Science of the Total Environment 905 (2023) 167095

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

A temporally and spatially explicit, data-driven estimation of airborne ragweed pollen concentrations across Europe

László Makra [a,*], István Matyasovszky [b,†], Gábor Tusnády [c], Lewis H. Ziska [d], Jeremy J. Hess [e], László G. Nyúl [f], Daniel S. Chapman [g], Luca Coviello [h], Andrea Gobbi [i], Giuseppe Jurman [i], Cesare Furlanello [j,ap], Mauro Brunato [k], Athanasios Damialis [l], Athanasios Charalampopoulos [l], Heinz Müller-Schärer [m], Norbert Schneider [a], Bence Szabó [a], Zoltán Sümeghy [a], Anna Páldy [n], Donát Magyar [n], Karl-Christian Bergmann [o], Áron József Deák [a], Edit Mikó [p], Michel Thibaudon [q], Gilles Oliver [q], Roberto Albertini [r], Maira Bonini [s], Branko Šikoparija [t], Predrag Radišić [t], Mirjana Mitrović Josipović [u], Regula Gehrig [v], Elena Severova [w], Valentina Shalaboda [x], Barbara Stjepanović [y], Nicoleta Ianovici [z], Uwe Berger [aa], Andreja Kofol Seliger [ab], Ondřej Rybníček [ac], Dorota Myszkowska [ad], Katarzyna Dąbrowska-Zapart [ae], Barbara Majkowska-Wojciechowska [af], Elzbieta Weryszko-Chmielewska [ag], Łukasz Grewling [ah], Piotr Rapiejko [ai], Malgorzata Malkiewicz [aj], Ingrida Šaulienė [ak], Olexander Prykhodo [al], Anna Maleeva [al], Victoria Rodinkova [am], Olena Palamarchuk [am], Jana Ščevková [an], James M. Bullock [ao]

[a] Institute of Economics and Rural Development, Faculty of Agriculture, University of Szeged, 6800 Hódmezővásárhely, Andrássy út 15, Hungary
[b] Department of Meteorology, Eötvös Loránd University, 1518 Budapest, P.O.B. 32, Hungary
[c] Alfréd Rényi Institute of Mathematics, 1364 Budapest, P.O.B 127, Hungary
[d] Mailman School of Public Health, Columbia University, New York, NY 10032, USA
[e] Department of Global Health, University of Washington, Seattle, WA 98105, USA
[f] Department of Image Processing and Computer Graphics, University of Szeged, 6701 Szeged, P.O.B. 652, Hungary
[g] Biological and Environmental Sciences, Faculty of Natural Sciences, University of Stirling, Stirling FK9 4LA, Scotland, UK
[h] University of Trento and Enogis s.r.l., Trento, Italy
[i] Bruno Kessler Foundation, Trento, Italy
[j] HK3 Lab, Rovereto, Italy
[k] Department of Information Engineering and Computer Science, University of Trento, Trento, Italy
[l] Terrestrial Ecology and Climate Change, Department of Ecology, School of Biology, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece
[m] Departement of Biology, Unit of Ecology and Evolution, University of Fribourg, CH-1700 Fribourg, Switzerland
[n] National Institute of Environmental Health, 1097 Budapest, Albert Flórián út 2-6, Hungary
[o] Institute of Allergology, Charité, 12203 Berlin, Hindenburgdamm 30, Germany
[p] Institute of Animal Science and Wildlife Management, Faculty of Agriculture, University of Szeged, 6800 Hódmezővásárhely, Andrássy út 15, Hungary
[q] Réseau National de Surveillance Aérobiologique, 11 chemin de la Creuzille, Le Plat du Pin, 696905 Brussieu, France
[r] Laboratory of Hygiene and Aerobiology, Department of Medicine and Surgery, University of Parma, U.O. Medicina Interna di Continuità, Azienda Ospedaliero-Universitaria di Parma, Via Gramsci 14, 43126 Parma, Italy

* Corresponding author at: Institute of Economics and Rural Development, Faculty of Agriculture, University of Szeged, Hungary, HU-6800 Hódmezővásárhely, Andrássy út 15, Hungary.

E-mail addresses: makralaszlo@szte.hu (L. Makra), matya@ludens.elte.hu (I. Matyasovszky), tusnady.gabor@renyi.mta.hu (G. Tusnády), lhz2103@cumc.columbia.edu (L.H. Ziska), jjhess@uw.edu (J.J. Hess), nyul@inf.u-szeged.hu (L.G. Nyúl), daniel.chapman@stir.ac.uk (D.S. Chapman), luca.coviello@unitn.it (L. Coviello), agobbi@fbk.eu (A. Gobbi), jurman@fbk.eu, giuseppe.jurman@fbk.eu (G. Jurman), cesare.furlanello@hk3lab.ai (C. Furlanello), mauro.brunato@unitn.it (M. Brunato), dthanos@bio.auth.gr (A. Damialis), athchara@bio.auth.gr (A. Charalampopoulos), heinz.mueller@unifr.ch (H. Müller-Schärer), paldy.anna@oki.antsz.hu (A. Páldy), aron@geo.u-szeged.hu (Á.J. Deák), miko.edit@szte.hu (E. Mikó), gilles.oliver@rnsa.fr (G. Oliver), roberto.albertini@unipr.it (R. Albertini), Mbonini@ats-milano.it (M. Bonini), sikoparijabranko@biosense.rs (B. Šikoparija), mirjana.mitrovic@sepa.gov.rs (M.M. Josipović), Regula.Gehrig@meteoswiss.ch (R. Gehrig), barbara.stjepanovic@stampar.hr (B. Stjepanović), nicoleta.ianovici@e-uvt.ro (N. Ianovici), uwe.berger@meduniwien.ac.at (U. Berger), andreja.kofol.seliger@nlzoh.si (A.K. Seliger), dorota.myszkowska@uj.edu.pl (D. Myszkowska), katarzyna.dabrowska-zapart@us.edu.pl (K. Dąbrowska-Zapart), bmw@csk.umed.lodz.pl (B. Majkowska-Wojciechowska), grewling@amu.edu.pl (Ł. Grewling), piotr@rapiejko.pl (P. Rapiejko), malgorzata.malkiewicz@uwr.edu.pl (M. Malkiewicz), ingrida.sauliene@sa.vu.lt (I. Šaulienė), rodinkova@vnmu.edu.ua (V. Rodinkova), jana.scevkova@uniba.sk (J. Ščevková), jmbul@ceh.ac.uk (J.M. Bullock).

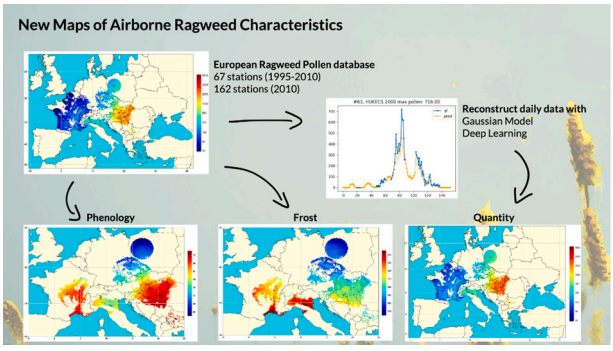
https://doi.org/10.1016/j.scitotenv.2023.167095
Received 11 July 2023; Received in revised form 29 August 2023; Accepted 13 September 2023
Available online 23 September 2023

[s] Department of Hygiene and Health Prevention, ATS (Agency for Health Protection of Metropolitan Area of Milan), Hygiene and Public Health Service, via Spagliardi 19, Parabiago, 20015 Milan, Italy
[t] BioSense Institute - Research Institute for Information Technologies in Biosystems, University of Novi Sad, Dr. Zorana Đinđića 1, 21000 Novi Sad, Serbia
[u] Ministry of Environmental Protection, Environmental Protection Agency, 11000 Belgrade, Ruže Jovanoviüa 27a, Serbia
[v] Federal Department of Home Affairs FDHA, Federal Office of Meteorology and Climatology MeteoSwiss, Operation Center 1, P.O. Box, CH-8058, Zurich-Airport, Switzerland
[w] Lomonosov Moscow State University, Biological Faculty, 1-12 Leninskie Gory, 119991 Moscow, Russia
[x] State Institution (Scientific and Practical Center (SPC) of the State Forensic Examination Committee of the Republic of Belarus, Akademicheskaya Str. 27, 220072 Minsk, Belarus
[y] Teaching Institut of Public Health "Dr Andrija Štampar", 10000 Zagreb, Croatia
[z] West University of Timişoara, Blvd. V. Parvan 4, 300223 Timişoara, Romania
[aa] Department of Oto-Rhino-Laryngology, HNO Klinik, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria
[ab] National Laboratory of Health, Environment and Food, Center for Environment and Health, Department for Air, Noise, Environmental Impact Assessment and Aerobiology, Grablovičeva ulica 44, 1000 Ljubljana, Slovenia
[ac] Pediatric Department, University Hospital and Masaryk University, Brno, Jihlavská 20, 00 Brno, Czech Republic
[ad] Jagiellonian University, Medical College, Department of Clinical and Environmental Allergology, 31-531 Kraków, ul. Kopernika 15A, Poland
[ae] Institute of Earth Sciences, Faculty of Natural Sciences, University of Silesia in Katowice, Bedzinska 60, 41-200 Sosnowiec, Poland
[af] Aeroallergen Monitoring Centre "AMoC", Department of Immunology, Rheumatology and Allergy, Medical University of Lodz, Pomorska 251, 92-213 Lodz, Poland
[ag] Department of Botany, University of Life Sciences, 13 Akademicka Street, 20-950 Lublin, Poland
[ah] Laboratory of Aerobiology, Department of Systematic and Environmental Botany, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland
[ai] Allergen Research Center Ltd., Warsaw, Poland
[aj] Department of Palaeobotany, Institute of Geological Sciences, University of Wroclaw, Poland
[ak] Vilnius University, Siauliai Academy, Vytauto 84, LT-76352, Siauliai, Lithuania
[al] Department of Medical Biology, Zaporizhia State Medical University, 69035 Zaporizhia, Ukraine
[am] National Pirogov Memorial Medical University, Vinnytsya, 56 Pirogov street, Vinnytsia 21018, Ukraine
[an] Department of Botany, Comenius University, Šafárikovo námestie 6, 81806 Bratislava, Slovakia
[ao] UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford OX10 8BB, UK
[ap] Orobix Life Srl, Bergamo-Rovereto, Italy

HIGHLIGHTS

- Temporal maps for the distribution of ragweed pollen for Europe were developed.
- Missing daily pollen data were restored with the Gaussian method and deep learning.
- A web page was created to upload newly measured or restored daily pollen data.
- These are the first maps of pollen phenology, -quantity and frost events for Europe.
- DL improves the accuracy more than GM, although it requires the use of weather data.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Anastasia Paschalidou

Keywords:
Ambrosia
Aerobiology
Flowering phenology
Artificial intelligence
Climate change
Data reconstruction
Health risk
Invasive species

ABSTRACT

Ongoing and future climate change driven expansion of aeroallergen-producing plant species comprise a major human health problem across Europe and elsewhere. There is an urgent need to produce accurate, temporally dynamic maps at the continental level, especially in the context of climate uncertainty. This study aimed to restore missing daily ragweed pollen data sets for Europe, to produce phenological maps of ragweed pollen, resulting in the most complete and detailed high-resolution ragweed pollen concentration maps to date. To achieve this, we have developed two statistical procedures, a Gaussian method (GM) and deep learning (DL) for restoring missing daily ragweed pollen data sets, based on the plant's reproductive and growth (phenological, pollen production and frost-related) characteristics. DL model performances were consistently better for estimating seasonal pollen integrals than those of the GM approach. These are the first published modelled maps using altitude correction and flowering phenology to recover missing pollen information. We created a web page (http://euragweedpollen.gmf.u-szeged.hu/), including daily ragweed pollen concentration data sets of the stations examined and their restored daily data, allowing one to upload newly measured or recovered daily data. Generation of these maps provides a means to track pollen impacts in the context of climatic shifts, identify geographical regions with high pollen exposure, determine areas of future vulnerability, apply spatially-explicit mitigation measures and prioritize management interventions.

[†] He passed away in 2015.

1. Introduction

Common ragweed *(Ambrosia artemisiifolia)* (Asteraceae family) is an annual herbaceous plant, native to North America, which has been introduced and subsequently naturalized in a large part of Europe (Bullock et al., 2010; Makra et al., 2015; Montagnani et al., 2017, 2023), Asia (Chen et al., 2007a, 2007b; Singh and Mathur, 2021) and Australia (Bass et al., 2000; Beggs, 2018) following its introduction to many places in the world.

The expansion of *Ambrosia* into Europe began after the First World War (Comtois, 1998). Seeds of different *Ambrosia* species were transferred to Europe from America by purple clover seed shipments and grain imports. Its expansion started probably from the European ports: from Rijeka toward Croatia and Transdanubia (a region of Hungary), from Trieste and Genoa toward northern Italy, and from Marseilles toward the Rhône valley (Makra et al., 2005; Kazinczi et al., 2008).

The most frequently occurring *Ambrosia* species in Europe are *A. artemisiifolia, A. trifida, A. psilostachya,* and *A. tenuifolia*. However, the most widespread of them is *A. artemisiifolia*. Its estimated distribution rate vs. other ragweed species in Europe is approx. 90:10, while in Hungary approx. 95:5 in favour of *A. artemisiifolia* (Bartha et al., 2019, 2022; Makra, 2022).

In Europe, the most important habitat areas of common ragweed in decreasing order of the measured seasonal pollen integral are as follows: (I) the south-western part of European Russia, (II) the southern and eastern parts of Ukraine, (III) the Pannonian Plain in central Europe, (IV) the Rhône-Alpes region in France, furthermore (V) the Po River valley in Italy (Makra et al., 2015).

There are substantial differences in the seasonal pollen integral per station. Between the two extreme years, five times or even higher differences may be detected at the individual stations (Makra et al., 2005). On the days with the maximum pollen concentration, a difference of one order of magnitude can be detected, while on the years with the maximum annual total pollen concentration, a difference of even two orders of magnitude can be shown between the ragweed pollen concentrations of the stations in the Pannonian Basin, which provides the climatic optimum for the life processes of common ragweed, and the stations of the peripheral areas (Makra et al., 2005).

Ragweed pollen grains can reach a height of 1 km due to the strong updrafts of the summer months, where typical wind speeds reach 40 km/h. Though their settling rate is low (approx. 1 cm/s) (Sofiev et al., 2006), strong updrafts can keep pollen grains at that height for longer periods of time. Based on this, ragweed pollen grains can travel hundreds of kilometres per day. However, every airborne pollen grain can osmotically rupture to produce sub-pollen particles (SPP) that can be long-distance transported much further away than the airborne ragweed pollen itself (Stone et al., 2021). Some examples of the ratio of the long-range transported pollen in the total annual pollen amounts are as follows. For the Galápagos Islands, 1000 km from South America, 5 % of the seasonal pollen integral (van der Knaap et al., 2012); for southern Greenland, 11 % (Rousseau et al., 2005, 2006); for Szeged, Hungary 7.5 % (Makra et al., 2016) and for eastern Germany 20 % (Zink et al., 2012). In addition, Prank et al. (2013) found that high-concentration areas are substantially more extensive than the heavily-infested territories. This is indirect proof of the long-distance transport of airborne ragweed pollen in Europe.

