the air that we breathe has long been known to impact human health (Sundell, 2004; Griffin et al., 2013). Particulate matter with a diameter ≤2.5 microns ($PM_{2.5}$) and ozone (O_3) pose major health problems and are therefore among the air pollutants regulated in most countries and also guided by World Health Organization (WHO) standards. According to WHO, an estimated 3.7 million premature deaths occurred in 2012 due to ambient air pollution (WHO, 2014). Among the air pollutants, $PM_{2.5}$ affects more people than any other pollutant and is also responsible for a wide variety of adverse health conditions including respiratory problems (Dominici et al., 2006), cardiovascular disease (Brook et al., 2010), cancer (Vineis et al., 2006), birth defects (Hooven et al., 2011) and neurological disorders (Kannan et al., 2007). Several reviews summarise health impacts due to particulate matters, including those by Rucker et al. (2011), Pope and Dockery (2006) and Burnett et al. (2014). When O_3 is present in the air it can cause breathing problems and reduced lung function as well as increased mortality due to heart disease (Bell et al., 2004; Peng et al., 2013; WHO, 2014).

Both $PM_{2.5}$ and O_3 are known to contribute to increased emergency room (ER) visits, hospitalisations

Corresponding author:
Fazlay S. Faruque
GIS and Remote Sensing, University of Mississippi Medical Center
2500 North State Street, Jackson, Mississippi 39216-450, USA
Telephone +1 601 815-9027; Fax +1 601 815-5957
E-mail: ffaruque@umc.edu

and even deaths due to respiratory and cardiovascular disease (Fann et al., 2012; Kheirbek et al., 2013). In order to better understand the effects of air pollution on health and also to more efficiently address the issue, there is a need for an integrated surveillance system that can link air pollution, health and demographic data. In this paper, we describe the development of "GeoMedStat", a system that incorporates these three categories of data (pollution, health and demography) in a common geospatial format to enable routinely tracking pollution, health outcome and exposed population. Data generated from this integrated surveillance system can also be used for analysing associations between air pollution and health conditions in the population.

Asthma and air pollution

Although the fundamental causes of asthma are not completely understood, we know the condition is a chronic inflammatory disorder of the airways due to hyper-responsiveness to a variety of stimuli (Mannino et al., 1998; Reinke and Hoffman, 2000). The strongest risk factors for developing asthma are a combination of genetic predisposition and environmental exposure to inhaled substances and particles that may provoke allergic reactions or irritate the airways. While medication can control asthma, avoiding the asthma triggers can significantly reduce the severity of symptoms. WHO estimates that some 235 million people, mainly children, suffer from asthma, (WHO, 2013). Air pollution can exacerbate asthma symptoms particularly after exposure to atmospheric ozone, nitrogen dioxide, sulphur dioxide and $PM_{2.5}$ (Dockery et al., 1993; Anderson et al., 1998; D'Amato et al., 2005). Studies have shown that reducing exposure to airborne triggers eases asthma symptoms, the need for medication and can, in some cases, also improve lung function (Institute of Medicine-IOM, 2000; Carpenter, 2004). In order to reduce asthma, WHO recommends: (i) surveillance of asthma and analysis of its determinants along with population characteristics; (ii) reduction of the level of exposure; and (iii) improvement of access to intervention (WHO, 2013). As an integrated surveillance system, GeoMedStat tracks asthma cases, population demographics and exposure level to pollutants. Researchers and policymakers can utilise these data to develop intervention strategies including resource allocations and pollution control.

In the United States of America (USA), the asthma burden is higher than in many developing countries and the general trend of asthma prevalence, mortality and hospitalisation rates is rising (Beasley, 2002; Masoli et al., 2004; Braman, 2006; Moorman et al., 2007, 2012). According to a Pew Environmental Health Commission report (2000), highlighting limitations of the public health system in preventing asthma in USA, the current epidemic of asthma will grow worse over the next generation unless appropriate action is taken now. The Commission recommends actions for tracking asthma rates and studying environmental risks, and also addressing environmental causes of asthma development and exacerbation in order to achieve the goal of cutting the number of asthma cases in half by 2020. Another Pew Environmental Health Commission report (2000) states that the U.S. is facing a gap in critical knowledge regarding environmental health that hinders its national efforts to reduce or eliminate diseases that might be prevented by better managing environmental factors (McGeehin et al., 2004). GeoMedStat can be considered as an example of a system designed to address the recommendations of both these reports.

Public health surveillance systems and geographical information systems

Public health surveillance, by the most common definition in USA, is the ongoing and systematic collection, analysis and interpretation of data essential to the planning, implementation and evaluation of public health practices, which closely integrate with the timely dissemination of these data to persons responsible for disease prevention and control (Thacker and Berkelman, 1988). Geographical information systems (GIS) have been successfully used in many areas of public health including surveillance (Boulos, 2004; Shuai et al., 2006, Yiannakoulias et al., 2009; Horst and Coco, 2010). GIS-based real-time surveillance systems can be instrumental in determining when and where interventions are critical to implement. Further, such spatially enabled surveillance systems empower users to: a) better understand health outcomes in the community influenced by environmental pollution or other factors; b) optimize resource allocations; and c) respond efficiently in the event of manmade or natural health disasters in the community. However, to date, GIS-based real-time, or near real-time, surveillance systems are not common due to limitations in real-time data access, complexity in integration of different components in different formats and the difficulty in application development.

Satellite data in air pollution assessment

In current environmental public health surveillance practice, air quality data typically are obtained from fixed-site, ground-level air quality monitoring stations that tend to be located strategically in areas where high levels of pollutants are expected or where the population density is high (Nerriere et al., 2005; Bell, 2006). The main purpose of these sites is to monitor regulatory compliance. Such sites are very limited in number and are unable to provide seamless air quality data over the full range of space and time necessary for population-based studies. Furthermore, in developing countries, such ground monitoring facilities are very sporadic, irregularly spread, even non-existent. In USA, a country of 3.8 million miles^2 (about 9.8 million km^2) and a population of more than 300 million, there are only approximately 320 monitoring facilities providing daily $PM_{2.5}$ data. Remote sensing can potentially help overcoming these limitations by providing daily data for the entire globe. Since a number of satellites are collecting global data on a daily basis, remote sensing can help supplement ground monitored data resulting in seamless coverage, which should eventually generate reliable estimates of all ground level pollutants of interest.

There are multiple space-borne sensors measuring total aerosol optical depth (AOD) with daily global coverage. Scientists have explored the use of these data to estimate ground level $PM_{2.5}$ (GLP), which can fill gaps for the areas lacking monitoring stations (Engel-Cox et al., 2004; Al-Saadi et al., 2005; Kloog et al., 2012; Lary et al., 2014). Application of satellite-derived data for estimating ground-level particulates in order to study health impacts started fairly recently, but its use is increasing rapidly due to the critical need for seamless air quality data covering locations lacking ground monitoring facilities (Crouse et al., 2012; Hystad et al., 2012; Kloog et al., 2012, 2014; Madrigano et al., 2013; Wang et al., 2013). By integrating satellite data with ground-monitored air quality data, it is possible to generate continuous surface models of $PM_{2.5}$, eliminating spatial and temporal data gaps in ground data. The resulting datasets can therefore provide suitable air quality information necessary for public health studies.

Integrated surveillance systems for tracking environmental risk and asthma

Surveillance of both air pollution and asthma is crucial for developing effective prevention strategies. Systems tracking asthma are not generally compatible with systems tracking environmental hazards (whether using ground-based or remotely sensed data) because of data source type, data attributes, data capture strategy, spatial resolution and data transfer mechanisms. Due to this incompatibility, it is difficult to integrate the two systems to assist in tracking air pollution to predict, prevent and intervene with respect to diseases exacerbated by pollutants. In recent years, several initiatives and programmes have aimed to provide information on atmospheric conditions and events affecting human health, including Health and Environment Linked for Information Exchange (HELIX), GeoMedStat, Public Health Applications in Remote Sensing (PHAiRS) http://phairs.unm.edu/news/healthalert.html and the 3-Dimensional Air Quality System (3D-AQS) (Morain and Budge, 2008).