Among the 42 species of the ragweed *(Ambrosia)* genus, common ragweed (*A. artemisiifolia* L.), an invasive alien species in Europe, is among the most dangerous for public health. Based on clinical investigations, common ragweed pollen is the most serious and persistent cause of allergy-associated respiratory diseases (Schaffner et al., 2020; Bonini et al., 2022; Montagnani et al., 2023). Damialis et al. (2021) found that simultaneous exposure to SARS-CoV-2 (via other infected human carriers) and airborne pollen may, under 'favourable' weather conditions, also promote viral infection. Currently, 23.2 million people (3.13 %) are sensitised to airborne ragweed pollen in Europe (Schaffner et al., 2020). The prevalence of allergic respiratory conditions has increased over the last three decades, especially in industrialised countries (Damialis et al., 2019; Ziska et al., 2019; Rauer et al., 2020). Economic costs per year, associated with ragweed and allergic rhinitis, can run into billions of dollars (AAAAI, 2006; ASCIA, 2007; Bullock et al., 2010; Richter et al., 2013; Schaffner et al., 2020; Hillerich et al., 2023).

Recent anthropogenic warming is associated with changes in the phenological and quantitative parameters of pollen dispersion of different plant species (Recio et al., 2010; Rodríguez-Rajo et al., 2011; Hess, 2019; Anderegg et al., 2021). Warming might already be contributing to extended pollen season duration and increased pollen load for multiple aero-allergenic taxa across the northern hemisphere (Ziska et al., 2019; Xian et al., 2023), which can potentially be attributed to increases in pollen productivity and/or higher reproductive capacity in general, or expansion of plant ranges (Damialis et al., 2019). Recent models, in accordance with the warming climate, predict a northward and eastward shift of common ragweed habitats in central and northern Europe in the coming decades (Cunze et al., 2013, ecological niche modelling (ENM); Chapman et al., 2014, a phenology model; Storkey et al., 2014, a process-based model; Chapman et al., 2016, a simulation model to understand and describe plant invasion at a continental scale; Leiblein-Wild et al., 2016, a physiological model; Chapman et al., 2017, a forward mechanistic species distribution model; Lake et al., 2017, a process-based model, and a simulation model; Rasmussen et al., 2017, species distribution models). Anderegg et al. (2021) found that the contribution of human forcing on the climate system – due to the accumulation of anthropogenic greenhouse gases, especially CO_2 (Wayne et al., 2002; Ziska and Beggs, 2012) – accounted for ~50 % of the trend in lengthening pollen seasons and ~8 % of the trend in increasing pollen concentrations. Based on these climatic trends, sensitisation to ragweed is projected to more than double in Europe, from 33 million people (1986–2005) to 77 million by 2041–2060 (Lake et al., 2018).

The changing frequency, strength, and duration of extreme atmospheric events such as heat waves, droughts, floods, and thunderstorms may also be associated with the changing climate (D'Amato and D'Amato, 2023). Due to heat waves and droughts pollen production decreases (Deák et al., 2013). At the same time, this decrease in pollen production is far exceeded by the surplus of pollen, which is caused by the extension of the pollen season due to global warming and the expansion of the habitats of pollen-producing plants. Nevertheless, certain pollen-producing plants can be sensitive not only to drought but also to floods by changing their pollen properties (Yamburov et al., 2014). On the other hand, thunderstorms during the pollen season can increase the intensity of asthma attacks in pollinosis patients (D'Amato et al., 2007).

Further models for estimating airborne ragweed pollen emission and dispersion are as follows. Zink et al. (2012) used the COSMO-ART meteorological and dispersion model and combined it with a manually harmonized inventory of ragweed habitat in Germany, Austria, Czechia, and Hungary. However, the model was applied only to a single short episode. Prank et al. (2013) presented a new model for ragweed pollen release and dispersion over the European continent, and reported the first estimates of the European-wide ragweed pollen load. They used the SILAM (System for Integrated modeLling of Atmospheric coMposition) model (http://silam.fmi.fi (Sofiev et al., 2008, Sofiev et al., 2013), which is a chemical transport model. Liu et al. (2015, 2016) implemented a pollen emission and transport module in the Regional Climate Model (RegCM4) using the Community Land Model (CLM4.5) for calculating pollen emissions for Europe. Menut et al. (2021) used the CHIMERE chemistry-transport model to improve the daily release of ragweed pollen emission for nine stations in Hungary, Croatia and France.

Recent maps of airborne ragweed pollen concentrations for Europe presented by Smith et al. (2013) taken from Skjøth et al. (2012), were based on mean annual *Ambrosia* pollen levels from 368 stations using the

European Aeroallergen Network (EAN) Pollen Database (https://ean.polleninfo.eu/ean). Another map of airborne ragweed pollen concentrations for Europe was published by Prank et al. (2013), who used data for the period 2005–2011. Liu et al. (2015, 2016) used data for the period 2000–2010, simulating the production and dispersion of airborne ragweed pollen for Europe. Hamaoui-Laguel et al. (2015) used the chemistry transport model CHIMERE and the RegCM4 regional climate model for mapping historical average pollen concentrations based on annual pollen concentration data of 51 sites across Europe. The mean annual ragweed pollen concentration map for Europe, prepared by Schaffner et al. (2020), used 5–9-year long pollen concentration data sets for 296 stations.

Concerning airborne pollen concentrations of common ragweed, there are no accurate maps representing Europe in toto [see airborne ragweed pollen maps of the European Pollen Information Service (EPI) (http://www.polleninfo.org), the European Food Safety Authority (EFSA, 2010, http://www.efsa.europa.eu/fr/scdocs/doc/1566.pdf), the European Aeroallergen Network (EAN) (https://ean.polleninfo.eu/Ean), Buttenschøn et al., 2010, and Mányoki et al., 2011]. Existing maps: (I) are low resolution, due to the small number of aerobiological stations and significant gaps in data sets, (II) show only regional distributions of relative pollen concentrations, (III) are not particularly sensitive for areas with high pollen levels (Csépe et al., 2019), and (IV) do not eliminate the effect of elevation.

Reliable ragweed pollen concentration maps for North America, the native area of ragweed, are also missing from the literature. Only predictions on the start and length of the ragweed pollen season have been reported so far (Ziska et al., 2011; Zhang et al., 2015). One of the reasons for this may be that the pollen measurement was not done with a uniform sampling technique, but partly with Rotorod and partly with Burkard sampling equipment (oral communication: Lewis H. Ziska).

To date, these efforts have resulted in different estimates for representing ragweed distribution and pollen production for Europe. However, actual in situ quantification of ragweed pollen concentration or associated temperature characteristics related to ragweed growth and temporal pollen production (e.g. number of frost-free days) are currently unavailable. In addition, maps of phenological characteristics of the ragweed pollen season, that can affect its impacts on public health, such as the start, end and duration of the season, are also missing.

Consequently, there is a need to generate a biogeographical assessment of ragweed pollen concentrations for Europe using in situ generated data. Such an assessment could be used to generate a number of features including: (i) mean annual total ragweed pollen concentrations for a given time interval, (ii) mean first and last day exposure for a given season; (iii) mean duration of the ragweed pollen season, (iv) mean maximum daily ragweed pollen concentrations, (v) day of the year in which the maximum pollen concentrations occur; (vi) temperature related data on first and last frost and extent of the frost-free period. Overall, generation of these maps can identify geographical regions with high pollen exposure risk, determine areas of future vulnerability and, provide a means to distribute resources to manage and/or prevent ragweed introduction and spread.

The primary objective of this work was to produce the most reliable quantity- and phenology-related ragweed pollen maps for Europe. To achieve this, we set ourselves the following tasks: (a) restoring missing daily pollen concentration data for 67 stations for the years 1995–2010 and for 162 stations for the year 2010, by using two procedures (Gaussian method and deep learning), (b) comparing the accuracy of the two methods, (c) analyzing whether the data restoration with these two methods increased the accuracy of the ragweed pollen concentration maps at a statistically significant level compared to those prepared with raw data sets, and (d) applying an easy-to-use method to allow others to restore missing daily pollen concentration data with as much accuracy as possible.

In the following, we present the collected data and restoring methodologies and map generation techniques (Section 2), we show the results obtained in restoring missing pollen concentration data and the generated maps (Section 3), we discuss the relevance of the obtained results (Section 4), and present the conclusions and future research lines. To help readers follow the study, the overall methodology is also reported in a flow chart depicting the various stages of the analysis (Appendix, Fig. A8).

2. Materials and methods

2.1. Locations and selected years

In our initial approach, we used pollen concentration data of common ragweed from 625 aerobiological stations in Europe [https://ean.polleninfo.eu/Ean/ (European Aeroallergen Network Pollen Database)]. However, many of these stations have limited or sporadic data, may be absent from key geographical regions, or have ceased operations altogether. Consequently, significant temporal and spatial gaps for determining pollen characteristics are evident. For instance, most stations have been operating since the 1980s in the western part of Europe, while all aerobiological stations in Ukraine and the south-western part of Russia came on line circa 2010. Therefore, in order to prepare a satisfactory pollen portrait for Europe trade-offs among data coverage, number of stations and length of the period are evident. To address this issue, we decided to select a common study period covering the flowering and pollen-producing period of *A. artemisiifolia* (July 15 – October 15) and used only stations for which not more than 25 % of the years 1995–2010 (i.e. at most 4 years) had data coverage <40 %. This criterion is subjective, but by clarifying it, our goal was to include as many of the stations with incomplete databases as possible in the study, even if we have no restrictions regarding the temporal distribution of the data gaps per station according to the pollen season. As a result, only 67 out of the above 625 aerobiological stations were selected for a 16-year period (1995–2010). The year 2010 had the most complete data sets over this period, and using a similar approach for the 1995–2010 period, we selected 162 stations for this year.

Finally, we generated two data sets from the above database (Fig. 1, 1b; Table 1; European Aeroallergen Network Pollen Database, https://ean.polleninfo.eu/Ean). Namely, (a) 16-year (1995–2010) annual total ragweed pollen counts based on daily data for 67 stations and (b) annual total ragweed pollen counts for the year 2010 using daily data from 162 stations.

2.2. Study periods and variables in the selected years

Ambrosia pollen is prevalent over the late summer - early autumn period, although it can continue producing pollen until the first frost. Therefore, when calculating mean seasonal pollen integrals (67 stations, 1995–2010) and seasonal pollen integrals (162 stations, 2010), we used a uniform study period (July 15 – October 15) for each station involved in the analysis. The rationale for selecting this period is that it covers the pollen season of the stations in the Pannonian Plain within the Carpathian Basin (the south-eastern part of central Europe) and that of the other main distribution centres (Rhône-Alpes region in France and western Lombardy in Italy) (Skjøth et al., 2019). With this choice, we intended to focus on the most important centres of common ragweed distribution in relation to generating pollen concentration maps.

All pollen data from every aerobiological station were measured applying the same methodology, using the Burkard sampling equipment (Crisp et al., 2013) (Appendix, Sections A1 and A2).

Daily weather data were downloaded from the European Climate Assessment and Dataset website (https://www.ecad.eu/dailydata/predefinedseries.php#, Blended ECA dataset, Cornes et al., 2018). The website provides daily meteorological data for a large number of stations over Europe.

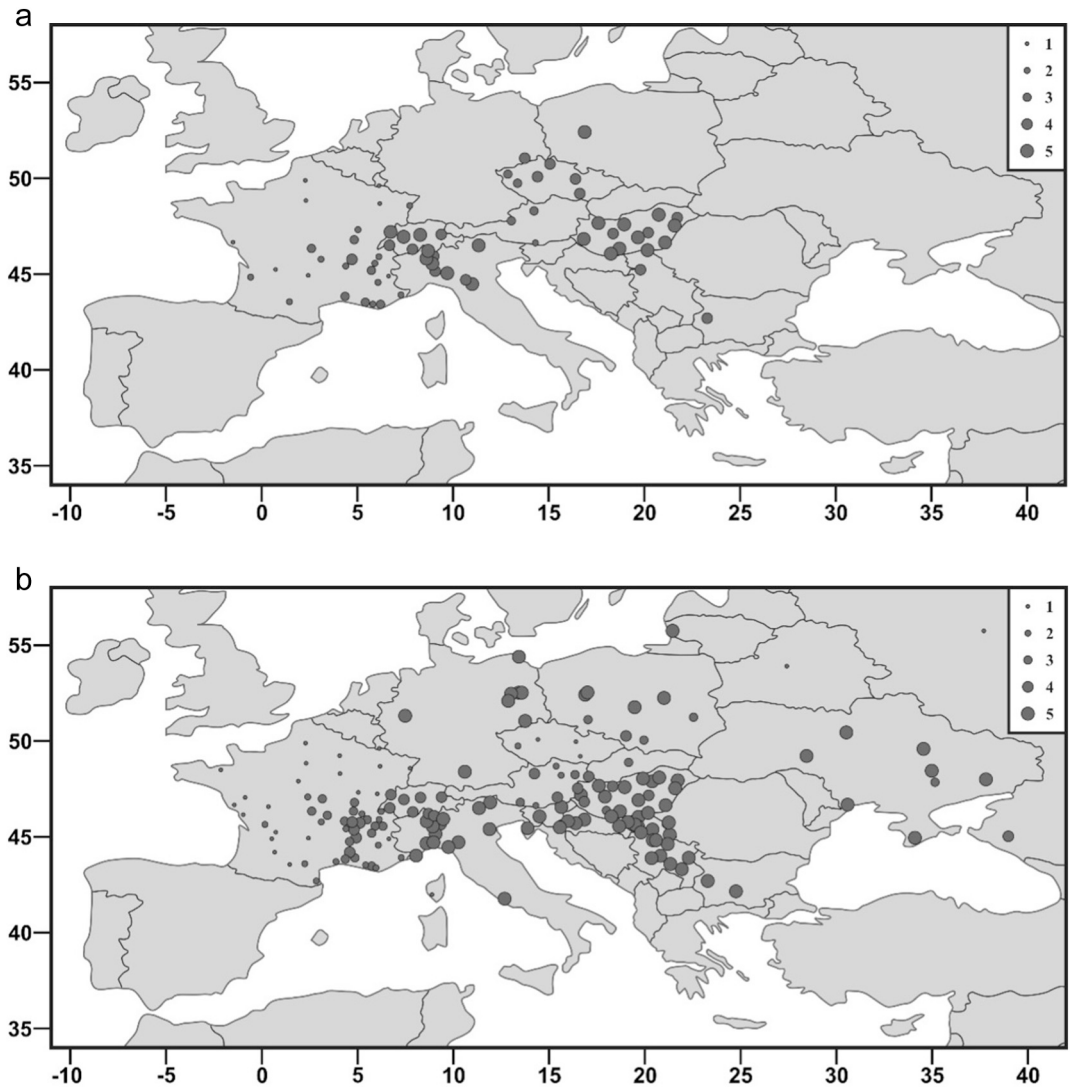


Fig. 1. a. Geographical location of the 67 aerobiological stations used in the study comprising 16-year (1995–2010) mean annual total ragweed pollen concentrations. Locations of the aerobiological stations are indicated by black dots. Horizontal axis: longitude, angle in degrees; vertical axis: latitude, angle in degrees. The data density (x, %) of the stations is proportional to the size of the black dots (see Table 1). Legends: 1: $0 < x \leq 20$; 2: $20 < x \leq 40$; 3: $40 < x \leq 60$; 4: $60 < x \leq 80$; 5: $80 < x \leq 100$.

b. Geographical location of the 162 aerobiological stations used in the study comprising the total annual ragweed pollen concentrations for the year 2010. Locations of the aerobiological stations are indicated by black dots. Horizontal axis: longitude, angle in degrees; vertical axis: latitude, angle in degrees. The data density (x, %) of the stations is proportional to the size of the black dots (see Table 1). Legends: 1: $0 < x \leq 20$; 2: $20 < x \leq 40$; 3: $40 < x \leq 60$; 4: $60 < x \leq 80$; 5: $80 < x \leq 100$.

2.3. Restoration of missing ragweed pollen data

2.3.1. Data restoration by the Gaussian method

Pollen data were unavailable for many stations and on several days for the study period (Fig. 1, 1b; Table 1). Due to the relatively few numbers of aerobiological stations with full ragweed pollen data sets, we intended to keep stations with incomplete data but to use what was available to estimate their missing data. The Gaussian curve was fitted to the seasonal distribution of the daily ragweed pollen concentrations at the given station, and the missing daily pollen concentration values were estimated from the fitted daily values for the Gaussian curve (Kasprzyk and Walanus, 2014).

In order to estimate the mean of the seasonal pollen integrals (MATPCs), individual ATPCs were estimated separately for every station and every year. Where enough data were available in a suitable temporal distribution, this was done in one step. Otherwise, an additional step, spatial interpolation (SI) was used (Appendix, Section A3). Then, averaging ATPCs over years for each station gave mean ATPC (MATPC) values.

In the first step, a Gaussian curve was fitted to the daily pollen concentration data for every year and every station separately (Appendix, Section A4).

Evidently, the fitted curve does not always provide a good estimate for actual daily concentrations, but we are interested only in the ATPC. An experiment was performed for stations and years where data coverage was 100 %, i.e., pollen concentration data were available for the entire pollen seasons. Typically, the relative difference between C and $\widehat{C}$ (C denotes measured ATPC and $\widehat{C}$ is the estimate of the expected ATPC if data on all days were available) was just slightly above 5 %. Fig. 2 shows two examples as the worst estimations, namely for a station (Nyíregyháza, Hungary) (Fig. 2a) having very high daily concentrations and for another station (Dresden, Germany) (Fig. 2b) having very low daily concentrations. The measured ATPC in 2010 is 14,223 and 100 pollen grains $\cdot$ m^{-3} at Nyíregyháza and Dresden respectively, and the relative difference $(C - \widehat{C})/C$ is −7.7 % and 10 % for these two stations. In a further experiment we cut portions of data from the data sets which

Table 1
Number of the stations with data coverage x (%) for the selected periods[1].

*Data density, x (%)	*Graphic display of the data density		Period			
			Pollen season, 1995-2010		Pollen season, 2010	
	Serial no.	Size	Number of stations	%	Number of stations	%
0<x≤ 20	1		8	11.9	27	16.7
20<x≤ 40	2		13	19.4	17	10.5
40<x≤ 60	3		9	13.4	22	13.6
60<x≤ 80	4		17	25.4	22	13.6
80<x≤100	5		20	29.9	74	45.7
Total	–		67	100.0	162	100.0

*Data density of the stations is proportional to the size of the black dots (see Fig. 1a and Fig. 1b);
[1]Note: the table concerns the station-specific pollen seasons but not the selected study period, July 15 – October 15. Since the study period does not cover the pollen season for many stations, the station selection criteria defined in Section 2.1, paragraph 1, does not contradict with the result of Table 1, according to which >25 % of the years (actually, 31.3 % of the years) 1995–2010 has data coverage <40 %.

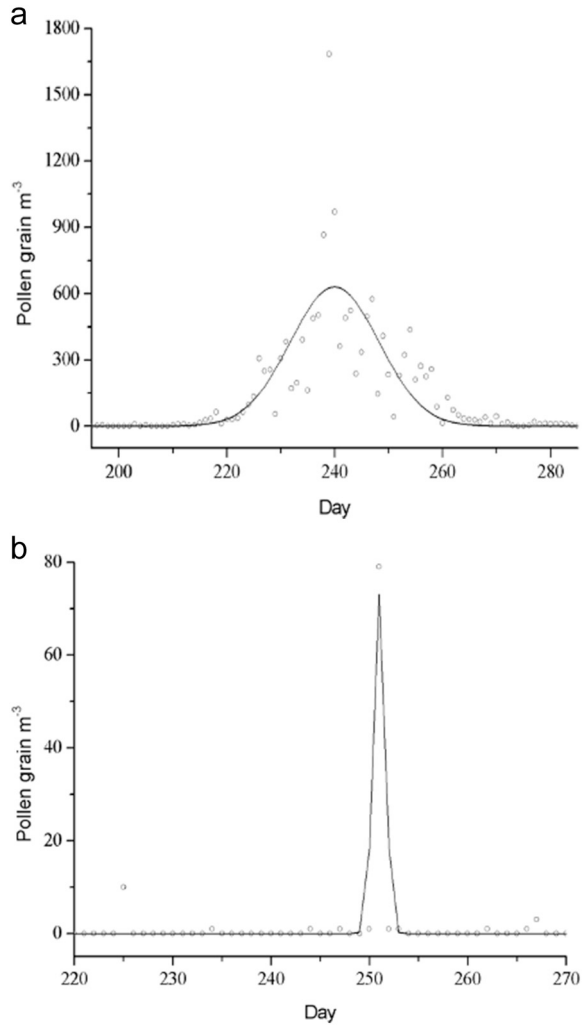


Fig. 2. Daily ragweed pollen concentrations at (a) Nyíregyháza (Hungary) and (b) Dresden (Germany) in 2010 and their approximation by a Gaussian Curve.

had 100 % data coverage thus creating data sets with missing values. Our finding is that the relative error $(C-\widehat{C})/C$ increases slowly as the portion of omitted data increases when missing values are not clustered but are distributed nearly uniformly in time. In contrast, the accuracy of $\widehat{C}$ substantially decreases when missing values are clustered, i.e. they are concentrated on one or more subperiods of the pollen season.

Unfortunately, data sets for a number of stations and years have clustered missing data structures and hence the estimated values $\widehat{C}$ for these are likely highly inaccurate. These estimated ATPCs were omitted and substituted with values obtained by interpolating ATPCs from other stations where reliable ATPCs were available (see Section 2.3.1.1). The decision of whether a $\widehat{C}$ is accurate enough or not is based on subjective considerations because an objective decision requires a measure of $C-\widehat{C}$, but C is not known. Specifically, the location parameter m relating to the date of the annual peak pollen concentration should be around the middle of the pollen season. An $\widehat{m}$ strongly deviating from the day of the middle of the pollen season shows that missing values are concentrated in subperiods of the pollen season with increasing or decreasing pollen concentrations. Additionally, the dispersion parameter σ relates to the length of the pollen season. It happens in a number of cases that $\widehat{C}$ is small, but with a large $\widehat{\sigma}$. This situation shows that missing values are concentrated in the subperiod with the highest daily pollen concentrations. Finally, we compared $\widehat{C}$ in a given station and year to ATPC values of other years at the same station and to the values of the neighbouring stations of the same year. When $\widehat{C}$ looked like an outlier ($\widehat{C}>2\widehat{\sigma}$), then $\widehat{C}$ was omitted.

A web page (http://euragweedpollen.gmf.u-szeged.hu/) was created, which includes the daily ragweed pollen concentration data sets, restored with the Gaussian method, for both the 67 selected stations (16-year data series, 1995–2010) and the 162 selected stations (annual data sets, 2010), with the possibility to upload newly measured or recovered daily data (Appendix, Section A5).

2.3.1.1. Interpolation. In the interpolation step, refined kriging techniques (Chiles and Delfiner, 1999; Oteros et al., 2019) cannot be used as they require the variogram of the underlying random field, which cannot be accurately estimated due to missing values. Therefore, our original intention was to use an inverse distance method. The procedure, however, produced poor results as the interpolation led to large over- or underestimations for areas of sharp spatial changes in concentrations. Hence, the interpolation was performed with the help of the geographical coordinates (latitude ϕ, longitude λ and the height above sea surface z) of stations using the Nadaraya-Watson estimate (Simonoff, 1996):

$$\widehat{y}_i=\frac{\sum_j y_j K\left(\frac{\phi_j-\phi_i}{h_\phi}\right)K\left(\frac{\lambda_j-\lambda_i}{h_\lambda}\right)K\left(\frac{z_j-z_i}{h_z}\right)}{\sum_j K\left(\frac{\phi_j-\phi_i}{h_\phi}\right)K\left(\frac{\lambda_j-\lambda_i}{h_\lambda}\right)K\left(\frac{z_j-z_i}{h_z}\right)}. \quad (2.1)$$

Here y_i denotes ATPC at the ith target station, and the summation runs over stations where ATPC is available from the first step. Note that the interpolated ATPC is a weighted average of ATPC values with weights depending on the distance between the geographical coordinates of the target station and other stations. The weights are generated by a Gaussian kernel $K(u)=exp\left(-(u/h)^2/2\right)\Big/\left(\sqrt{2\pi}h\right)$ and the bandwidth h controls the width of averaging. The choice of bandwidths h_ϕ, h_λ, h_z has an important role as small bandwidths provide a small bias with a big variance, while large bandwidths deliver a large bias with a small variance of the estimate Eq. (2.1). Optimal bandwidths resulting in the smallest mean squared error of the interpolation are obtained by cross-validation (Simonoff, 1996), i.e., by minimizing $\sum_i\left(y_i-\widehat{y}_i^i\right)^2$ with respect to h_ϕ, h_λ, h_z, where $\widehat{y}_i^i$ is obtained from Eq. (2.1) but omitting ith data from the summation. In fact, the procedure is a combination of three one-dimensional Gaussian kernels, namely one Gaussian kernel for each dimension (ϕ, λ and z), and the best bandwidths parameters (for each dimension) are estimated with cross-validation. This way the Gaussian kernels can be applied to variables with very different value ranges. The Gaussian kernel was developed

with the formula reported in Eq. (2.1).

Using this technique for the years when ATPC is available can provide a result for the ratio of the mean squared error of the interpolation to the variance. Hence, it is concluded that the spatial mean of the variance explained by this interpolation is 43.6 %.

2.3.2. Data restoration by artificial intelligence

Another way of reconstructing daily ragweed pollen concentration data is an AI approach employing deep neural networks. Specifically, the deep learning (DL) strategy is implemented through a Denoising Convolutional Auto-encoder (ConvAE). ConvAE is an auto-encoder composed of only convolutional layers that aim at modelling the temporal dependency in the time series and reconstructing the input data. Considering the low availability of data (i.e. the significant lack of data), Convolutional Neural Networks are easier to train compared to more complex recurrent architectures, e.g. Long-Short Term Memory components (LSTM) or Gated Recurrent Units (GRU) (Duan et al., 2016; Liu and Chen, 2017; Che et al., 2018; Navares and Aznarte, 2019; Zewdie et al., 2019a, 2019b; Zhang et al., 2019; Zhao et al., 2020).

ConvAE architecture takes as input daily pollen concentrations between June 1 and October 31 (in combination with additional daily weather data, namely daily values of maximum temperature, minimum temperature, mean temperature and precipitation). Other spatio-temporal variables are not used as input, with the objective of constructing a general model. The model is trained by randomly removing daily input concentrations and reconstructing the original time series. Models were trained with different data corruption schemes, input features, loss functions, and other hyper-parameters. The best configuration is selected with 5-fold cross-validation (Kohavi, 1995). A more detailed definition of the ConvAE model and the training procedure (Appendix, Fig. A1) is given in the appendix (Appendix, Section A6).

Results are reported for eight different model configurations: (1) with and without weather information, (2) with original or normalized loss function, and (3) with a perturbation scheme working on days or windows. Performances are reported for two evaluation settings: restoring single (i.e. *points*) and windows of consecutive (i.e. *windows*) missing days. Among the generated models, two (with and without weather channel) were selected to reconstruct daily pollen concentration data for each setting, based on cross-validation results. The two resulting models were then compared on the *test* and *test103* data sets to assess their behaviour. The *test* dataset was generated from the 67 stations, while the *test103* dataset was generated by selecting 103 stations that belong to the 162 stations dataset but are not included in the 67 stations dataset. The result of the best DL model was then compared to that of the restoration process implemented with the one-step and two-step ATPC restoration models.

The model performances are assessed through four metrics: Root Mean Squared Error *(aRMSE)*, the average of Normalized Root Mean Squared Error *(aNRMSE)*, the coefficient of determination *(R^2)* and the Pearson Correlation Coefficient (*PCC*). While *PCC* and R^2 are standard measures computed on the whole set of the reconstructed values, *aRMSE* and *aNRMSE* are computed by averaging *RMSE* and *NRMSE* values obtained on each reconstructed series (Mentaschi et al., 2013).

2.4. Cartographical background

The maps are shown using parts of the European grid. ETRS89/LAEA Europe is a projected coordinate reference system (CRS) suitable for use in Europe (Annoni et al., 2003). ETRS89/LAEA Europe uses the ETRS89 geographic 2D CRS as its base CRS and the Europe Equal Area 2001 (Lambert Azimuthal Equal Area) as its projection.

The borders of the study area were determined on the basis of the geographical positions of the pollen-measuring stations having sufficient pollen concentration data. Geographical coordinate corners for the maps were set to cover a region with ample margin beyond the pollen measuring stations of extreme coordinates. The reason is that in this way the given airborne ragweed pollen concentrations and ragweed pollen-related characteristics can be better described in the neighbourhood of the stations with extreme positions as well.

For the Europe-scale maps the map corners were set to cover most of Europe since, in addition to the many stations in the western and central part of the continent, we have data from stations in Russia and Scandinavian countries, as well. For these Europe-scale maps of mean annual total (1995–2010) (Fig. 3, 3*) and annual total (2010) (Appendix, Figs. A10, A10*) ragweed pollen concentrations a spherical trapezoid with 24° difference of latitudes (34°N-58°N) and 53° difference of longitudes (11°W-42°E) was selected.

The maps of the phenological characteristics (Fig. 4), peak pollen-related characteristics (Fig. 5) and frost-related characteristics (Appendix, Fig. A11) are based on the data from 67 stations only, located mostly in eastern and central Europe. The geographical coordinate corners of these maps were inclusive of those parts of Europe where data were available. This allowed better visualization for pollen relevancy of the maps. Using the same projection as detailed above, a spherical trapezoid with 15° difference in latitudes (40°N-55°N) and 32° difference of longitudes (5°W-27°E) was selected.

The Europe-scale pollen concentration maps were constructed using the interpolated pollen data from 67 and 162 stations, respectively. The maps of ragweed pollen-related characteristics are based on the data from 67 stations. Spatial predictions were made using latitude, longitude, as well as elevation information. Elevation information was available for the measuring stations. Elevation data for the interpolation sites were extracted from the ETOPO1 Global Relief Model (Amante and Eakins, 2009). ETOPO1 is a 1 arc-minute global relief model of Earth's surface that integrates land topography and ocean bathymetry. It was built from numerous global and regional data sets.

All interpolation was executed and the maps were created using the R software (version 3.1.2) (R Core Team, 2014). The maps were generated in two stages. First, using the scattered measured data points a high-resolution raster map was computed using Bayesian Gaussian Process Modelling as implemented in the spTimer package (Bakar and Sahu, 2014). The model was fitted by the spt.Gibbs function using the package default settings ("GP" model, "geodetic:km" distance method). The raster resolution was set to 15 pixels per geographical degree in both latitude and longitude dimensions, which corresponds to a spatial resolution of about 4 arcminutes. Because the predictions were made using a stochastic modelling process, the resulting pixels show the predicted trends with a stochastic pattern. To facilitate a better interpretation of the results, the predicted values were smoothed using a 3×3 averaging filter. The plots were generated using the ggplot package (Kahle and Wickham, 2013).