GeoMedStat is a GIS-based integrated surveillance system that tracks both air pollution and asthma hospital visits (Gotz et al., 2009; Li et al., 2009; Chen et al., 2010). This system links health data with geospatial attributes, census data and satellite air quality observations (Morain and Budge, 2008). The University of Mississippi Medical Center (UMMC) initially developed a prototype system for bioterrorism-related disease outbreak surveillance in 2004 (Li et al., 2005). In 2006, UMMC received funding support from NASA's Public Health Application Program through the Mississippi Research Commission (MRC) to develop GeoMedStat as a part of a decision-support tool. It is important to note that an institute's own resources and support are essential to sustain a system like GeoMedStat. The long-term goal of developing such a surveillance system requires continuous support for maintenance and updates with changing technology and standards for the multiple components of the system (e.g., both environmental and healthcare aspects) beyond the funding period. Over time, GeoMedStat has evolved into a more general environmental health surveillance system with current focus on respiratory disease/asthma surveillance and air quality. The GeoMedStat framework is capable of expanding to other diseases as its system architecture is flexible and can accommodate other disease records. As electronic health systems are becoming common in practice, other disease records, potentiality associated with $PM_{2.5}$ and O_3, can be integrated aiding the further explorations of their possible spatio-temporal associations. In addition, new models to estimate pollutants as environmental risk factors can easily be integrated into GeoMedStat.

Materials and methods

GeoMedStat is a web-based geospatial solution developed using ArcGIS Servers, the Web Application Developer Framework (ADF), Ajax, JavaScript and HTML technologies in a Microsoft .NET framework. GeoMedStat data are stored in an Oracle database and the system is hosted via a Microsoft IIS server. Through a user interface, pollutants and health data can be displayed as maps and charts. GeoMedStat has been developed as a real-time (daily) spatial surveillance system. The three principal data components of GeoMedStat are: air pollution data, asthma hospital visit data and census demographic data. The air pollution component includes $PM_{2.5}$ and ozone estimates. The former are based on modelling and surfacing techniques developed and provided by the NASA's Marshall Space Flight Center (MSFC). Since ground level ozone cannot be modelled from currently available satellite data, only EPA's ground data were used to generate daily ozone surfaces.

The hospital and clinic system at UMMC, known as the University of Mississippi Health Care (UMHC) system, has a live link with the GeoMedStat system allowing real-time access to daily patient visits, including those associated with respiratory problems. In addition, we also obtain annual asthma patient visit data for the entire state of Mississippi from the Asthma Program of the Mississippi State Department of Health (MSDH). We use demographic data from the ESRI Tapestry database (ESRI, 2010) to characterise the neighbourhood of the patients' residence and to identify the exposed population.

Patient encounter data are identified as records with International Classification of Disease (ICD) codes for asthma and chief complaints as respiratory syndrome (Table 1). UMMC hospital and clinic data are accessible to the GeoMedStat server via a real-time connection, the Health Level 7 (HL7). Although these data are live-linked, for the sake of process efficiency, the algorithm includes a trigger mechanism that starts communication between the UMHC and GeoMedStat data servers once a day for querying and then geocoding the selected data. Selected data from UMHC are automatically geocoded on a daily basis. In contrast, the state-wide asthma hospital visit data are geocoded separately and inserted into the database annually. Through the GeoMedStat user interface, users can display both pollutants and health data as maps and charts. For research purposes, users with privileged login can download daily data, aggregated to 10 km x 10 km grids.

Pollutant data acquisition and integration

Currently, data for the two major ambient air pollutants impacting asthma and other respiratory diseases, $PM_{2.5}$ and O_3, are integrated into GeoMedStat. However, the framework of this spatial surveillance

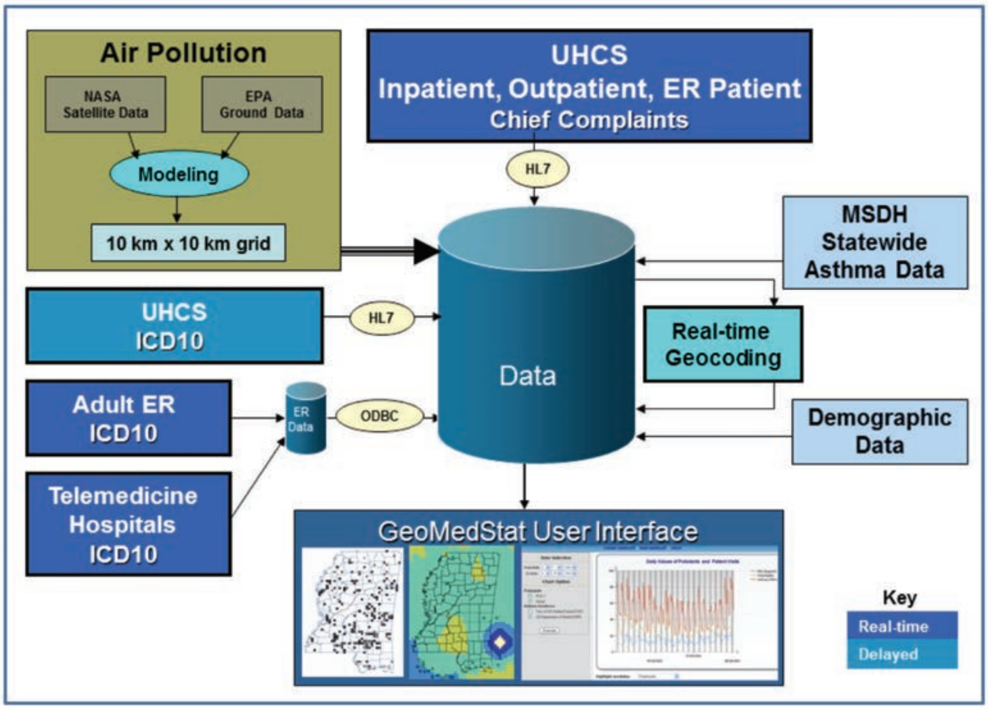

Fig. 1. Schematic description of GeoMedStat.

system is sufficiently flexible to accommodate also other relevant pollutant data with suitable formats. The availability of daily data on these pollutants can assist healthcare professionals in resource preparation and can also help individuals in limiting their outdoor activities in the event of elevated level of pollutant abundance.

As illustrated in Fig. 2, the main process for pollutant modelling is triggered to begin at a certain time of the day, currently at 2:00 AM every day. Ground-monitored $PM_{2.5}$ and ozone data for the preceding day are downloaded from the EPA AirNow website (http://www.airnowgateway.org). Request for an account to access AirNow data is required so that EPA can monitor the data access. It could be noted that although the old File Transfer Protocol (FTP)-based AirNow gateway is still available to download data, there is a new website (airnowapi.org) which is more efficient including web services, data file outputs and XML (RSS) feeds.

As Fig. 2 shows, all available monitoring stations within the study area are identified first in order to procure only the required data. After downloading, using the .NET FTP service, data are saved locally. ASCII files are prepared from the procured data to serve as input for the surface models. Utilizing these ground monitored data from EPA, daily surface models are created using a Basis Spline or B-spline method (De Boor, 2001) for ozone and $PM_{2.5}$ for the winter season (October through March) and also when no satellite data are available.

Similar to ground data, daily AOD data are also automatically procured using the .NET FTP service from the NASA website (ftp://ladsftp.nascom.nasa.gov/allData). These satellite data products are available from the Goddard Earth Sciences Data (https://earthdata.nasa.gov/data/data-centers/ges-disc) and Information Services Center (GES DISC). Files are identified for the study area based on the date and satellite path. Reference ID from the HDF