For the pollen concentration maps, the measured data were mapped with the generalized log transform before interpolation and the interpolated results were mapped back to the original scale using the inverse log transform. This is a commonly used technique to ensure that the interpolated values fall within a sensible range (in our case only non-negative values can be interpreted).

Standard rectangular shaped raster maps were created. In addition, the region covered by our maps extended the minimal enclosing rectangle of the measuring stations to show some influential regions around the stations at extreme locations. It should be stressed that interpolation techniques can only produce values at locations that lie within the convex hull of the data points with known values. For locations outside the convex hull, only extrapolation can be applied (using some fitting models). Extrapolation is less reliable than interpolation and the error increases with increasing geographical distance from the data points. Therefore, although visually appealing, some maps may show obviously false information, especially in regions where extrapolation was used. To address this problem, the maps were modified in several respects, as follows. (i) We masked out the regions further than 150 km away from any of the measuring stations (Skjøth et al., 2012; Smith et al., 2013). The neighbouring circles can be observed around stations further than

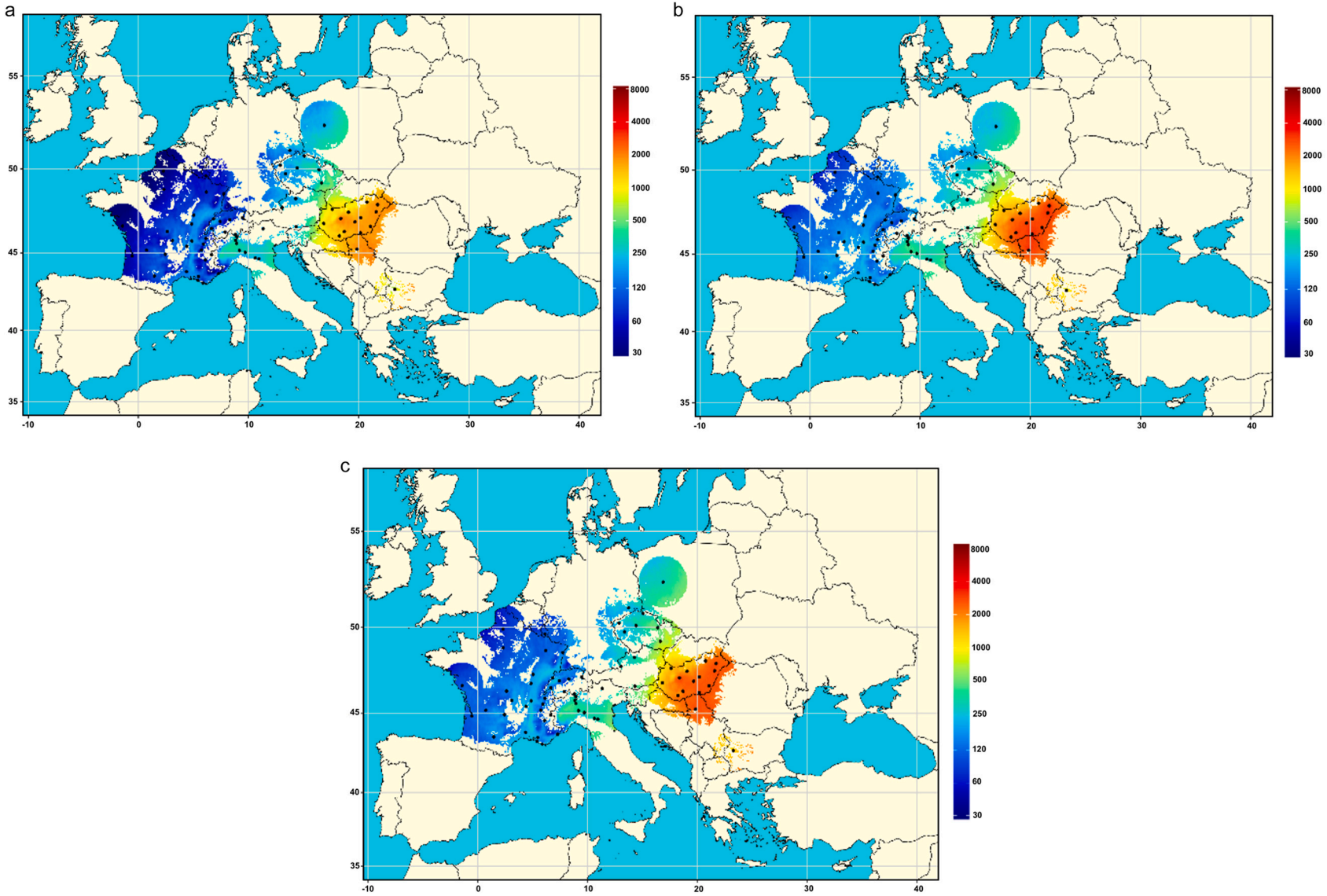


Fig. 3. 16-year mean annual total ragweed pollen concentrations for Europe using a *base-10* logarithmic colour scale, with the geographical location of the 67 aerobiological stations with (a) raw data sets, (b) restored (GM) data sets, and (c) restored(DL) data sets, July 15 – October 15, 1995–2010, (pollen grains · m^{-3} of air). The maps (1) include all stations, (2) the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping, and (3) the maximum distance is 150 km to each station. Horizontal axis: longitude, angle in degrees; vertical axis: latitude, angle in degrees. Locations of the aerobiological stations are indicated by black dots. **Restored(GM) data sets:** restoration of the missing daily data occurred on the assumption that the seasonal distribution of the daily pollen concentrations is normal (Gaussian Method) (see Section 2.3.1). **Restored(DL) data sets:** restoration of the missing daily data occurred by using convolutional auto-encoder models with weather data[1] (ConvAE + Weather) (see Section 2.3.2). Weather data[1]: daily values of maximum temperature, minimum temperature, mean temperature and precipitation.

Fig. 3*. 16-year mean annual total ragweed pollen concentrations for Europe on *base-10* logarithmic colour scale, with the geographical locations of the aerobiological stations, using the (a) raw data sets, (b) restored (GM) data sets, and (c) restored(DL) data sets, July 15 – October 15, 1995–2010 (pollen grains · m^{-3} of air). The stations (1) with mean annual pollen concentration smaller than 100 pollen grains · m^{-3} of air, as well as (2) the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping, and (3) the maximum distance is 150 km to each station. Horizontal axis: longitude, angle in

degrees; vertical axis: latitude, angle in degrees. Locations of the aerobiological stations are indicated by black dots. **Restored(GM) data sets:** restoration of the missing daily data occurred on the assumption that the seasonal distribution of daily pollen concentrations is normal (Gaussian Method) (see Section 2.3.1). **Restored(DL) data sets:** restoration of the missing daily data occurred by using convolutional auto-encoder models with weather data[1] (ConvAE + Weather) (see Section 2.3.2). Weather data[1]: daily values of maximum temperature, minimum temperature, mean temperature and precipitation.

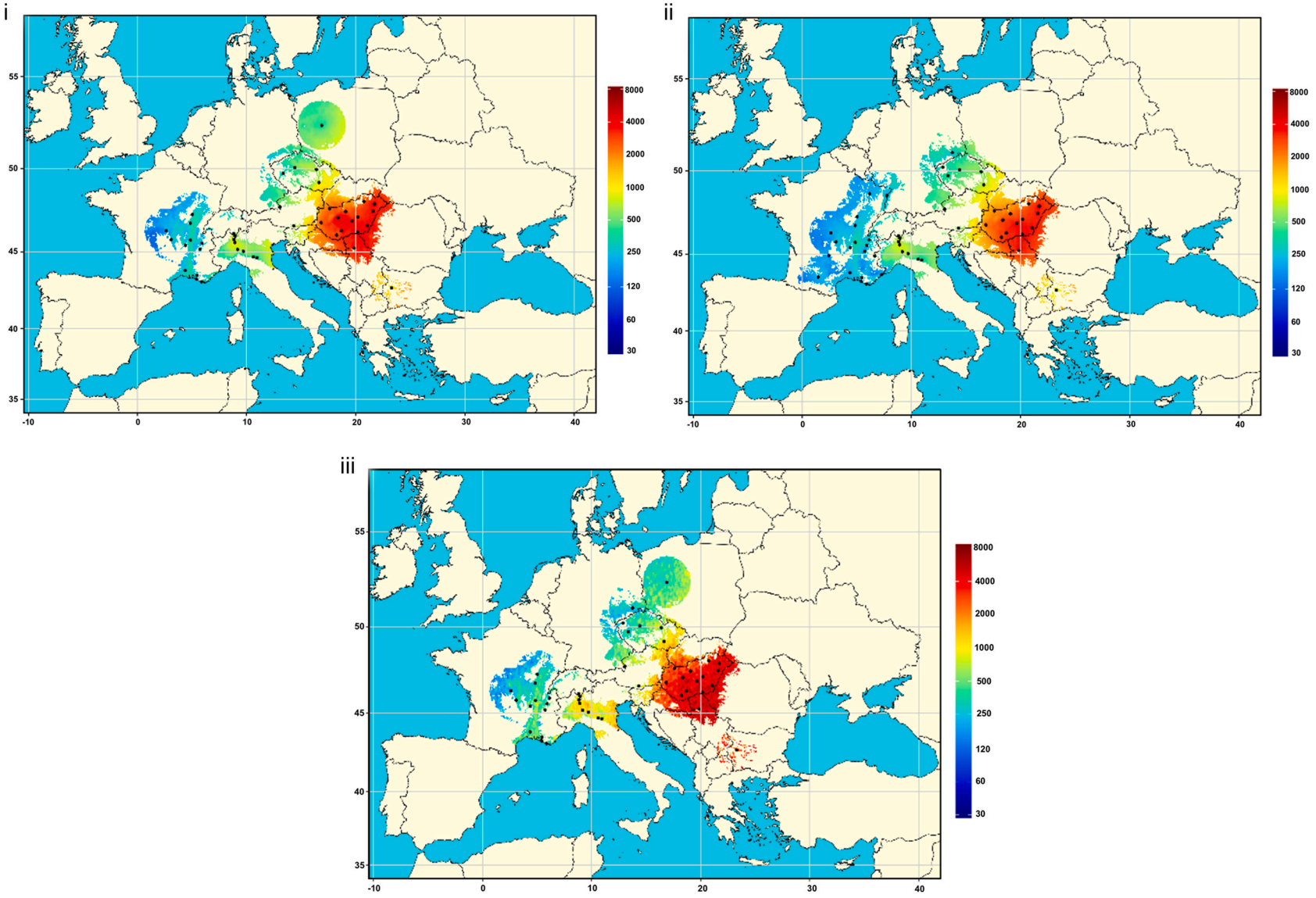


Fig. 3. (*continued*).

150 km from the other stations. These circles appear larger at higher latitudes than at lower latitudes due to the choice of the LAEA projection, though their areas are equal. (ii) We applied an altitudinal buffer, namely we masked out those areas within a 150 km radius of every aerobiological station that are beyond ±100 m from the altitude of the station. Removing these areas (e.g. the Alps and the Carpathians) (a) makes it possible to address uncertainties in predicting very late start and end dates of the pollen season and (b) allows a distinct colour scale to better differentiate eastern and western Europe. (iii) Studying seasonality requires a considerable observed pollen concentration. To this end, seasonal parameter maps were created using only those aerobiological stations where the mean annual pollen concentration reached at least 100 pollen grains · m^{-3}. This threshold also addresses the high altitude problem mentioned in (ii).

The statistical computations were performed with the following softwares: SPSS (version 15.0), MATLAB (version 7.7.0.471), Python (version 3.7), PyTorch (version 1.5), NumPy (version 1.18.4), and SciPy (version 1.4.1). The source code for training the models is as follows: https://gitlab.fbk.eu/dsip/dsip_dlresearch/ragweed-pollen-reconstruction-europe.

2.5. Evaluation of the map accuracy with the original and restored datasets

The final aim of the GM and DL models was to estimate a more accurate ATPC value that leads to an accuracy improvement in the generated maps. We compared the reconstruction abilities of the two models considering a set of 25 stations from the "*test103*" dataset. These stations had no missing data for the 2010 season, therefore their ATPC is considered correct. We removed different percentages of data (from 10 % to 80 %) from the time series of the 25 stations and asked the model to reconstruct the real ATPC. We evaluated two scenarios. First, we assessed model accuracy in reconstructing ATPC values with R^2, *RMSE* and Mean Absolute Error *(MAE)*. Second, we compared the Mean Absolute Percentage Error *(MAPE)* between the maps generated with real and perturbed data, and the *MAPE* between the maps generated with real and reconstructed data (Appendix, Section A7: A7.1, A7.2). Both experiments were repeated 10 times, and the results were averaged. Note that since the SI method (spatial interpolation) only considers ATPC of the closest 6 stations, a single experiment was performed, in contrast with the 10 experiments performed for GM and DL.

3. Results

3.1. Data restoration by deep learning

Deep learning generalization performances on the validation sets were used to select the best training configurations to use (Appendix, Fig. A2). The best-performing models were then evaluated on the *"test"* and *"test103"* data sets. Results are reported in terms of *aRMSE*, *aNRMSE*, R^2 and *PCC*. Results are presented for two scenarios: removing the small window of consecutive days (Table 2) and removing large windows of consecutive days (Table 3). The use of the weather input feature impacted positively on the second scenario, while models performed similarly in the first one. This difference is also present when examining the distribution of *NRMSE* over the predicted missing values (Appendix, Figs. A3, A4). Intuitively, the improvement in *NRMSE* is mainly associated with high pollen concentration values, i.e. $>$ 500 pollen grains m^{-3} (Appendix, Fig. A5). To reconstruct missing data in the considered data set, the best-performing architecture was obtained, when training the model to reconstruct small windows of missing data using weather data (ConvAE + weather, normloss, Appendix, Fig. A2). An example of the reconstructed daily pollen concentration data with the best DL model is presented in Appendix, Fig. A6.

3.2. Comparison of the maps prepared on the raw and restored pollen concentration data sets

The original, namely raw, data sets consist of annual average pollen concentrations for the 16-year period (1995–2010), without restoring the missing daily pollen concentration data. For every station in the restored databases, the missing data for given days were restored by entering the values generated by both the Gaussian method (GM) (see Section 2.3.1) and the deep learning (DL) (see Section 2.3.2). For the study period, eight stations (11.9 %), showed the lowest data coverage, while 20 stations (29.9 %) indicated the highest data density (Table 1).

We prepared mean annual total ragweed pollen concentration maps for Europe for the 16-year period, July 15 – October 15, 1995–2010, some including all stations (Fig. 3) and some excluding those for which the mean annual pollen concentration was smaller than 100 pollen grains m^{-3} (Fig. 3*). For both cases, the maps were produced using both the raw data sets (Fig. 3a; *a) and the data sets restored with either the Gaussian method (GM) (Fig. 3b; *b) and the deep learning (DL) (Fig. 3c; *c).

For mapping, we used (a) a concentration filter (the stations with mean annual pollen concentration smaller than 100 pollen grains m^{-3} were excluded from mapping (Figs. 3*a, *b, *c, 4a, b, c, 5a, b; Appendix, Figs. A10*a, *b, *c, A11a, b, c); (b) an altitude filter (the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping (Figs. 3a, b, c, *a, *b, *c, 4a, b, c, 5a, b; Appendix, Fig. A10a, b, c, *a, *b, *c, A11a, b, c); and (c) a station distance filter (the maximum distance is 150 km to each station; Figs. 3a, b, c, *a, *b, *c, 4a, b, c, 5a, b; Appendix, Figs. A10a, b, c, *a, *b, *c, A11a, b, c).

The question arises as to whether restoration of the missing data changed the mapped pollen patterns. When comparing the maps based on the raw vs. restored data, the following results were obtained. If we take the 16-year (1995–2010) mean annual total ragweed pollen data sets including all stations (Fig. 3a; b; c), we find that the maps of the raw data set (Fig. 3a) were very similar to both the restored (GM) (Fig. 3b) and the restored (DL) (Fig. 3c) data sets. However, the maps for both the restored (GM) (Fig. 3b) and the restored (DL) (Fig. 3c) data sets show higher pollen concentrations both for western Europe and the Pannonian Basin. This suggests that data restoration increases the mean seasonal pollen integrals per station (Fig. 3a; b; c). When the stations with mean annual pollen concentrations lower than 100 pollen grains m^{-3} are excluded from the mapping (Fig. 3*a; b; c), then the maps based on the raw data set (Fig. 3*a) show substantial similarity to both the restored (GM) (Fig. 3*b) and the restored (DL) (Fig. 3*c) data sets. In addition, neither the mean seasonal pollen integral in western Europe and the Pannonian Basin, nor the gradient of the pollen counts from east to west were changed (Fig. 3*a; *b). However, the map of the restored (DL) data set (Fig. 3*c) shows a slight increase in these two characteristics compared to the two other maps (Fig. 3*a; *b).

The maps generated after the raw and restored pollen concentration data sets, 2010, including 162 stations were also prepared and compared (Appendix, Section A8).

3.3. Maps of raw phenological data

Based on the raw data sets, the earliest mean start dates of the pollen season occur in the central and southern part of the Pannonian Plain in central Europe at the end of July. For the stations in the southernmost part of France and in the Po River estuary in Italy, airborne ragweed pollen release starts at the beginning of August. In northern Czechia and in Poland the pollen release starts in the middle of August (Fig. 4a). Pollen dispersion occurs at the latest in the middle of October in the Rhône River estuary in France, in western Lombardy in Italy and in the north-eastern part of the Pannonian Plain in Hungary. Meanwhile, the earliest mean end date of the pollen season is observed within the first ten days of October in the north-eastern part of the Pannonian Plain in Hungary (Fig. 4b). The pollen season is the longest in the southern part

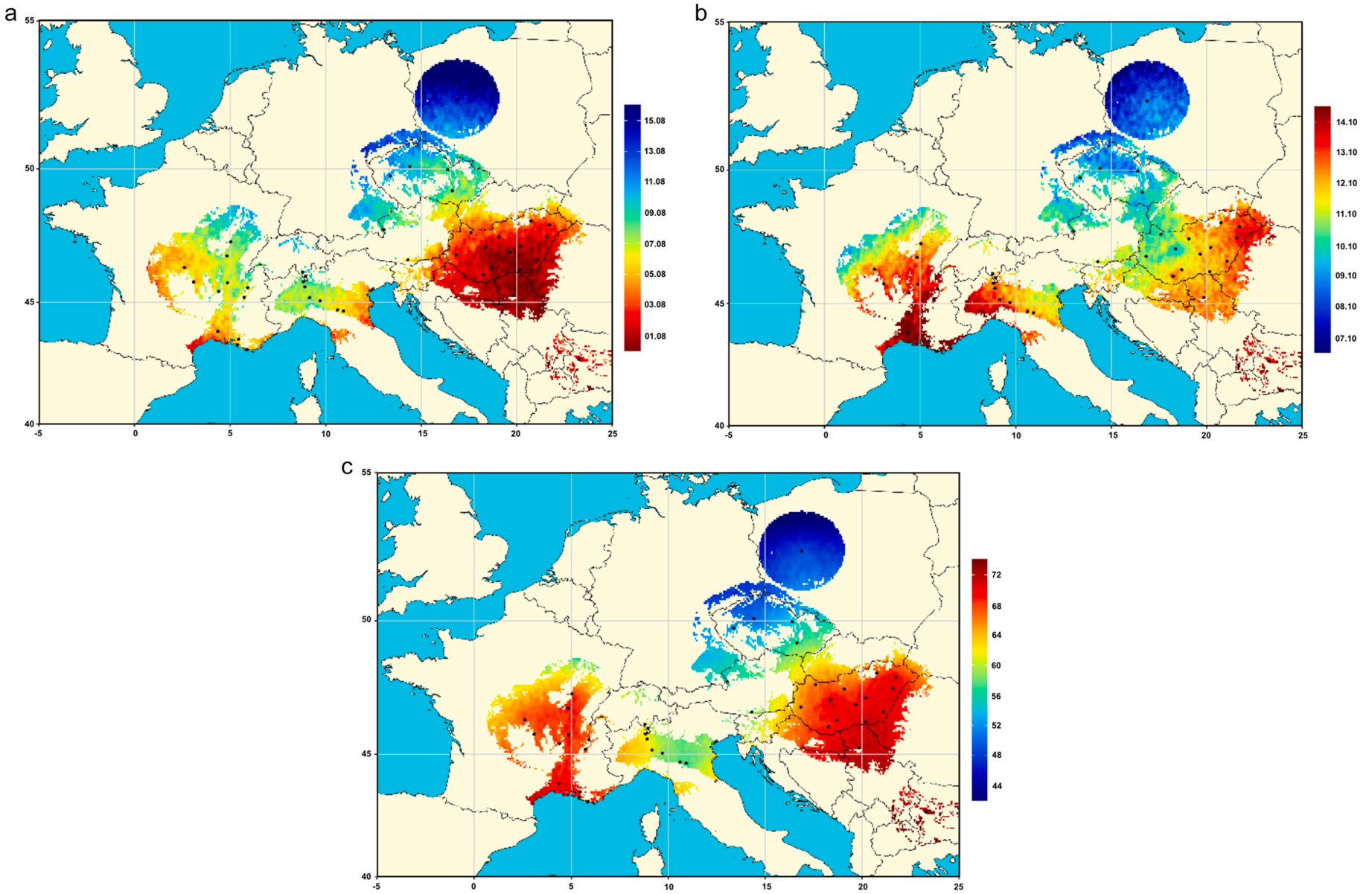


Fig. 4. Mean start date (a), mean end date (b), and mean duration (days) (c) of the ragweed pollen season for Europe, with the geographical locations of the aerobiological stations, using the raw data sets, 16 years, 1995–2010. The stations (1) with mean annual pollen concentration smaller than 100 pollen grains · m^{-3} of air, as well as (2) the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping, and (3) the maximum distance is 150 km to each station. Horizontal axis: longitude, angle in degrees; vertical axis: latitude, angle in degrees. Locations of the aerobiological stations are indicated by black dots.

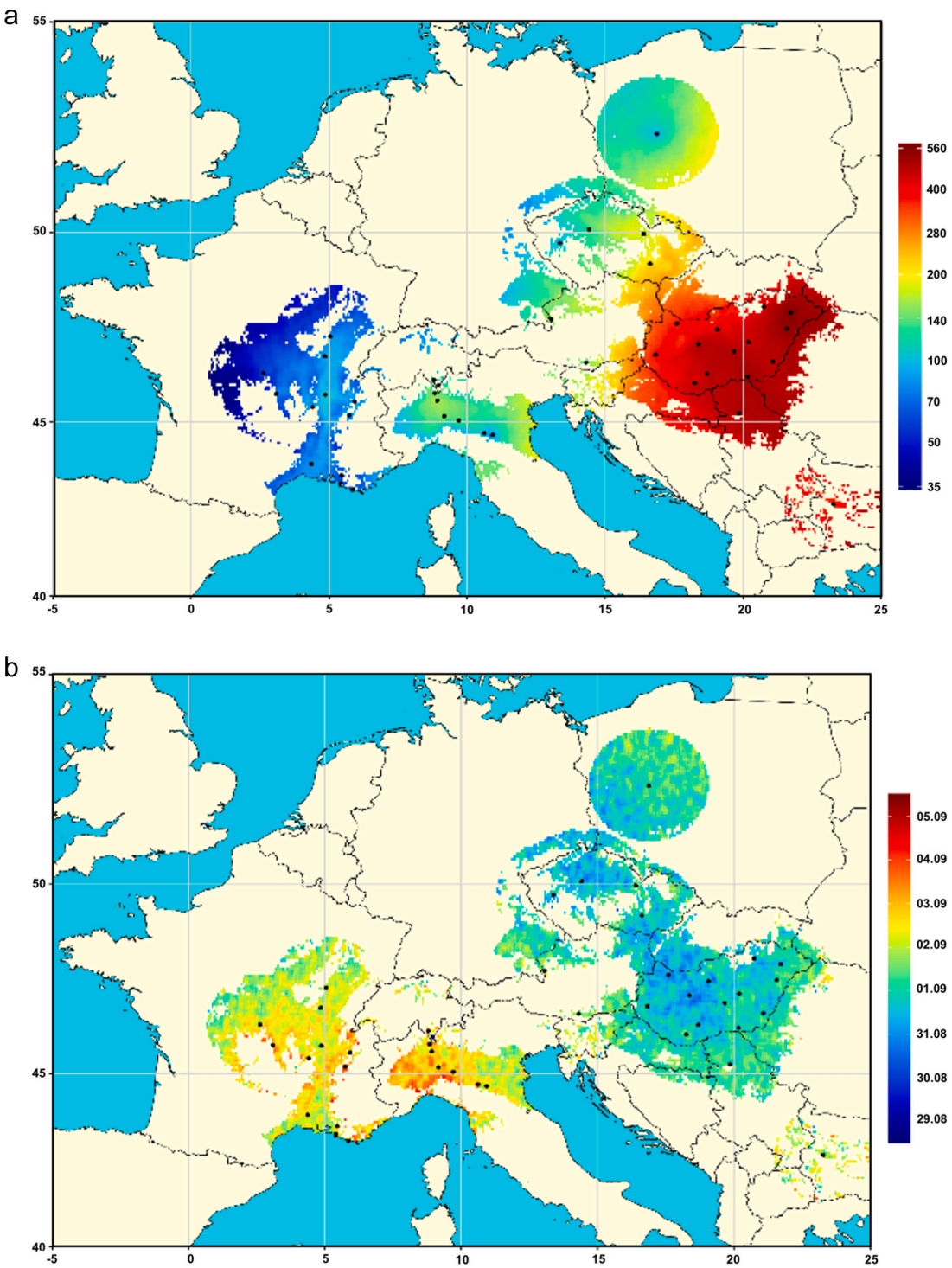


Fig. 5. Mean maximum daily ragweed pollen concentrations (pollen grains · m^{-3} of air) (a), and date of the mean maximum daily ragweed pollen concentration (b) for Europe, using a *base-10* logarithmic colour scale, with the geographical location of the aerobiological stations, using the raw data sets, 16 years, from 1995 to 2010. The stations are: (1) with mean annual pollen concentration smaller than 100 pollen grains · m^{-3} of air, as well as (2) the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping, and (3) the maximum distance is 150 km to each station. Horizontal axis: longitude, angle in degrees; vertical axis: latitude, angle in degrees. Locations of the aerobiological stations are indicated by black dots.

of the Pannonian Plain, followed by the southern part of France, while it is the shortest in northern central Europe (Fig. 4c). Detailed information on the characteristics of the frost-free period (the last frost day in spring, the first frost day in fall and the duration of the frost-free period), as well as the spatial patterns of pollen abundance and flowering phenology in association with the geographical coordinates and the frost-free period are presented in the Appendix (Sections A9 and A10).

3.4. Comparison of the maps of raw and restored maximal pollen counts and their date

Peak values of the mean maximum daily ragweed pollen concentrations were observed in the Pannonian Plain within the Carpathian Basin (Fig. 5a). Areas exceeding 600 pollen grains m^{-3}/day cover the central and southern parts, as well as the north-eastern part of this area,

Table 2
Missing data imputation performances with weather information (ConvAE + Weather) and without weather information (ConvAE). Original data were perturbed removing **small windows** of consecutive days. Metrics are reported for the 1995–2010 data set (5-fold validation and test sets) and for the 2010 test data set (103 stations not present in the training distribution). Reported metrics are the average of Root Mean Squared Error *(aRMSE)*, the average of Normalized Root Mean Squared Error *(aNRMSE)*, the coefficient of determination *(R^2)* and the Pearson Correlation Coefficient *(PCC)*. Pearson's *p*-values are calculated to be significantly close to 0.

Models	Data sets	*aRMSE*	*aNRMSE*	R^2	*PCC*
*ConvAE	[1]val	18.13 ± 30.34	0.23 ± 0.14	0.67 ± 0.04	0.82 ± 0.02
	[2]test	16.90 ± 32.35	0.23 ± 0.15	0.62	0.80
	[3]test103	28.61 ± 40.34	0.20 ± 0.11	0.64	0.82
*ConvAE + Weather	[1]val	18.09 ± 30.58	0.23 ± 0.14	0.67 ± 0.04	0.82 ± 0.02
	[2]test	16.84 ± 31.32	0.23 ± 0.14	0.64	0.81
	[3]test103	28.20 ± 38.93	0.21 ± 0.12	0.65	0.82

[*] Convolutional auto encoder models.
[1] Validation data set.
[2] "*test*" data set.
[3] 103 stations from the 2010 test data set.

Table 3
Missing data imputation performances with weather information (ConvAE +Weather) and without weather information (ConvAE). Original data were perturbed removing **large windows** of consecutive days. Metrics are reported for the 1995–2010 data set (5-fold validation and test sets) and for the 2010 test data set (103 stations not present in the training distribution). Reported metrics are the average of Root Mean Squared Error *(aRMSE)*, the average of Normalized Root Mean Squared Error *(aNRMSE)*, the coefficient of determination *(R^2)* and the Pearson Correlation Coefficient *(PCC)*. Pearson's *p*-values are calculated to be significantly close to 0.

Models	Data sets	*aRMSE*	*aNRMSE*	R^2	*PCC*
*ConvAE	[1]val	23.99 ± 45.94	0.34 ± 0.18	0.39 ± 0.07	0.63 ± 0.06
	[2]test	19.92 ± 33.02	0.32 ± 0.17	0.40	0.65
	[3]test103	39.83 ± 91.38	0.29 ± 0.15	0.19	0.46
*ConvAE + Weather	[1]val	24.58 ± 48.95	0.33 ± 0.17	0.31 ± 0.06	0.60 ± 0.04
	[2]test	16.93 ± 31.00	0.23 ± 0.14	0.65	0.82
	[3]test103	28.45 ± 39.90	0.20 ± 0.12	0.64	0.81

[*] Convolutional auto encoder models.
[1] Validation data set.
[2] "*test*" data set.
[3] 103 stations from the 2010 test data set.

respectively. [The daily ragweed pollen concentration above 100 pollen grains m^{-3} during the main flowering period is many times higher than the minimum concentration values causing symptoms (Déchamp et al., 1997; Makra et al., 2005)].

The date of the mean maximum daily ragweed pollen counts changes within a very narrow time frame (between late August and early September), with the earliest peak values in the Pannonian Plain, while the latest ones occur in western Lombardy and France, respectively (Fig. 5b).