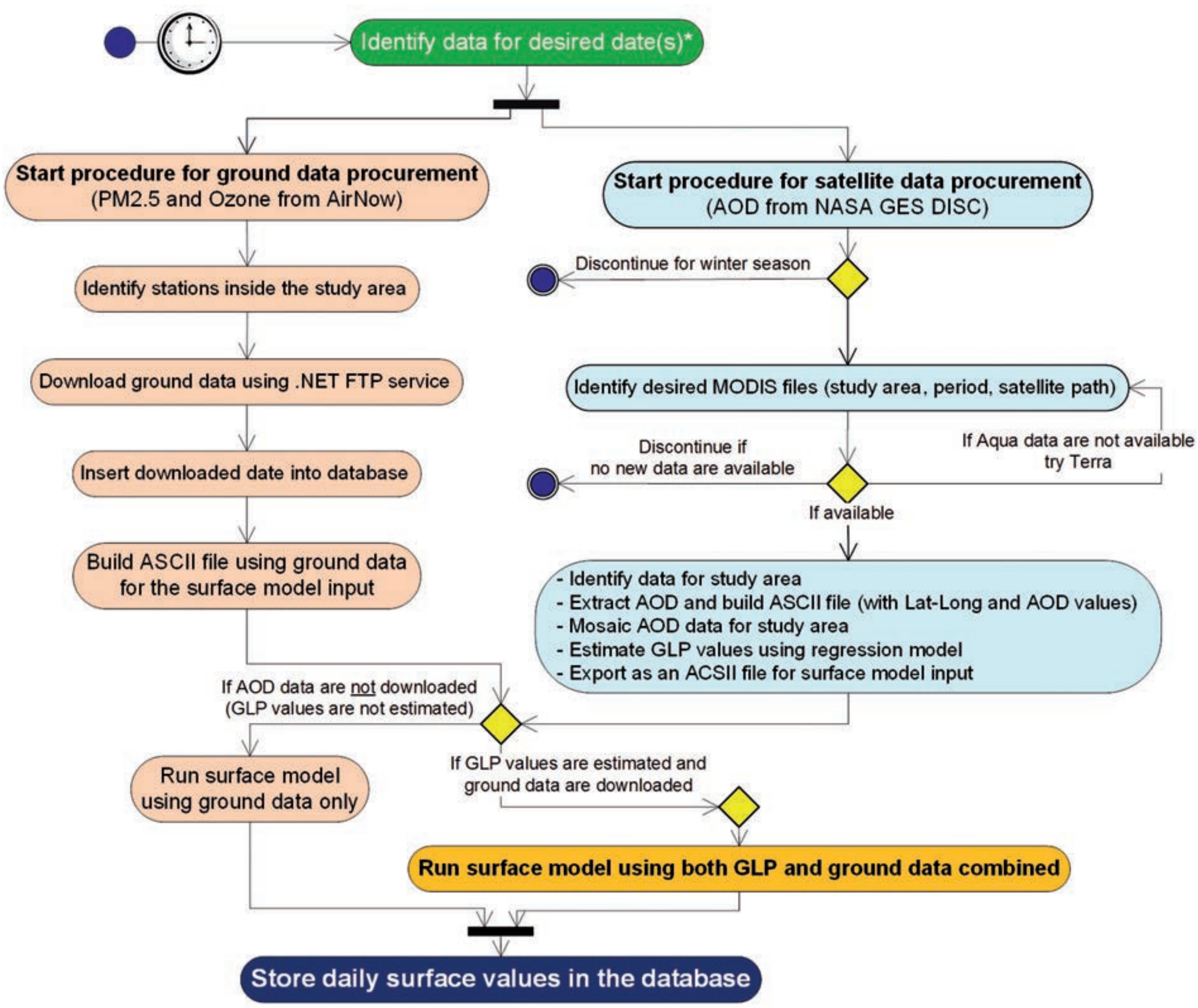

Fig. 2. Process diagram for daily air pollutant geospatial surface modeling. *In addition to data for the current day, previous days data up to 5 days are looked for if previously missed.

header file of the MODIS AOD data are used to obtain the latitude-longitude values which are used to extract data for the desired area. Individual AOD data files mosaic into a single file serving as the input for GLP estimation model for the specific study area. The next step is to run a regression model to estimate $PM_{2.5}$ from the AOD values. Finally, spatial surface model is built with estimated and ground $PM_{2.5}$ data using the B-spline method (De Boor, 2001; Al-Hamdan et al., 2009). These grid-based surface values are stored in an Oracle database to be accessed through the GeoMedStat user interface for any desired days.

While downloading, GeoMedStat first checks for availability of MODIS data from the Aqua satellite (MYD04_L2); if none are available, it checks for data from Terra (MOD04_L2) for download. If no MODIS data are available from either satellite for the desired day and location, the processes discontinues and only ground data are used to create the surface model without contribution from satellite data. We note that, due to inconsistent relationships between AOD and $PM_{2.5}$ during the winter season (October through March), the regression model is less reliable; hence GeoMedStat does not use satellite data to create the surface model during the winter. Similarly, when satellite data for given days are not available, only ground data are used for the surface model. It should be noted that the algorithm first checks if there are data missing from the previous day(s). MODIS data occasionally are missing for certain days but, if acquired, may be processed and added in later. The GeoMedStat algorithm checks if data for previous day(s) are available from the sources (the Goddard website); if so, GeoMedStat downloads the new data and completes the processing for the missing day(s). The system continues checking for missing data for up to 5 days, if data are still not available after this time then GeoMedStat shows data gaps.

$PM_{2.5}$ and O_3 from ground monitoring stations

The EPA in USA has a nationwide air-monitoring network which includes ground monitoring of $PM_{2.5}$ and O_3, the two pollutants of interest in this project. $PM_{2.5}$ data from ground monitoring stations are used: (i) for quality control assessments of satellite-derived $PM_{2.5}$ estimates; and (ii) during periods when MODIS data do not correlate with ground data. The EPA maintains two air pollutant data repositories, i.e. AQS (http://www.epa.gov/ttn/airs/airsaqs/) and AirNow (http://www.airnow.gov/). In USA, state, local and tribal agencies collect ground data from monitoring stations across the nation and submit them electronically to the EPA AQS on a periodic basis. Since monitoring pollution from ground stations is expensive and labour-intensive, stations are sparsely located resulting in the frequency of data collection not being consistent but varying from hourly to every 6 days. Archived air pollutant data are available for downloading from the AQS web site. The data in AQS are considered to be of the highest quality available because these data pass several quality assurance tests (http://www.epa.gov/ ttn/airs/aqsdatamart/basic_info.htm). AQS $PM_{2.5}$ data are used to establish the regression models for estimating GLP of the same dimension from MODIS AOD data. On the other hand, GeoMedStat employs data from AirNow, a "real-time" air pollutant data portal in USA for developing its daily geospatial surface models of $PM_{2.5}$. The AirNow system has been developed by the EPA, the National Oceanic and Atmospheric Administration (NOAA), National Park Service (NPS) and tribal, state and local agencies to provide the public with easy access to national air quality information on a "real-time" basis. Since the purpose of the AirNow gateway is to provide information as close to "real-time" as possible, these data are reported with preliminary data quality assessments that are not as stringent as AQS. An FTP client class allows automatic download of data from the AirNow gateway using .NET FTP API functions. Ground measurements of air quality are stored in an ASCII file for each observed day. Because an air quality data file contains nationwide ground observations from over 400 cities, for this project it is necessary to extract data for only those stations inside Mississippi and a 220-km buffer around its borders. Ground data beyond the border of Mississippi were used to maintain data quality consistency by avoiding spatial edge effects.

$PM_{2.5}$ estimation from satellite data sources

NASA's Terra and Aqua satellites orbit the globe every day, each carrying a MODIS sensor. Because both satellites pass the study area at a defined time every day, the data acquisition time-stamp of MODIS is used to identify data for the study area so that only these data are downloaded. AOD is one of the products generated from the electromagnetic radiation values captured by MODIS. The AOD data are stored in a hierarchical data format and are available for automatic download from the NASA-Goddard Earth

Sciences FTP site using .NET FTP API functions. Based on a regression model, developed at NASA MSFC, ground level $PM_{2.5}$ values are estimated from these AOD data (Al-Hamdan et al., 2009). Due to inherent differences in data quality, regression models were developed separately for Terra and Aqua MODIS data.

Data integration and surface modelling

Both GLP, the estimated ground level particle matter as a proxy for $PM_{2.5}$, and O_3 are mapped as continuous daily spatial surfaces using B-spline algorithm. Currently, the ozone source data are obtained only from the ground monitoring stations. Due to extrapolation between monitoring stations, the quality of ozone estimations is compromised in the areas without nearby monitoring stations. It is expected that when reliable ground level ozone can be estimated from satellite data, that model will be integrated into the GeoMedStat to improve the quality of the ozone surface data. For GLP, the system merges the estimated data and EPA's ground-monitored $PM_{2.5}$ data (from AirNow) resulting in seamless coverage of the entire study area. The original codes for running the $PM_{2.5}$ estimation model, B-Spline surfacing and bias removal were originally developed by NASA's MSFC (Al-Hamdan et al., 2009), which were recoded at UMMC to implement as an automated real-time system for the GeoMedStat. The pollutant surface outputs are resolved at a spatial resolution comparable to that of the original MODIS-based AOD data product, i.e., approximately 10 km x 10 km or 0.1^{o} x 0.1^{o} grids.

Health and demographic data: sources and integration

Health data from two sources were integrated into GeoMedStat - UMMC Hospital and Clinic System (UHCS) data from UMMC's own hospitals and clinics and MSDH data from hospitals across the state.