We found that DL model performances were consistently better than GM for estimating ATPC with 10 % to 70 % data removal considering R^2 and *RMSE*. With 80 % missing data, the GM approach works better, since DL underestimates the real value (Table 4, Fig. 6). Compared to GM and DL, the SI method performed poorly, due to over/under-estimation of the real ATPC value (Table 4). The simulation shows that with large percentages of missing data, performances become increasingly similar, and tend to overlap in terms of standard deviation. However, the DL model exhibits a more robust behaviour, and reaches consistently better performances when considering the generated maps (Appendix, Section A7: A7.1, A7.2).

Table 4
ATPC estimation results reported in terms of R^2, *RMSE*, and *MAE*. The DL model used takes as input the additional weather information. For each metric, mean and standard deviation are calculated on the 10 simulations. [Results of the two-step spatial interpolation (SI) are reported in the last row of the table.] (Bold indicates that the given method works better.)

Removed data, %	Method	R^2	*RMSE*	*MAE*
0.1	DL	**0.9971 ± 0.00**	**203.47 ± 44.72**	**123.03 ± 21.53**
	GM	0.9654 ± 0.01	705.25 ± 120.77	460.60 ± 58.38
0.2	DL	**0.9912 ± 0.01**	**343.39 ± 109.26**	**204.52 ± 50.10**
	GM	0.9503 ± 0.05	789.72 ± 333.28	489.66 ± 84.82
0.3	DL	**0.9802 ± 0.01**	**521.27 ± 146.65**	**323.33 ± 59.59**
	GM	0.9130 ± 0.10	1002.28 ± 529.73	545.31 ± 164.75
0.4	DL	**0.9678 ± 0.01**	**680.54 ± 114.38**	**449.96 ± 66.05**
	GM	0.9299 ± 0.05	976.64 ± 286.94	595.34 ± 98.10
0.5	DL	**0.9302 ± 0.02**	**1008.30 ± 118.74**	**644.69 ± 51.88**
	GM	0.8430 ± 0.09	1460.19 ± 433.07	861.59 ± 181.14
0.6	DL	**0.8745 ± 0.03**	**1348.70 ± 187.85**	917.29 ± 100.88
	GM	0.8574 ± 0.07	1410.63 ± 341.93	**886.81 ± 140.82**
0.7	DL	**0.8149 ± 0.06**	**1635.62 ± 244.74**	1196.53 ± 168.97
	GM	0.7901 ± 0.08	1732.98 ± 314.95	**1042.87 ± 158.84**
0.8	DL	0.5262 ± 0.12	2626.26 ± 322.68	1971.28 ± 217.17
	GM	**0.6071 ± 0.25**	**2292.06 ± 742.99**	**1325.30 ± 271.27**
Two-step SI	SI	0.3052	3204.25	2297.74

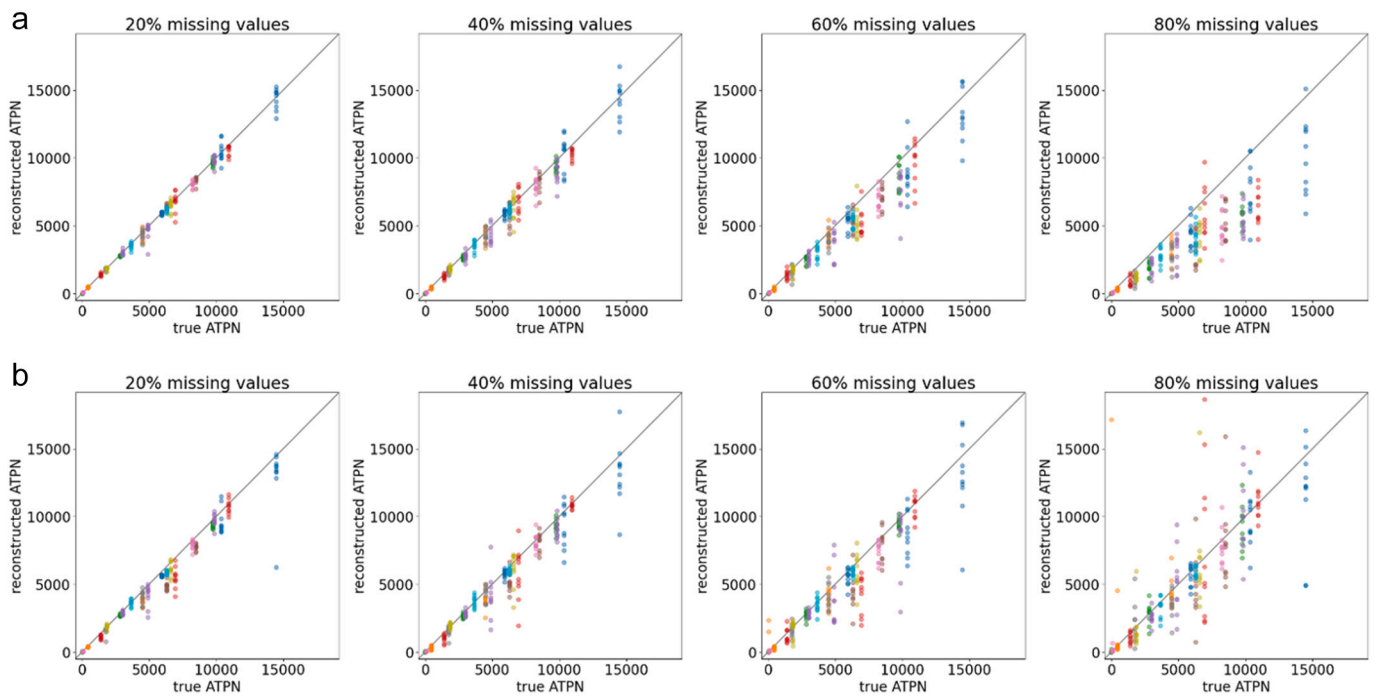


Fig. 6. Comparison of ATPN restoration between (a) DL and (b) GM methods with 20 % to 80 % of missing values. Restoration is performed 10 times by removing data from 25 time series (identified by the colour) extracted from the *"test103"* data set. With a growing percentage of missing data, the DL method tends to underestimate the observed ATPN, while the GM method occasionally yields ATPN values largely different from the target.

4. Discussion

Biodiversity loss and changes under climate change are a major concern according to the IPCC 2022. In addition, the agenda of IPBES (2018) also covered non-native species. Invasive non-native species are strongly linked to biodiversity changes and the expansion of many is favoured by ongoing climate change. *A. artemisiifolia* is one non-native species that has been continuously spreading in central and other parts of Europe (Bullock et al., 2010; Smith et al., 2013; Xian et al., 2023). Many of the current maps used to determine ragweed distribution and pollen release have significant information gaps. With this background, we developed the most up-to-date and accurate ragweed pollen abundance maps currently available for Europe. The maps provide the basis for improved management of pollen-associated respiratory diseases in the future.

The novelty of our research can be summarized in the fact that we applied two approaches, i.e. the Gaussian method and deep learning, to restore incomplete ragweed pollen data sets to improve pollen mapping. As such, this study improves on previous mapping approaches which used interpolation-oriented methods (e.g. European Aeroallergen Network Pollen Database, https://ean.polleninfo.eu/ean; Skjøth et al., 2012; Smith et al., 2013). To the best of our knowledge, this is the first attempt in the literature to restore incomplete pollen data sets.

In addition, we developed maps of ragweed phenology (means of start, end and duration of the pollen season), quantity-related characteristics (means of maximum daily pollen concentrations and the day of the mean maximum daily pollen concentrations) and frost-related parameters (means of last frost day in spring, first frost day in fall and the duration of the frost-free period) for Europe that have not been generated by previous research (e.g. Skjøth et al., 2012; Smith et al., 2013; Storkey et al., 2014). The maps created by our study can therefore be considered the most accurate, most complete and most detailed, with the highest resolution for Europe with the currently available information.

As climatic alterations are shaping habitat structure under the effects of climate change, the biological cycles of plants will also be affected (Ziska and Beggs, 2012). Therefore, it is crucial to be able to track such changes and understand plant responses via monitoring of various ecological metrics, such as their reproductive output (Ziska et al., 2011) and distribution patterns (Chapman et al., 2016). The maps presented in our study are the first to be generated following a detailed cartographical technique. A set of processes was implemented to obtain detailed results with increased accuracy. The methodological approach we used involved: (i) map projection, (ii) refining the border selection technique for the study area; (iii) incorporation of latitude, longitude and elevation information when interpolating [where the interpolation procedure is a combination of three one-dimensional Gaussian kernels]; and (iv) creation of rectangular raster maps to make the maps familiar to the readers. Complementarily, the following approaches improved the information available: (v) the region covered with our maps extends the minimal enclosing rectangle of the measuring stations to show some influential regions around the stations at extreme locations, (vi) the longest data sets (16-year-long station data sets for the period 1995–2010) were used for creating the maps of quantity- and quality-related characteristics of ragweed, (vii) a concentration filter (mean annual pollen concentration smaller than 100 pollen grains m^{-3} are excluded from mapping) and an elevation filter (the areas beyond ±100 m from the altitude of the stations within their 150 km radius are excluded from mapping) were incorporated, and finally (viii) the maps are interpreted by using a wide sliding colour scale, to better reflect the spatial differences in the examined pollen-related characteristics.

The spreading of *A. artemisiifolia* is associated with a large number of local environmental factors including climatic conditions, ecological and habitat preferences, competition with other plants, relief and elevation, as well as changes in land use and affected by management interventions such as timing of mowing. We did not deal with these additional local environmental and human factors influencing ragweed pollen concentration. The reasons for this are: (I) Ragweed is mainly growing in disturbed and unmanaged habitats across Europe, with little competitive vegetation. In the more heavily colonized areas of Europe, ragweed is most frequent on roadsides and in arable fields (Essl et al., 2015), (II) for the managed habitats, only very limited information is

available (e.g. on the date of mowing), and information is missing for some areas (e.g. the CORINE Land Cover Database, as a pan-European land cover inventory does not contain data on the following countries: Croatia, Belarus, Ukraine, and Russia), and (III) information is missing for several years (the CORINE Land Cover Database inventory is only available for the following years based on the period we are examining: 1990, 2000, and 2006. In a previous paper, Deák et al. (2013) examined the potential change in land use for one of the stations included in this paper (Szeged, Hungary). They found that land use changes did not influence ragweed pollen concentration over the Szeged area in the period examined (http://www.eea.europa.eu/publications/COR0-landcover) (Deák et al., 2013).

4.1. Suitability of the distribution maps of common ragweed and their agreement with the pollen concentration maps of ragweed for Europe

Distribution databases may not accurately reflect pollen loads. For example, casual ragweed occurrences in northern Europe, where ragweed populations are small, produce a very little amount of pollen. As a result, pollen data give a better indication of variation in abundance and impacts (Dahl et al., 1999). In addition, distribution databases can under-record in regions like Ukraine (Rodinkova et al., 2018) where monitoring is absent. At the same time, the pollen data can, in principle, help fill the gaps now that more stations in these regions are available online. However, under-recording of distributions seems to correlate with poor pollen monitoring.

Although careful modelling of habitat suitability from the distribution data (e.g. envelope modelling) can be used to forecast ragweed demography. Pollen data could also be used for calibrating or validating habitat suitability models based on native or invasive distributions, providing a powerful new approach (Chapman et al., 2016). To the best of our knowledge, the paper of Chapman et al. (2016) is the only example, where modelling invasion and pollen counts have been correlated.

Distribution data and habitat suitability models do not provide temporal information, while the pollen data are useful for mapping plant phenology and possibly population trends over time (e.g. Ziska et al., 2011). Some habitat suitability models include or are based on flowering phenology (Chapman et al., 2014), so integrating large-scale spatial data on ragweed phenology with model predictions would be very useful for developing better models for ragweed invasion risk as a response to climate-induced changes in temperature.

Distribution maps of *A. artemisiifolia* for Europe using the DAISIE (Bullock et al., 2010; http://www.europe-aliens.org) and the GBIF (Bullock et al., 2010; Cunze et al., 2013; http://data.gbif.org) databases are less accurate due to lack of information (i.e. significant gaps in the data sets). Specifically, (a) the distribution of the European GBIF records for *A. artemisiifolia*, (b) the distribution of *A. artemisiifolia* in Europe either on 50 km × 50 km or 10 km × 10 km grid cells (based on the distribution maps of >40 national aerobiological services and international databases), and (c) the modelled habitat suitability of *A. artemisiifolia* for Europe under the current climatic conditions based on the European data (adventive range approach) are not precise either, and only (d) the modelled habitat suitability of *A. artemisiifolia* for Europe under current climatic conditions based on the North American data (native range approach) provides a good estimation of the habitat suitability of this taxon for Europe (Bullock et al., 2010; Cunze et al., 2013).

4.2. Suitability of the retained stations, data restoration and interpolation maps

Without appropriately accounting or compensating for missing data, data sets may be biased and, as a consequence, any projections may be imprecise. For ragweed over the majority of Europe, this is the current situation. Due to the relatively few aerobiological stations with full ragweed pollen data sets, stations with raw data sets satisfying certain conditions (see Section 2.1) were retained and their missing data were restored through two novel procedures, i.e. the Gaussian method (GM) (see Section 2.3.1) and deep learning (DL) (see Section 2.3.2).

Annual pollen concentrations of several retained stations with complete databases are very low, not even exceeding 100 pollen grains m^{-3}. For such stations, it seems excessive to speak about a pollen season, even in some cases about a longer pollen season reaching 60 days or so. In addition, at stations situated quite far from the source areas, long-range transported ragweed pollen may represent a substantial part of the measured small seasonal pollen integral, reaching even 20 % of the total ragweed pollen load in Germany (Zink et al., 2012). In contrast, in the centre of a source area (e.g. Szeged in Hungary, Pannonian Plain, south-eastern part of central Europe), only 7.5 % is added to the annual total ragweed pollen concentration, due to long-range transport (Makra et al., 2016). However, 7.5 % of Hungarian pollen levels represent a substantially higher absolute amount than 20 % of German pollen levels (Makra et al., 2016). Accordingly, stations with mean annual pollen concentrations smaller than 100 pollen grains m^{-3} were masked when preparing the maps (see Section 3.2). This threshold also concerns the high altitude problem mentioned in Section 2.4.

5. Conclusions

For the first time, we have developed detailed and up-to-date maps of the airborne pollen concentrations of ragweed across the continent of Europe. We applied two statistical approaches, i.e. the Gaussian method (GM) and deep learning (DL) model for restoring missing daily ragweed pollen data sets for Europe. Based on these techniques, we used large ragweed pollen data sets for Europe to produce maps of the mean annual ragweed pollen concentrations for the 16-year period 1995–2010 and for the year 2010. Both types of concentration maps were prepared for the raw and restored data sets, respectively. For maximum comprehensibility and dissemination, we created a web page (http://euragweedpollen.gmf.u-szeged.hu/), including the restored daily ragweed pollen concentration data sets of both the 67 selected stations (16-year data series, 1995–2010) and the 162 selected stations (annual data sets, 2010) (Appendix, Section A5).

The ragweed pollen related maps produced and presented in this paper are unique for Europe. Maps of phenology-related characteristics (means of start, end and duration of the pollen season), quantity-related parameters (means of maximum daily pollen concentration and the day of the maximum daily pollen concentration) and frost-related characteristics (means of last frost day in spring, first frost day in fall and duration of the frost-free period) are novelties in the international literature. In addition, these are the first maps in the literature prepared using altitude correction.

We found that DL model performances are consistently better than those of GM for estimating annual total pollen amounts (ATPC). Ragweed pollen concentration maps derived from the raw data sets were very similar to those based on the restored data sets produced by both the Gaussian method (GM) and deep learning (DL). The restored maps based on both the GM and DL methods show slightly higher values over France, producing a smaller east-to-west gradient in the pollen concentrations after the restoration. This may be attributed to the fact that the pollen concentrations in western Europe have been underestimated due to the higher volume of missing values in this region.

We suggest that the maps developed here, covering a wider geographical and altitudinal range than previously available, are of considerable value in assessing ragweed pollen characteristics for those areas where airborne ragweed pollen is either an existing, or a potential, health threat.

The occurrence of ragweed and the pollen concentrations reported here are, overall, consistent with the known biogeography of ragweed and reflect the temperature (frost-free) conditions under which this species flourishes.

A future prospect will be to associate (i) the actual daily ragweed pollen concentrations reported in this paper and, as an extension, (ii) the cumulative daily pollen concentrations of all remaining allergenic taxa, with public health effects for Europe.

CRediT authorship contribution statement

L.M., †I.M., G.T., L.C., A.G. designed the research; L.M., G.T., L.C., A. G. L.H.Z., J.J.H., L.G.Ny., D.S.Ch., G.J., C.F., H.M.S., M.B., J.M.B., A.D., N.Sch., B.Sz., Á.J.D., Z.S., E.M., P.J. performed research; L.M., L.C., A.G., L.G.Ny. contributed analytic tools; L.M., L.C., A.G., G.J., C.F., M.B. analyzed data; L.M., L.C., A.G., L.H.Z., D.S.Ch., C.F., M.B., J.M.B., A.D., A.C., H.M.S. wrote the paper; L.M. coordinated the study; A.P., D.M., K-C.B., M.T., G.O., R.A., M.B. performed data preparation; B.Š., P.R., M.M. J., R.G., E.S., V.R., O.P., B.S., N.I., U.B., A.K.S., O.R., D.M., K.D-Z., B.M-W., E.W-Ch., Ł.G., P.R., M.M., I.Š., V.S., A.P., A.M., R.P., J.Š. provided airborne pollen data.

Declaration of competing interest

The authors have declared that they have no potential conflict of interest.

Data availability

Data will be made available from data owners per station upon reasonable request.

Acknowledgements

The authors appreciate the advice of Marje Prank and Mikhail Sofiev (Finnish Meteorological Institute, Helsinki, Finland). In addition, the authors would like to thank the European Aeroallergen Network (EAN, https://ean.polleninfo.eu/Ean) for providing ragweed pollen data from Europe, and the Polish Aerobiological Team for the provision of ragweed pollen data from Poland. The study was partly implemented in the frame of the European Cooperation in Science and Technology (EU-COST) program, "New approaches in detection of pathogens and aeroallergens (ADOPT)," Grant CA18226 (EU Framework Program Horizon 2020).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.167095.

References

AAAAI, 2006. (American Academy of Allergy, Asthma and Immunology) 1996-2005. The allergy report: science-based findings on the diagnosis and treatment of allergic disorders. 2006. http://www.theallergyreport.com/reportindex.html. (Accessed 2 December 2022).

Amante, C., Eakins, B.W., 2009. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA, 2009. https://doi.org/10.7289/V5C8276M [2015.04.13].

Anderegg, W.R.L., Abatzoglou, J.T., Anderegg, L.D.L., Bielory, L., Kinney, P.L., Ziska, L., 2021. Anthropogenic climate change is worsening North American pollen seasons. PNAS 118 (7), e2013284118. https://doi.org/10.1073/pnas.20132841188.

Annoni, A., Luzet, C., Gubler, E., Ihde, J. (Eds.), 2003. Map Projections for Europe. European Communities, EUR 20120 EN.

ASCIA, Australian Society of Clinical Immunology and Allergy, 2007. The economic impact of allergic disease in Australia: not to be sneezed at. In: CSIRO, 2007; Canberra, Australia, 111 pp.

Bakar, K.S., Sahu, S.K., 2014. spTimer: Spatio-temporal Bayesian modeling using R. J. Stat. Softw. 63(15), 1-32. URL. http://www.jstatsoft.org/v63/i15/. (Accessed 21 March 2021).

Bartha, D., Schmidt, D., Tiborcz, V., 2019. Atlas florae Hungariae. Kitaibelia 24 (2), 238–252. https://doi.org/10.17542/kit.24.238.

Bartha, D., Bán, M., Schmidt, D., Tiborcz, V., 2022. Vascular Plants of Hungary - Online Database. Department of Botany and Nature Conservation. Faculty of Forestry, Sopron University. http://floraatlasz.uni-sopron.hu. (Accessed 24 October 2022).

Bass, D.J., Delpech, V., Beard, J., Bass, P., Walls, R.S., 2000. Ragweed in Australia. Aerobiologia 16, 107–111. https://doi.org/10.1023/A:1007696112953.

Beggs, P.J., 2018. Climate change and allergy in Australia: an innovative, high-income country, at potential risk. Public Health Res. Pract. 28 (4), e2841828 https://doi.org/10.17061/phrp2841828.

Bonini, M., Monti, G.S., Pelagatti, M.M., Ceriotti, V., Re, E.E., Bramè, B., Bottero, P., Tosi, A., Vaghi, A., Martelli, A., Traina, G.A., Rivolta, L., Rivolta, F., Ortolani, C.M., 2022. Ragweed pollen concentration predicts seasonal rhino-conjunctivitis and asthma severity in patients allergic to ragweed. Sci. Rep. 12, 15921. https://doi.org/10.1038/s41598-022-20069-y.

Bullock, J.M., Chapman, D., Schafer, S., Roy, D., Girardello, M., Haynes, T., Beal, S., Wheeler, B., Dickie, I., Phang, Z., Tinch, R., Čivić, K., Delbaere, B., Jones-Walters, L., Hilbert, A., Schrauwen, A., Prank, M., Sofiev, M., Niemelä, S., Räisänen, P., Lees, B., Skinner, M., Finch, S., Brough, C., 2010. Assessing and controlling the spread and the effects of common ragweed in Europe. Final report: ENV.B2/ETU/2010/0037, Natural Environment Research Council, UK, 456 p. https://circabc.europa.eu/sd/d/d1ad57e8-327c-4fdd-b908-dadd5b859eff/FinalFinalReport.pdf. (Accessed 15 November 2022).

Buttenschøn, R.M., Waldispühl, S., Bohren, C., 2010. Guidelines for management of common ragweed, *Ambrosia artemisiifolia*. University of Copenhagen. 54 p. EUPHRESCO project *AMBROSIA* 2008-09. ISBN: 9788779034549; http://www.EUPHRESCO.org. http://ign.ku.dk/ansatte/skov-natur-biomasse/?pure=files%2F32962432%2Fambrosia_rapport_uk.pdf. (Accessed 13 September 2022).

Chapman, D.S., Haynes, T., Beal, S., Essl, F., Bullock, J.M., 2014. Phenology predicts the native and invasive range limits of common ragweed. Glob. Chang. Biol. 20, 192–202. https://doi.org/10.1111/gcb.12380.

Chapman, D.S., Makra, L., Albertini, R., Bonini, M., Páldy, A., Rodinkova, V., Šikoparija, B., Weryszko-Chmielewska, E., Bullock, J.M., 2016. Modelling the introduction and spread of non-native species: international trade and climate change drive ragweed invasion. Glob. Chang. Biol. 22, 3067–3079. https://doi.org/10.1111/gcb.13220.

Chapman, D.S., Scalone, R., Štefanić, E., Bullock, J.M., 2017. Mechanistic species distribution modeling reveals a niche shift during invasion. Ecol. 98 (6), 1671–1680. https://doi.org/10.1002/ecy.1835.

Che, Z.P., Purushotham, S., Cho, K.H., Sontag, D., Liu, Y., 2018. Recurrent neural networks for multivariate time series with missing values. Sci. Rep. 8 (6085), 1–12. https://doi.org/10.1038/s41598-018-24271-9.

Chen, H., Chen, L., Albright, T.P., 2007a. Developing habitat-suitability maps of invasive ragweed (*Ambrosia artemisiifolia* L.) in China using GIS and statistical methods. In: Lai, P.C., Mak, A.S.H. (Eds.), GIS for Health and the Environment. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-71318-0_8.

Chen, H., Chen, L., Albright, T.P., 2007b. Predicting the potential distribution of invasive exotic species using GIS and information-theoretic approaches: a case of ragweed (*Ambrosia artemisiifolia* L.) distribution in China. Chinese Sci. Bull. 52 (9), 1223–1230. https://doi.org/10.1007/s11434-007-0192-2.

Chiles, J.P., Delfiner, P., 1999. Geostatistics, Modeling Spatial uncertainty. Wiley Series in Probability and statistics. John Wiley & Sons, New York, ISBN 978-0-470-18315-1.

Comtois, P., 1998. Ragweed (Ambrosia sp.): the Phoenix of allergophytes. In: Spieksma, FThM (Ed.), Ragweed in Europe. Satellite Symposium Proceedings of 6th International Congress on Aerobiology, Perugia, Italy. Alk-Abelló A/S, Horsholm, Denmark, pp. 3–5.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An ensemble version of the E-OBS temperature and precipitation data sets. J. Geophys. Res. Atmos. 123, 9391–9409. https://doi.org/10.1029/2017JD028200.

Crisp, H.C., Gomez, R.A., White, K.M., Quinn, J.M., 2013. A side-by-side comparison of Rotorod and Burkard pollen and spore collections. Ann. Allergy Asthma Immunol. 111 (2), 118–125. https://doi.org/10.1016/j.anai.2013.05.021.

Csépe, Z., Leelőssy, Á., Mányoki, G., Kajtor-Apaini, D., Udvardy, O., Péter, B., Páldy, A., Gelybó, G., Szigeti, T., Pándics, T., Kofol-Seliger, A., Simčič, A., Leru, P.M., Eftimie, A.-M., Šikoparija, B., Radišić, P., Stjepanović, B., Hrga, I., Večenaj, A., Vucić, A., Peroš Pucar, D., Škorić, T., Ščevková, J., Kmenta, M., Berger, U., Magyar, D., 2019. The application of a neural network-based ragweed pollen forecast by the ragweed pollen alarm system in the Pannonian biogeographical region. Aerobiologia 36, 131–140. https://doi.org/10.1007/s10453-019-09615-w.

Cunze, S., Leiblein, M.C., Tackenberg, O., 2013. Range Expansion of *Ambrosia artemisiifolia* in Europe is promoted by climate change. Ecol., 610126 https://doi.org/10.1155/2013/610126.

Dahl, Å., Strandhede, S.O., Wihl, J.Å., 1999. Ragweed – an allergy risk in Sweden? Aerobiologia 15, 293–297. https://doi.org/10.1023/A:1007678107552.

D'Amato, G., D'Amato, M., 2023. Climate change, air pollution, pollen allergy and extreme atmospheric events. Curr. Opin. Pediatr. 35 (3), 356–361, 8. https://doi.org/10.3390/atmos14050848.

D'Amato, G., Liccardi, G., Frenguelli, G., 2007. Thunderstorm-asthma and pollen allergy. Allergy 62 (1), 11–16. https://doi.org/10.1111/j.1398-9995.2006.01271.x.

Damialis, A., Traidl-Hoffmann, C., Treudler, R., 2019. Climate Change and Pollen Allergies. In: Marselle, M., Stadler, J., Korn, H., Irvine, K., Bonn, A. (Eds.), Biodiversity and Health in the Face of Climate Change. Springer, Cham, pp. 47–66. https://doi.org/10.1007/978-3-030-02318-8_3.

Damialis, A., Gilles, S., Sofiev, M., Sofieva, V., Kolek, F., Bayr, D., et al., 2021. Higher airborne pollen concentrations correlated with increased SARS-CoV-2 infection rates, as evidenced from 31 countries across the globe. PNAS 118 (e2019034118), 1–10. https://doi.org/10.1073/pnas.2019034118.

Deák, Á.J., Makra, L., Matyasovszky, I., Csépe, Z., Muladi, B., 2013. Climate sensitivity of allergenic taxa in Central Europe associated with new climate change related forces. Sci. Total Environ. 442, 36–47. https://doi.org/10.1016/j.scitotenv.2012.10.067.

Déchamp, C., Rimet, M.L., Meon, H., Deviller, P., 1997. Parameters of ragweed pollination in the Lyon's area (France) from 14 years of pollen counts. Aerobiologia 13, 275–279. https://doi.org/10.1007/BF02694495.

Duan, Y.J., Lv, Y.S., Liu, Y.L., Wang, F.Y., 2016. An efficient realization of deep learning for traffic data imputation. Transp. Res. Part C Emerg. Technol. 72, 168–181. https://doi.org/10.1016/j.trc.2016.09.015.

Essl, F., Bíró, K., Brandes, D., Broennimann, O., Bullock, J., Chapman, D., Chauvel, B., Dullinger, S., Fumanal, B., Guisan, A., Karrer, G., Kazinczi, G., Kueffer, Ch., Laitung, B., Lavoie, C., Leitner, M., Moser, D., Müller-Schärer, H., Petitpierre, B., Richter, R., Schaffner, U., Smith, M., Starfinger, U., Vautard, R., Vogl, G., von der Lippe, M., Follak, S., 2015. Biological Flora of the British Isles: *Ambrosia artemisiifolia*. J. Ecol. 103 (4), 1069–1098. https://doi.org/10.1111/1365-2745.12424.

Hamaoui-Laguel, L., Vautard, R., Liu, L., Solmon, F., Viovy, N., Khvorostyanov, D., Essl, F., Chuine, I., Colette, A., Semenov, M.A., Schaffhauser, A., Storkey, J., Thibaudon, M., Epstein, M.M., 2015. Effects of climate change and seed dispersal on airborne ragweed pollen loads in Europe. Nat. Clim. Chang. 5, 766–772. https://doi.org/10.1038/nclimate2652.

Hess, J.J., 2019. Another piece of the puzzle: linking global environmental change, plant phenology, and health. Lancet Planet. Health 3 (3), e103–e104. https://doi.org/10.1016/S2542-5196(19)30044-0.

Hillerich, V., Valbert, F., Neusser, S., Pfaar, O., Klimek, L., Sperl, A., Werfel, T., Hamelmann, E., Riederer, C., Wobbe-Ribinski, S., Neumann, A., Wasem, J., Biermann-Stallwitz, J., 2023. Quality of life and healthcare costs of patients with allergic respiratory diseases: a cross-sectional study. Eur. J. Health Econ. https://doi.org/10.1007/s10198-023-01598-3.

http://data.gbif.org (GBIF, Global Biodiversity Information Facility) (Accessed May 6, 2022).

http://euragweedpollen.gmf.u-szeged.hu/ (a website that involves both the row and restored data sets of the selected 67 stations (16-year data sets, 1995–2010), with the possibility of uploading new data and/or replacing missing data by using the above-mentioned data restoration process) (Accessed November 19, 2017).

http://www.eea.europa.eu/publications/COR0-landcover. (Accessed 12 August 2022).

http://www.efsa.europa.eu/fr/scdocs/doc/1566.pdf (EFSA Panel on Contaminants in the Food Chain (CONTAM), EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA) and EFSA Panel on Plant Health (PLH); Scientific Opinion on the effect on public or animal health or on the environment on the presence of seeds of *Ambrosia* spp. in animal feed. EFSA Journal 2010; 8(6):1566 [37 pp.]. doi:10.2903/j.efsa.2010.1566. Available online: www.efsa.europa.eu) (Accessed August 23, 2021).

http://www.europe-aliens.org (DAISIE, Delivering Alien Invasive Species Inventories for Europe) (Accessed April 17, 2021).

http://www.polleninfo.org (European Pollen Information) (Accessed June 2, 2018).

https://ean.polleninfo.eu/Ean/ (European Aeroallergen Network Pollen Database) (Accessed February 19, 2017).

https://www.ecad.eu/dailydata/predefinedseries.php# (Blended ECA dataset) (Accessed March 12, 2021).