Real-time UHCS data

Since the first implementation of GeoMedStat, the UHCS patient data management has gone through some changes which are not uncommon for hospital systems. The original schema is described here because it was more complicated than the current system and many hospitals may still use such disparate components of patient data management system. Hospital utilisation data from UHCS, which include Admit-Discharge-Transfer (ADT) data from all hospitals and clinics at UMMC, are accessed and categorised based on the chief complaint of patients during their registration. ADT data, particularly the ER data, are commonly used as a proxy for disease incidence data in many exposure association studies (Zanobetti, 2009). The ADT data are considered here as the healthcare events. When GeoMedStat was first implemented, UMMC ADT data were managed by a system known as the 'Invision Hospital Information System', developed by the Siemens Company. HL7, an industry-standard application protocol for secure exchange, integration, sharing and retrieval of electronic health information, was used as the real-time interface to transfer data from the Invision system to the GeoMedState database server.

Since real-time ADT records contain free text of the patient's presenting complaints without a standardised specific disease diagnosis, it was necessary to categorise complaints into appropriate disease syndromes. A Bayesian-based syndrome classifier (Wagner et al., 2004) developed at the University of Pittsburgh was used to classify free-text chief complaints into different syndrome categories, including asthma, the respiratory disease syndrome of current interest. Shortly after patient diagnosis in the emergency department of the Adult Emergency Hospital at UMMC, physicians generate a specific ICD9 diagnosis code using "EMstation", an electronic record-keeping system. ICD9 data from this source are unique in that they are available in real-time (similar data from other UMMC sources are delayed since codes first must be trans-

Table 1. Patient data input and characteristics.

Source	Disease attribute	Temporal integration	Access/transfer
UMHC (all)	ICD-9 for asthma	Approximately a week delay	HL7
UMHC (ER and Telemedicine)	ICD-9 for asthma	Daily	ODBC
UMHC (all)	Syndromic, chief complaints for respiratory	Daily	HL7
MSDH	ICD-9 for asthma	Annual	Manual

*This table shows the original schema. After recent implementation of the EPIC system, it has been simplified as ER and Telemedicine data are also accessed and transferred by EPIC.

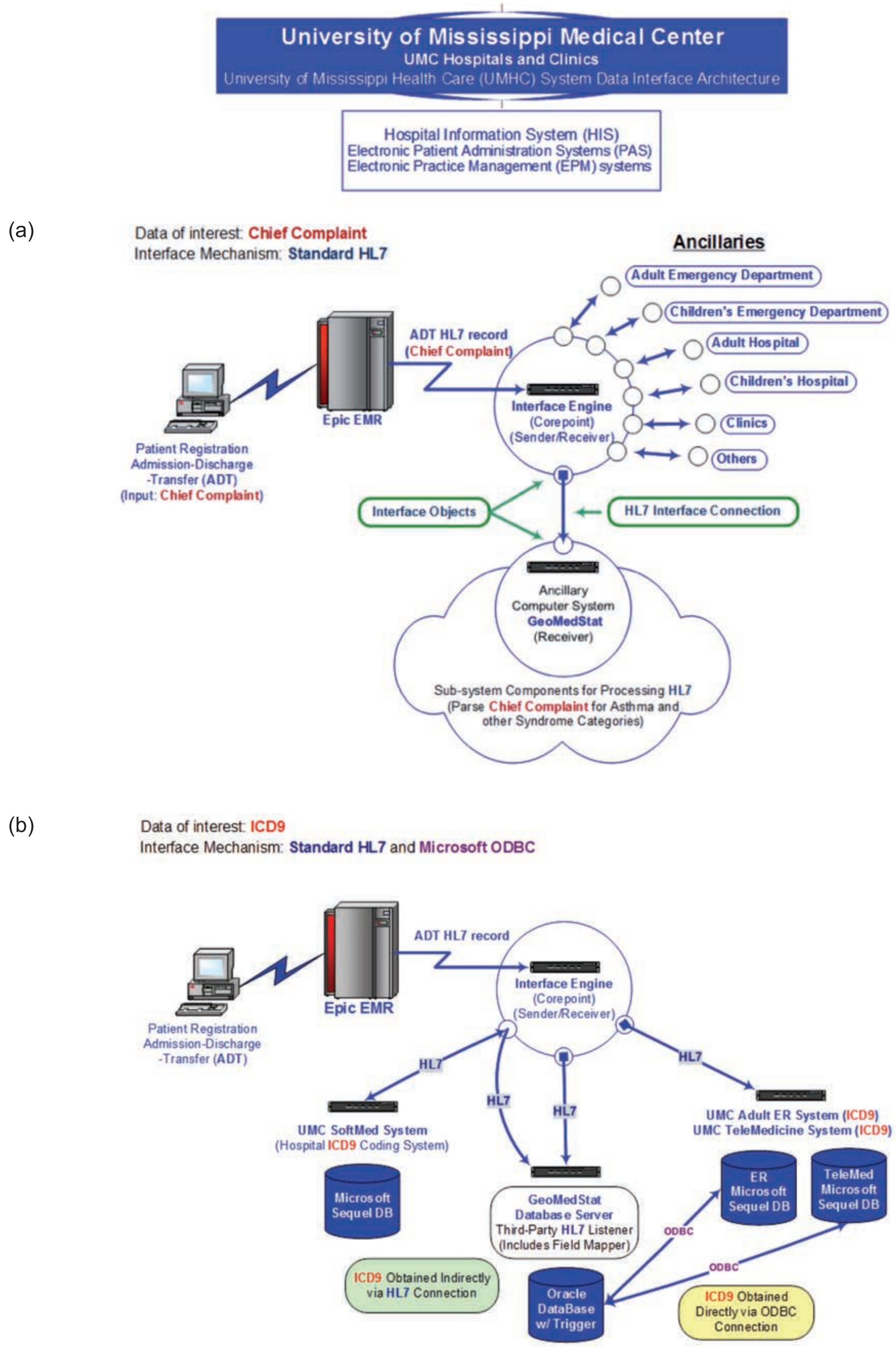

Fig. 3. (a) Data access and transfer mechanism for UMHC patient data with chief complaints. (b) Data access and transfer mechanism for UMHC patient data with ICD codes. *This figure shows the original schema. After implementing EPIC system, all UMHC data, including ICD 9 from ER and TeleMedicne, are accessed and transferred by EPIC.

ferred from written charts to electronic record-keeping systems). Each patient's specific ICD9 diagnosis codes are directly integrated into GeoMedStat using Open Database Connectivity (ODBC). Both the ADT data with chief complaints and ER data with ICD9 codes are stored in an Oracle database for further management and utilisation by GeoMedStat. Like most health care systems, the UMMC hospitals and clinics system has many components to ensure a secured and efficient flow of data from multiple sources. Fig. 3a shows a schematic diagram of the UMMC data transfer mechanism for chief complaint data, from patient registration to incorporation of data into GeoMedStat. Fig. 3b illustrates a similar mechanism for the transfer of ICD9 data from the hospital utilisation system via HL7 and from the Adult ER and TeleMedicine systems directly via Microsoft ODBC. It should be noted that after recent implementation of the EPIC system, EMstation is no longer in use for ER and TeleMedicne patient data input; these data are now directly accessed (without requiring an additional ODBC) by the GeoMedStat via EPIC to store and operate in its Oracle database.

MSDH data

The main objective of incorporating MSDH data into GeoMedStat was to generate hospital admission data and air pollution data at the same spatio-temporal scale that can be utilised for association analyses. GeoMedStat has been integrating state-wide asthma hospital visit data since the MSDH Asthma Program began collecting statewide data in 2003. So far, we have obtained hospital visit data from MSDH with ICD9 code for 493.xx (.xx is a sub classifier for asthma) for the years 2003-2011. These data include inpatient, outpatient and ER hospital visits, but not clinic visits. The number of patients ranges between 20 and 30 thousands per year. Along with the ICD codes, these patient records contain fields for the residence address, race, gender, date of admission, date of discharge and type of admission. Out of the three types of admission (inpatient, outpatient and ER) the latter is considered as the most representative admission in response to exposure to pollutants or other risk factors. While the state-wide asthma data provide good representations of the population in the state of Mississippi, it should be noted that these data do not reflect 100% representation of the state for two reasons: first, although a majority of the state's hospitals participated in reporting their asthma patients, a few did not until reporting became mandatory by Mississippi Code Annotated § 41-63-4 (https://law. resource.org/pub/us/code/ms/ms.xml.2010/2010/title-41/63/41-63-4/index.html) effective from January 1, 2010; and second, patients near the border of other states may have visited hospitals in neighbouring states.

Geocoding

Both MSDH state-wide data and UHCS data were geocoded at the street address level. Geocoding by street address instead of ZIP code provides more spatial sensitivity and accurate placement within the correct 10 x10 km^2 grid cells of the dataset. However, records not matched at the street level were geocoded at the ZIP code level, if possible. We obtained street level geocoding for 85% to 92% of the records; with better matching in recent years, indicating an improvement in data quality.

In order to provide a real-time capability for mapping patient distributions, we developed a real-time address-based service for geocoding UHCS patient data. For performance improvement purpose, the service runs once a day in batch mode on the patient records admitted on the previous day. This service includes the following five steps (Fig. 4):

(i) Trigger - a scheduled service that triggers the process at a fixed time (now at 2:00 AM) each day;
(ii) Identify required data - query patient records from the hospital database based on ICD 9 code and chief complaints;
(iii) Record extraction - extract selected patient records;
(iv) Geocoding - first conduct street-level and then at ZIP code level geocoding for the remaining unmatched records; and
(v) De-identification - obtain grid ID and county FIPS by overlaying the geocoded point features on the grid and county polygons.

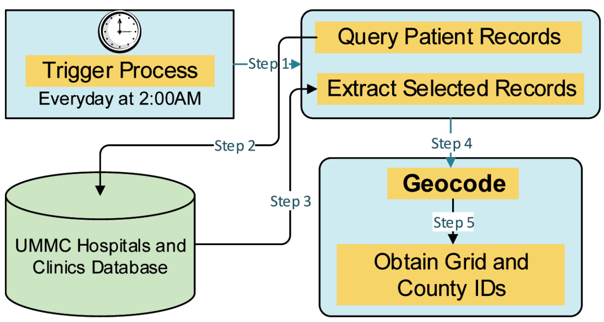

Fig. 4. Schematic representation of the autmated geocoding process.

The GeoMedStat database thus stores the grid ID and county FIPS code for each patient record instead of their address and point location. The entire process runs automatically and is no longer accessible even by the developers unless there is a need for trouble shooting. GeoMedStat users do not have access to any personal health information (PHI) but only the respective county and grid information. Since the pollutants and demographic data are transferred into grids, knowing the respective grid ID is most essential.

Demographic data

Population and socioeconomic demographics impact the hospital visits, requiring data on demographic distributions within each grid cell available to the researchers. A recent study in Mississippi found asthma hospitalisation rates to be significantly higher among all demographic groups in the rural Delta region compared with the urban Jackson metropolitan statistical area, but with the highest rates among blacks and females in both regions (Roy et al., 2010). Block group level demographic data from the ESRI Tapestry dataset (ESRI, 2010), based on the USA Census Bureau, were used for GeoMedStat.

Data integration into a common resolution

GeoMedStat prepares daily GLP surface values from MODIS data at a 10 km x 10 km spatial resolution, analogous to that of the original MODIS-based AOD data product. The ozone surfaces from the ground monitored data are also generated at the same scale to maintain the spatial compatibility. In order to link and analyse all data of interest, other datasets are also scaled to the same resolution (Fig. 5). Patient data are aggregated in the same size grid by spatial overlay operation. Block group demographic data are interpolated to the 10 km x 10 km grid resolution based using an approach suggested by Deichmann and colleagues (2001).

User interface

Maps and graphs (charts) represent the two major user interface functions in GeoMedStat. The interface was designed to include mapping options of pollutants and patient data on different temporal and spatial scales. Patient visits can be mapped by age, race and gender on grid, ZIP code or county. Secondary layers, such as streets or cities can be mapped as background. Common mapping tools, such as "zoom in", "zoom out", "identify" and "scale", are also available for the users (Fig. 6). The interface allows data graphics (chart) capability (Figs. 7a and 7b) with options of displaying asthma and pollution data at varying temporal scales; by age, race or gender groups; and also by one or multiple datasets at the same time.

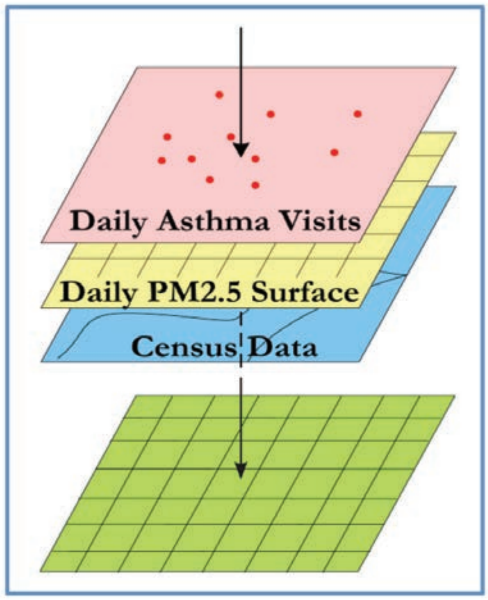

Fig. 5. Data integration at 10 km x 10 km grid resolution.

Software used for application solutions

GeoMedStat was developed using ASP.NET, Microsoft IIS and ArcGIS Server technology. GeoMedStat can be accessed remotely through a web interface using Microsoft Internet Explorer or any other compatible web browser. Once the application is loaded, users can make multiple requests through different menu options without refreshing the page or having to navigate to different pages. This is made possible through the Ajax capability of ASP.NET. This capability increases web application performance and response times. Language support in .NET allowed the FORTRAN code of the surface model to be modified, compiled and run, although the language service (Intel® Fortran Compiler 9.1 with International Mathematics and Statistics Libraries (IMSL) must be purchased separately. Microsoft IIS was used as the foundation for the web server.

User access, authentication and authorisation is validated through user accounts maintained in the server and predefined account roles are used to enable access to differing levels of GeoMedStat features for doctors, epidemiologists, patients and asthma coalition personnel depending on their needs. Oracle was chosen as the storage database for all GeoMedStat data. Table views were created to provide behind-the-screen queries, such as unions and joins for aggregating data. Real-time interfaces for ground-level and satellite data were handled through FTP services. To acquire hospital data, the real-time interface utilises a third-party HL7 interface program that selects records containing the desired syndrome during the desired period.

The ArcGIS server provided GIS capabilities such as spatial joins, mapping, mapping tools and symbols. Web ADF Common Data Source APIs were used to access ArcGIS server data sources (e.g. ArcGIS server map services). The ArcGIS server ArcObjects API was used to perform geoprocessing tasks, such as overlays and spatial analysis. Charting capabilities for displaying patient visits or pollutants over time were developed using Dundas Chart for ASP.NET.

Outcome examples

As a real-time spatial and integrated surveillance system, GeoMedStat currently allows the users to map and graph daily: (i) hospital visits reporting respiratory syndrome or diagnosed with asthma; (ii) grid level ozone and GLP values; and c) historical state-wide asthma hospital visits. These data can be accessed either for adults or for children, or combined. GeoMedStat aggregates all data on the same 10 km x 10 km grid as the pollutants. No individual record can be accessed and no link with the admitting hospital is identified. Record level individual data can be made available to researchers through a special protocol.

Although patient data are live-linked with the GeoMedStat server, for practical purposes, these data are accessed and added to GeoMedStat once a day. As detailed above, the AOD data from MODIS are available on the next day. $PM_{2.5}$ and O_3 data are obtained from EPA's AirNow corresponding to the satellite data collection time. Hence, GeoMedStat displays both patient and pollutant data delayed by one day. The map in Fig. 6 shows the distribution, by residence location at grid level, of patients, who visited the UMMC hospital from January 6 to March 6, 2014. The same period is shown as a graph in Fig. 7b along with the daily $PM_{2.5}$ and ozone.

Fig. 7a shows daily $PM_{2.5}$, ozone and asthma hospital visit data for the period of December 1, 2012 to July 31, 2014 revealing a general trend of higher number of admissions during cold seasons (could be due to infections instead of pollution levels). The graph of a shorter period, from January 6 to March 6, 2014, of the same variables is shown on Fig. 7b. Note that dur-

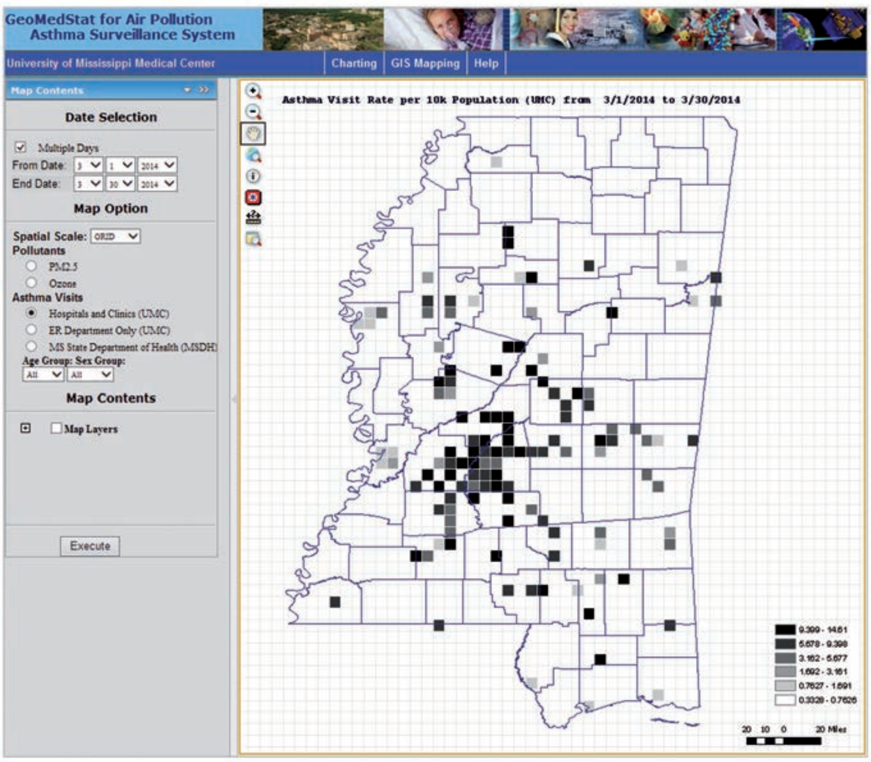

Fig. 6. Example of mapping capability showing patient rates per ten thousand populations at the grid level who visited the UMMC hospital for asthma from January 6 to March 6, 2014.

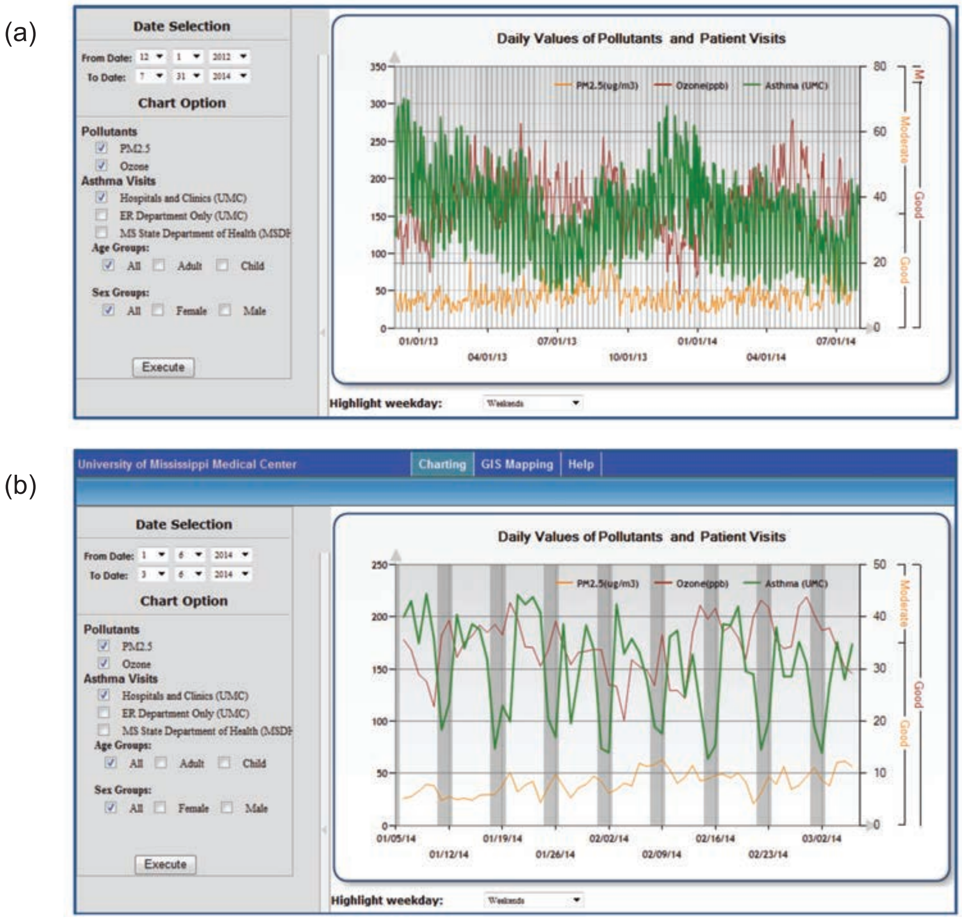

Fig. 7. (a)Daily $PM_{2.5}$, ozone, and asthma hospital visit data for the period of December 1, 2012 to July 31, 2014 revealing a general trend of higher number of admissions during the winter. (b) Graph showing daily $PM_{2.5}$, ozone and asthma hospital visit data from January 6 to March 6, 2014 for the same period as shown on the map of Fig. 6. The left Y-axis shows the raw count of hospital visits per grid. The right Y-axis shows $PM_{2.5}$ in micrograms per cubic meter and ozone in parts per billion.

ing the weekend (shaded bar) the number of hospital visits decreases. We suspect this is not related to pollution exposure, but due to perceptions that fewer healthcare resources are available during weekends. This is consistent with studies of health service utilisation by day of the week that report fewer visits to in-hospital ambulatory urgent care clinics on weekend days (Batal et al., 2001; Becker, 2007). It could be noted that while studying the association between asthma hospital visits and pollutant exposures, the SES of the patients and their community and the local meteorological conditions along with the lag time between the exposure and hospital admission need to be accounted for. However, GeoMedStat does not include any tool for association analysis. While GeoMedStat can provide spatio-temporal data on pollutant, patient and community characteristics, researchers need to develop their study designs and conduct their analyses accounting for all other necessary factors to examine the association.

Discussion

The framework presented in this paper serves as guidance for others in developing a real-time integrated spatial surveillance system for pollutant and health data. To the best of our knowledge, GeoMedStat is the first system directly accessing ADT data via HL7 for real-time mapping and charting patient encounter data. Accessing patient data, automation of all processes, maintaining data confidentiality, application development for both patient and pollutant data interface, as well as sustaining the system for routines and updates are some of the major challenges that were encountered during the developmental and operational phases of GeoMedStat.

GeoMedStat is an example of a state-wide surveillance system that tracks health outcome and environmental risks, aligned with the recommendation of the Pew commission report for reducing knowledge gaps in America's environmental health by tracking disease rates and environmental risk factors. This surveillance system also aligns with the US Government's initiative on environmental tracking (a follow up of the Pew Commission's report), which is being implemented by the Centers for Disease Control (CDC) (McGeehin, et al., 2004).

Since GeoMedStat utilises patient data, it is mandatory to maintain patient confidentiality during data processing. It is essential to pre-determine in the development phase who will access data, how much data can be accessed, how the data will be de-identified and how data should be stored. In addition, the geographic context of the patient data, patient's street address and five digit zip code are both considered as protected information. To address these concerns, GeoMedStat includes password protection and geocoded data are aggregated before they become accessible. Patient data used in GeoMedStat are securely stored by the Division of Information System (DIS) maintaining same protocol as other confidential data. All patient data processing, including geocoding, is automatically performed in the background without requiring any data access. No record level, individually identifiable data can be accessed through GeoMedStat. However; such data can be requested from the original source through proper protocol. Although privacy protection policies may differ in other countries, any system like GeoMedStat must be designed to fully accommodate the protection of confidentiality of the patients. In addition, a system like GeoMedStat has to be accommodating contemporary scientific and technological developments and standards. For example, more GLP estimation models are becoming available, both globally (van Donkelaar et al. 2010; Lary et al., 2014) and regionally (Liu et al., 2009; Weber et al., 2010). Most of these new estimates account for many more parameters including non-liner relationships, and are therefore more accurate than simple regression-based models and can be used throughout the year irrespective of the season "(Lary et al., 2014; Nguyen et al., 2014). The GeoMedStat's pollution estimation module has to be capable of integrating the best available model that is practically possible to implement as real-time (in this case daily) data input. Patient data management technology and standards are evolving quickly, as well. As a result, surveillance systems must be flexible and able to adapt to changes in patient data management systems. For example, GeoMedStat has recently adapted to changing from Siemens Invision System to EPIC (http://www.epic.com/), a more modern integrated electronic health system. During this change, it took about four months to implement new hospital patient data management system within GeoMedStat, but the system had to remain operational during the transition. In another case, it is taking more than two months for GeoMedStat to transit from ICD9 to ICD10 and the transition is currently on hold due to some compatibility issues. It is likely that there will be additional changes in GeoMedStat to adjust to the associated patient data management systems or to further improve its functionality. A major criterion of success and sustainability of this type of system utilising disparate data and technology is the ability to adapt to new standards and changing technology.

Electronic health records are becoming common practice in most of developed countries, a situation that has a tremendous potential for wider use of surveillance systems like GeoMedStat. Association of multiple diseases with possible environmental exposures can be explored when spatio-temporally stamped data become readily available to researchers. In the future, if these electronic health records become real-time, such surveillance system will be instrumental for fruitful interventions knowing where and what is happening related to public health. It is recommended that such surveillance systems be scalable. Modelled pollutants, such as $PM_{2.5}$, are already available at the global scale. With newer satellites and improved retrieval algorithms for satellite data, it is also possible to generate pollutant estimates at different resolutions for local applications. It is desirable that the resolution be generic enough to accommodate other possible satellite-derived estimates, such as mold spores and pollen. In the future, multiple independent surveillance systems may need to be integrated to generate better resolution information. However, until a uniform standard for ADT data is adopted by the hospitals, patient data integration on real-time will remain a challenge.

There are certain limitations of GeoMedStat. Hospital data linkage and utilisation are still quite complicated requiring continuous checks for any interruption data flow interuption. Quality of hospital patient data should be verified before making any critical decision. One common issue is the duplication of records due to registration of the same patient in more than one category, e.g., inpatient, outpatient or ER. In addition, interface speed is still not optimal for rapid analysis and decisions. A new application development environment is under consideration to improve speed and user-friendliness. Pollutant data gaps occur when sensors are unable to capture images due to cloud cover or malfunctions; hence no AOD data are generated by NASA for those days. Most of these limitations will improve over time with the availability of newer technology and data products.

Conclusions

GeoMedStat has been successfully established as a framework for tracking pollutants and diseases on a daily basis. It generates spatio-temporal data to study the association of healthcare events (both real-time and historical) with environmental exposures in order to assist in prevention and resource management for improved public health. Within the current framework, GeoMedStat has demonstrated the potential of public health surveillance systems to track disease incidence and environmental risks as well as facilitating the explorations of important environmental factors that may contribute to disease development or exacerbations. In the future, satellite data that make possible more accurate estimates of ground-level $PM_{2.5}$ and other pollutants will allow GeoMedStat to play an expanded role in pollution tracking, environmental hazard prediction and the development of appropriate public health responses.

Acknowledgments

Development of GeoMedStat was funded by NASA through MRC. The original $PM_{2.5}$ modeling and surfacing methodologies were developed by NASA MSFC and provided to UMMC through a subcontract. MSDH Asthma Program has provided the state-wide asthma data. During the original system design, Dr. Judith Qualters from CDC provided valuable suggestions. Since 2008, UMMC's DIS is supporting the maintenance of GeoMedStat utilising UMMC's internal resources.

References

Al-Hamdan MZ, Crosson WL, Limaye AS, Rickman DL, Quattrochi DA, Estes Jr MG, Qualters JR, Sinclair AH, Tolsma DD, Adeniyi KA, 2009. Methods for characterizing fine particulate matter using ground observations and remotely sensed data: potential use for environmental public health surveillance. J Air Waste Manag Assoc 59, 865-881.

Al-Saadi J, Szykman J, Pierce RB, Kittaka C, Neil D, Chu DA, Remer L, Gumley L, Prins E, Weinstock L, 2005. Improving national air quality forecasts with satellite aerosol observations. Bulletin of the American Meteorological Society 86, 1249-1261.

Anderson H, de Leon AP, Bland J, Bower J, Emberlin J, Strachan D, 1998. Air pollution, pollens, and daily admissions for asthma in London 1987-1992. Thorax 53, 842-848.

Batal H, Tench J, McMillan S, Adams J, Mehler PS, 2001. Predicting patient visits to an urgent care clinic using calendar variables. Acad Emerg Med 8, 48-53.

Beasley R, 2002. The burden of asthma with specific reference to the United States. J Allergy Clin Immunol 109, S482-S489.

Becker DJ, 2007. Do hospitals provide lower quality care on weekends? Health Serv Res 42, 1589-1612.

Bell ML, 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: a case study of ozone in the Northern Georgia region of the United States. Environ Int 32, 586-593.

Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F, 2004. Ozone and short-term mortality in 95 US urban communities, 1987-2000. JAMA 292, 2372-2378.

Boulos MN, 2004. Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom. Int J Health Geogr 3, 1.

Braman SS, 2006. The global burden of asthma. Chest 130, S4-S12.

Brook RD, Rajagopalan S, Pope III CA, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin F, Hong Y, Luepker RV, Mittleman MA, et al., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 121, 2331-2378.

Burnett RT, Pope III CA, Ezzati M, Olives C, Lim SS, Mehta S, Shin HH, Singh G, Hubbell B, Brauer M, et al., 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect 122, 397-403.

Carpenter DO, 2004. Clearing the air: asthma and indoor air exposure. J Natl Med Assoc 96, 1115.

Chen H, Zeng D, Yan P, 2010. Data visualization, information dissemination, and alerting, infectious disease informatics. Springer, 73-87 pp.

Crouse DL, Peters PA, van Donkelaar A, Goldberg MS, Villeneuve PJ, Brion O, Khan S, Atari DO, Jerrett M, Pope CA, et al., 2012. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. Environ Health Perspect 120, 708-714.

D'Amato G, Liccardi G, D'Amato M, Holgate S, 2005. Environmental risk factors and allergic bronchial asthma. Clin Exp Allergy 35, 1113-1124.

De Boor C, 2001. A Practical Guide to Splines. New York: Springer, 348 pp.

Deichmann U, Balk D, Yetman G, 2001. Transforming population data for interdisciplinary usages: from census to grid. Washington (DC): Center for International Earth Science Information Network.

Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, Ferris Jr BG, Speizer FE, 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329, 1753-1759.

Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL, Samet JM, 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases.

JAMA 295, 1127-1134.

Engel-Cox JA, Hoff RM, Haymet A, 2004. Recommendations on the use of satellite remote-sensing data for urban air quality. J Air Waste Manag Assoc 54, 1360-1371.

Environmental Health Tracking Project Team, 2000. America's environmental health gap: why the country needs a nationwide health tracking system. Baltimore: John Hopkins School of Hygiene and Public Health.

ESRI, 2010. 2010 Methodology Statement: Esri® Data-Tapestry™ Segmentation for Block Groups and ZIP+4s. Redlands, CA. Available at: http://downloads.esri.com/support/whitepapers/other_/17032010_USA_ESRI_Tapestry_Segmentation_Methodology.pdf (accessed on November 2014).

Fann N, Lamson AD, Anenberg SC, Wesson K, Risley D, Hubbell BJ, 2012. Estimating the national public health burden associated with exposure to ambient $PM_{2.5}$ and ozone. Risk Anal 32, 81-95.

Gotz C, Frates J, Weschler S, 2009. Geographic information systems (GIS) in health organizations: emerging trends for professional education. California Geographer 49, 87-106.

Griffin DW, Naumova EN, McEntee JC, Castronovo D, Durant JL, Lyles MB, Faruque FS, Lary DJ, 2013, Air quality and human health. In: Environmental tracking for public health surveillance. Morain, SA, Budge AM (eds). CRC Press, 129-185 pp.

Horst MA, Coco AS, 2010. Observing the spread of common illnesses through a community: using geographic information systems (GIS) for surveillance. JABFP 23, 32-41.

Hu Z, Rao KR, 2009. Particulate air pollution and chronic ischemic heart disease in the eastern United States: a county level ecological study using satellite aerosol data. Environ Health 8, 26.

Hystad P, Demers PA, Johnson KC, Brook J, van Donkelaar A, Lamsal L, Martin R, Brauer M, 2012. Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study. Environ Health 11, 22.

Institute of Medicine-IOM, Committee on the Assessment of Asthma and Indoor Air, 2000. Clearing the air: asthma and indoor air exposures. Washington DC: National Academy Press.

Kannan S, Misra DP, Dvonch JT, Krishnakumar A, 2007. Exposures to airborne particulate matter and adverse perinatal outcomes: a biologically plausible mechanistic framework for exploring potential. Cien Saude Colet 12, 1591-1602.

Kheirbek I, Wheeler K, Walters S, Kass D, Matte T, 2013. $PM_{2.5}$ and ozone health impacts and disparities in New York City: sensitivity to spatial and temporal resolution. Air Qual Atmos Health 6, 473-486.

Kloog I, Melly SJ, Ridgway WL, Coull BA, Schwartz J, 2012. Using new satellite based exposure methods to study the association between pregnancy $PM_{2.5}$ exposure, premature birth and birth weight in Massachusetts. Environ Health 11, 40.

Kloog I, Nordio F, Zanobetti A, Coull BA, Koutrakis P, Schwartz JD, 2014. Short term effects of particle exposure on hospital admissions in the Mid-Atlantic states: a population estimate. PLoS One 9, e88578.

Kloog I, Ridgway B, Koutrakis P, Coull BA, Schwartz JD, 2013. Long- and short-term exposure to $PM_{2.5}$ and mortality: using novel exposure models. Epidemiology 24, 555-561.

Lary DJ, Faruque F, Malakar N, Moore A, Roscoe B, Adams Z, Eggelston Y, 2014. Estimating the global abundance of ground level presence of microscopic particulate matter. Geospat Health 8, S611-S630.

Li H, Faruque F, Williams W, Al-Hamdan M, Luvall J, Crosson W, Rickman D, Limaye A, 2009. Optimal temporal scale for the correlation of AOD and ground measurements of $PM_{2.5}$ in a real-time air quality estimation system. Atmos Environ 43, 4303-4310.

Li H, Faruque F, Williams W, Finley R, 2005. Application of geographic information systems to syndromic surveillance. In: Geographic information system: an introduction. Bandamutha RKJ, Saxena P (eds). Hyderabad: ICFAI University Press, 114-121 pp.

Liu Y, Paciorek CJ, Koutrakis P, 2009. Estimating regional spatial and temporal variability of $PM_{2.5}$ concentrations using satellite data, meteorology, and land use information. Environ Health Perspec 117, 886-892.

Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J, 2013. Long-term exposure to $PM_{2.5}$ and incidence of acute myocardial infarction. Environ Health Perspect 121, 192-196.

Mannino DM, Homa DM, Pertowski CA, Ashizawa A, Nixon LL, Johnson CA, Ball LB, Jack E, Kang DS, 1998. Surveillance for asthma-United States, 1960-1995. MMWR CDC Surveill Summ 47, 1-27.

Masoli M, Fabian D, Holt S, Beasley R, Global Initiative for Asthma (GINA) Program, 2004. The global burden of asthma: executive summary of the GINA Dissemination Committee report. Allergy 59, 469-478.

McGeehin MA, Qualters JR, Niskar AS, 2004. National environmental public health tracking program: bridging the information gap. Environ Health Perspect 112, 1409-1413.

Moorman JE, Akinbami LJ, Bailey CM, Zahran HS, King ME, Johnson CA, Liu X, 2012. National surveillance of asthma: United States, 2001-2010. Vital Health Stat 3, 1-67.

Moorman JE, Rudd RA, Johnson CA, King M, Minor P, Bailey C, Scalia MR, Akinbami LJ, 2007. National surveillance for asthma-United States, 1980-2004. MMWR Surveill Summ 56, 1-54.

Morain SA, Budge AM, 2008. Environmental sensing and human health, advances in photogrammetry, remote sensing and spatial information sciences: 2008 ISPRS Congress Book. CRC Press, 397 pp.

Nerriere É, Zmirou-Navier D, Blanchard O, Momas I, Ladner J, Le Moullec Y, Personnaz M-B, Lameloise P, Delmas V, Target A, 2005. Can we use fixed ambient air monitors to estimate population long-term exposure to air pollutants? The case of spatial variability in the Genotox ER study. Environ Res 97, 32-42.

Nguyen T, Ta V, Le T, Mantovani S, 2014. Particulate matter concentration estimation from satellite aerosol and meteorological parameters: data-driven approaches. In: Knowledge and Systems Engineering. Huynh VN, Denoeux T, Tran DH, Le AC, Pham SB (eds). Springer International Publishing 244, 351-362.

Peng RD, Samoli E, Pham L, Dominici F, Touloumi G, Ramsay T, Burnett RT, Krewski D, Le Tertre A, Cohen A, et al., 2013. Acute effects of ambient ozone on mortality in Europe and North America: results from the APHENA study. Air Qual Atmos Health 6, 445-453.

Pew, 2000. Pew environmental health commission attack asthma: why America needs a public health defense system to battle environmental threats. Pew environmental health commission, Baltimore, MD. The Pew Charitable Trust. Johns Hopkins School of Public Health. Available at: http://healthyamericans.org/reports/files/healthgap.pdf (accessed on June 2014).

Pope III CA, Dockery DW, 2006. Health effects of fine particulate air pollution: lines that connect. J Air Waste Manag Assoc 56, 709-742.

Reinke LF, Hoffman L, 2000. Asthma education: Creating a partnership. Heart Lung 29, 225-236.

Roy SR, McGinty EE, Hayes SC, Zhang L, 2010. Regional and racial disparities in asthma hospitalizations in Mississippi. J Allergy Clin Immunol 125, 636-642.

Ruckerl R, Schneider A, Breitner S, Cyrys J, Peters A, 2011. Health effects of particulate air pollution: a review of epidemiological evidence. Inhal Toxicol 23, 555-592.

Shuai J, Buck P, Sockett P, Aramini J, Pollari F, 2006. A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada. Int J Health Geogr 5, 17.

Sundell, J, 2004. On the history of indoor air quality and health. Indoor Air 14, 51-58.

Thacker SB, Berkelman RL, 1988. Public health surveillance in the United States. Epidemiol Rev 10, 164-190.

van den Hooven EH, de Kluizenaar Y, Pierik FH, Hofman A, van Ratingen SW, Zandveld PY, Mackenbach JP, Steegers EA, Miedema HM, Jaddoe VW, 2011. Air pollution, blood pressure, and the risk of hypertensive complications during pregnancy: the generation R study. Hypertension 57, 406-412.

van Donkelaar A, Martin RV, Brauer M, Kahn R, Levy R, Verduzco C, Villeneuve PJ, 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. Environ Health Perspect 118, 847.

Vineis P, Hoek G, Krzyzanowski M, Vigna-Taglianti F, Veglia F, Airoldi L, Autrup H, Dunning A, Garte S, Hainaut P, et al., 2006. Air pollution and risk of lung cancer in a prospective study in Europe. Int J Cancer 119, 169-174.

Wagner MM, Espino J, Tsui F, Gesteland P, Chapman W, Ivanov O, Moore A, Wong W, Dowling J, Hutman J, 2004. Syndrome and outbreak detection using chief-complaint data-experience of the real-time outbreak and disease surveillance project. MMWR Morb Mortal Wkly Rep 53, 28-31.

Wang Z, Liu Y, Hu M, Pan X, Shi J, Chen F, He K, Koutrakis P, Christiani DC, 2013. Acute health impacts of airborne particles estimated from satellite remote sensing. Environ Int 51, 150-159.

Weber SA, Engel-Cox JA, Hoff RM, Prados AI, Zhang H, 2010. An improved method for estimating surface fine particle concentrations using seasonally adjusted satellite aerosol optical depth. J Air Waste Manag Assoc 60, 574-585.

WHO, 2013. Asthma. Fact sheet N° 307. Geneva: World Health Organization. Available at: http://www.who.int/mediacentre/factsheets/fs307/en/ (accessed on August 2014).

WHO, 2014. Ambient (outdoor) air quality and health. Fact sheet N° 313. Geneva: World Health Organization. Available at: http://www.who.int/mediacentre/factsheets/fs313/en/ (accessed on August 2014).

Yiannakoulias N, Svenson LW, Schopflocher DP, 2009. An integrated framework for the geographic surveillance of chronic disease. Int J Health Geogr 8, 69.

Zanobetti A, Franklin M, Koutrakis P, Schwartz J, 2009. Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environ Health 8, 58.