IPBES, 2018. Plenary of the intergovernmental science-policy platform on biodiversity and ecosystem services, sixth session. Medellin, Colombia, 18–-24 March 2018. https://www.ipbes.net/sites/default/files/ipbes-6-inf-10_en.pdf.

Kahle, D., Wickham, H., 2013. Ggmap: spatial visualization with ggplot2. R J. 5(1), 144-161. http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf. (Accessed 8 April 2022).

Kasprzyk, I., Walanus, A., 2014. Gamma, Gaussian and logistic distribution models for airborne pollen grains and fungal spore season dynamics. Aerobiologia 30, 369–383. https://doi.org/10.1007/s10453-014-9332-8.

Kazinczi, G., Béres, I., Novák, R., Bíró, K., Pathy, Z., 2008. Common ragweed (Ambrosia artemisiifolia): a review with special regards to the results in Hungary. I. Taxonomy, origin and distribution, morphology, life cycle and reproduction strategy. Herbologia 9 (1), 55–91. https://api.semanticscholar.org/CorpusID:86005609.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceeding IJCAI'95 Proceedings of the 14th International Joint Conference on Artificial Intelligence. 2, 1137-1143. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, ISBN 1-55860-363-8.

Lake, I.R., †Jones, N.R., Agnew, M., Goodess, C.M., Giorgi, F., Hamaoui-Laguel, L., Semenov, M.A., Solomon, F., Storkey, J., Vautard, R., Epstein, M.M., 2017. Climate change and future pollen allergy in Europe. Environ. Health Perspect. 125 (3), 385–391. https://doi.org/10.1289/EHP173.

Lake, I., Colon, F., Jones, N., 2018. Quantifying the health effects of climate change upon pollen allergy: a combined cohort and modelling study. Lancet Planet. Health 2, S16. https://doi.org/10.1016/S2542-5196(18)30101-3.

Leiblein-Wild, M.C., Steinkamp, J., Hickler, T., Tackenberg, O., 2016. Modelling the potential distribution, net primary production and phenology of common ragweed with a physiological model. J. Biogeogr. 43, 544–554. https://doi.org/10.1111/jbi.12646.

Liu, L.J., Chen, R.C., 2017. A novel passenger flow prediction model using deep learning methods. Transp. Res. C: Emerg. Technol. 84, 74–91. https://doi.org/10.1016/j.trc.2017.08.001.

Liu, L., Solmon, F., Vautard, R., Hamaoui-Laguel, L., Torma, C.Z., Giorgi, F., 2015. Estimates of common ragweed pollen emission and dispersion over Europe using RegCM-pollen model. Biogeosci. Discuss. 12, 17595–17641. https://doi.org/10.5194/bgd-12-17595-2015.

Liu, L., Solmon, F., Vautard, R., Hamaoui-Laguel, L., Torma, C.Z., Giorgi, F., 2016. Ragweed pollen production and dispersion modelling within a regional climate system, calibration and application over Europe. Biogeosciences 13 (9), 2769–2786. https://doi.org/10.5194/bg-13-2769-2016.

Makra, L., 2022. Tackling ragweed: the international ragweed society held its 2022 world conference in Budapest. Ecocycles 8 (3), 1–5. https://doi.org/10.19040/ecocycles.v8i3.241. Special Issue.

Makra, L., Juhász, M., Béczi, R., Borsos, E., 2005. The history and impacts of airborne *Ambrosia* (Asteraceae) pollen in Hungary. Grana 44 (1), 57–64. https://doi.org/10.1080/00173130510010558.

Makra, L., Matyasovszky, I., Hufnagel, L., Tusnády, G., 2015. The history of ragweed in the world. Appl. Ecol. Environ. Res. 13 (2), 489–512. https://doi.org/10.15666/aeer/1302_489512.

Makra, L., Matyasovszky, I., Tusnády, G., Wang, Y.Q., Csépe, Z., Bozóki, Z., Nyúl, G.L., Erostyák, J., Bodnár, K., Sümeghy, Z., Vogel, H., Pauling, A., Páldy, A., Magyar, D., Mányoki, G., Bergmann, K.C., Bonini, M., Šikoparija, B., Radišić, P., Gehrig, R., Kofol Seliger, A., Stjepanović, B., Rodinkova, V., Prikhodko, A., Maleeva, A., Severova, E., Ščevková, J., Ianovici, N., Peternel, R., Thibaudon, M., 2016. Biogeographical estimates of allergenic pollen transport over regional scales: common ragweed and Szeged, Hungary as a test case. Agric. For. Meteorol. 221, 94–110. https://doi.org/10.1016/j.agrformet.2016.02.006.

Mányoki, G., Apatini, D., Magyar, D., Páldy, A., 2011. A parlagfűpollen becsült országos eloszlása a Parlagfű Pollen Riasztási Rendszer (PPRR) szerint. Assessed incidence of ragweed in Hungary according to the Ragweed Pollen Alarm System (RPAS). In: Apatini, D. (Ed.), Az ÁNTSZ Aerobiológiai Hálózatának tájékoztatója, éves jelentés, kézirat. (Report of the Aerobiological Network of ÁNTSZ, Annual Report) OKI, 2011; Budapest, 81 p. (in Hungarian).

Mentaschi, L., Besio, G., Cassola, F., Mazzino, A., 2013. Problems in RMSE-based wave model validations. Ocean Model. 72, 53–58. https://doi.org/10.1016/j.ocemod.2013.08.003.

Menut, L., Khvorostyanov, D., Couvidat, F., Meleux, F., 2021. Impact of ragweed pollen daily release intensity on long-range transport in Western Europe. Atmosphere 12, 693. https://doi.org/10.3390/atmos12060693.

Montagnani, C., Gentili, R., Smith, M., Guarino, M.F., Citterio, S., 2017. The worldwide spread, success, and impact of ragweed (Ambrosia spp.). Crit. Rev. Plant Sci. 36 (3), 139–178. https://doi.org/10.1080/07352689.2017.1360112.

Montagnani, C., Gentili, R., Citterio, S., 2023. Ragweed is in the air: *Ambrosia* L. (Asteraceae) and pollen allergens in a changing world. Curr. Protein Pept. Sci. 24 (1), 98–111. https://doi.org/10.2174/1389203724666221121163327.

Navares, R., Aznarte, J.L., 2019. Geographical imputation of missing poaceae pollen data via convolutional neural networks. Atmosphere 10 (11), 717. https://doi.org/10.3390/atmos10110717.

Oteros, J., Bergmann, K.C., Menzel, A., Damialis, A., Traidl-Hoffmann, C., Schmidt-Weber, C.B., Buters, J., 2019. Spatial interpolation of current airborne pollen concentrations where no monitoring exists. Atmos. Environ. 199, 435–442. https://doi.org/10.1016/j.atmosenv.2018.11.045.

Prank, M., Chapman, D.S., Bullock, J.M., Belmonte, J., Berger, U., Dahl, A., Jäger, S., Kovtunenko, I., Magyar, D., Niemelä, S., Rantio-Lehtimäki, A., Rodinkova, V., Sauliene, I., Severova, E., Sikoparija, B., Sofiev, M., 2013. An operational model for forecasting ragweed pollen release and dispersion in Europe. Agric. For. Meteorol. 182-183, 43–53. https://doi.org/10.1016/j.agrformet.2013.08.003.

R Core Team, 2014. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. (Accessed 11 May 2022).

Rasmussen, K., Thyrring, J., Muscarella, R., Borchsenius, F., 2017. Climate-change-induced range shifts of three allergenic ragweeds (*Ambrosia* L.) in Europe and their potential impact on human health. PeerJ 5 (11), e3104. https://doi.org/10.7717/peerj.3104.

Rauer, D., Gilles, S., Wimmer, M., Frank, U., Mueller, C., Musiol, S., Vafadari, B., Aglas, L., Ferreira, F., Schmitt-Kopplin, P., Durner, J., Winkler, J.B., Ernst, D., Behrendt, H., Schmidt-Weber, C.B., Traidl-Hoffmann, C., Alessandrini, F., 2020. Ragweed plants grown under elevated CO_2 levels produce pollen which elicit stronger allergic lung inflammation. Allergy 76 (6), 1718–1730. https://doi.org/10.1111/all.14618.

Recio, M., Docampo, S., García-Sánchez, J., Trigo, M.M., Melgar, M., Cabezudo, B., 2010. Influence of temperature, rainfall and wind trends on grass pollination in Malaga (western Mediterranean coast). Agric. For. Meteorol. 150, 931–940. https://doi.org/10.1016/j.agrformet.2010.02.012.

Richter, R., Berger, U.E., Dullinger, S., Essl, F., Leitner, M., Smith, M., Vogl, G., 2013. Spread of invasive ragweed: climate change, management and how to reduce allergy costs. J. Appl. Ecol. 50, 1422–1430. https://doi.org/10.1111/1365-2664.12156.

Rodinkova, V., Palamarchuk, O., Toziuk, O., Yermishev, O., 2018. Modeling hay fever risk factors caused by pollen from *Ambrosia* spp. using pollen load mapping in Ukraine. Acta Agrobot. 71 (3), 1742. https://doi.org/10.5586/aa.1742.

Rodríguez-Rajo, F.J., Aira, M.J., Fernández-González, M., Seijo, C., Jato, V., 2011. Recent trends in airborne pollen for tree species in Galicia, NW Spain. Clim. Res. 48 (2/3), 281–291. https://doi.org/10.3354/cr00966.

Rousseau, D.D., Schevin, P., Duzer, D., Cambon, G., Ferrier, J., Jolly, D., Poulsen, U., 2005. Pollen transport to southern Greenland: new evidences of a late spring long-distance transport. Biogeosci. Discuss. 2 (4), 829–847. https://doi.org/10.5194/bgd-2-829-2005.

Rousseau, D.D., Schevin, P., Duzer, D., Cambon, G., Ferrier, J., Jolly, D., Poulsen, U., 2006. New evidence of long-distance pollen transport to southern Greenland in late spring. Rev. Palaeobot. Palynol. 141 (3–4), 277–286. https://doi.org/10.1016/j.revpalbo.2006.05.001.

Schaffner, U., Steinbach, S., Sun, Y., Skjøth, C.A., de Weger, L.A., Lommen, S.T., Augustinus, B.A., Bonini, M., Karrer, G., Šikoparija, B., Thibaudon, M., Müller-

Schärer, H., 2020. Biological weed control to relieve millions from *Ambrosia* allergies in Europe. Nat. Commun. 11, 1745 https://doi.org/10.1038/s41467-020-15586-1.
Simonoff, J.S., 1996. Smoothing Methods in Statistics. Springer Series in Statistics, Springer, New York.
Singh, A.B., Mathur, Ch., 2021. Climate change and pollen allergy in India and South Asia. Immunol. Allergy Clin. N. Am. 41, 33–52. https://doi.org/10.1016/j.iac.2020.09.007.
Skjøth, C.A., Šikoparija, B., Jäger, S., EAN network, 2012. Pollen sources. In: Sofiev, M., Bergmann, K.C. (Eds.), Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts. Springer Science+Business Media, Dordrecht, Netherlands, pp. 9–27. https://doi.org/10.1007/978-94-007-4881-1_2.
Skjøth, C.A., Sun, Y., Karrer, G., Šikoparija, B., Smith, M., Schaffner, U., Müller-Schärer, H., 2019. Predicting abundances of invasive ragweed across Europe using a "top-down" approach. Sci. Total Environ. 686, 212–222. https://doi.org/10.1016/j.scitotenv.2019.05.215.
Smith, M., Cecchi, L., Skjøth, C.A., Karrer, G., Šikoparija, B., 2013. Common ragweed: A threat to environmental health in Europe. Environ. Int. 61, 115–126. https://doi.org/10.1016/j.envint.2013.08.005.
Sofiev, M., Siljamo, P., Ranta, H., Rantio-Lehtimäki, A., 2006. Towards numerical forecasting of long-range air transport of birch pollen: theoretical considerations and a feasibility study. Int. J. Biometeorol. 50, 392–402. https://doi.org/10.1007/s00484-006-0027-x.
Sofiev, M., Galperin, M., Genikhovich, E., 2008. A construction and evaluation of Eulerian dynamic Core for the air quality and emergency modelling system SILAM. In: Borrego, C., Miranda, A.I. (Eds.), Air Pollution Modeling and Its Application XIX. NATO Science for Peace and Security Series Series C: Environmental Security. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-8453-9_94.
Sofiev, M., Siljamo, P., Ranta, H., Linkosalo, T., Jaeger, S., Rasmussen, A., Rantio-Lehtimaki, A., Severova, E., Kukkonen, J., 2013. A numerical model of birch pollen emission and dispersion in the atmosphere. Description of the emission module. Int. J. Biometeorol. 57, 45–58. https://doi.org/10.1007/s00484-012-0532-z.
Stone, E.A., Mampage, C.B.A., Hughes, D.D., Lillian, M., Jones, L.M., 2021. Airborne sub-pollen particles from rupturing giant ragweed pollen. Aerobiologia 37, 625–632. https://doi.org/10.1007/s10453-021-09702-x.
Storkey, J., Stratonovitch, P., Chapman, D.S., Vidotto, F., Semenov, M.A., 2014. A process-based approach to predicting the effect of climate change on the distribution of an invasive allergenic plant in Europe. PloS One 9 (2), e88156. https://doi.org/10.1371/journal.pone.0088156.
van der Knaap, W.O., van Leeuwen, J.F.N., Froyd, C.A., Willis, K.J., 2012. Detecting the provenance of Galápagos non-native pollen: the role of humans and air currents as transport mechanisms. Holocene 22, 1373–1383. https://doi.org/10.1177/0959683612449.
Wayne, P., Foster, S., Connolly, J., Bazzaz, F., Epstein, P., 2002. Production of allergenic pollen by ragweed (*Ambrosia artemisiifolia* L.) is increased in CO_2-enriched atmospheres. Ann. Allergy Asthma Immunol. 88 (3), 279–282. https://doi.org/10.1016/S1081-1206(10)62009-1.
Xian, X.Q., Zhao, H.X., Wang, R., Huang, H.K., Chen, B.X., Zhang, G.F., Liu, W.X., Wan, F. H., 2023. Climate change has increased the global threats posed by three ragweeds (*Ambrosia* L.) in the Anthropocene. Sci. Total Environ. 859 (Part 2), 160252. https://doi.org/10.1016/j.scitotenv.2022.160252.
Yamburov, M.S., Astafurova, T.P., Zhuk, K.V., Romanova, S.B., Smolina, V.M., 2014. The effects of drought and flood stress on pollen quality and quantity in Clivia miniata (Lindl.) Bosse (Amaryllidaceae). Biomed. Pharmacol. J. 7 (2), 575–580. https://doi.org/10.13005/bpj/526.
Zewdie, G.K., Lary, D.J., Levetin, E., Garuma, G.F., 2019a. Applying deep neural networks and ensemble machine learning methods to forecast airborne *Ambrosia* pollen. Int. J. Environ. Res. Public Health 16 (11), 1–14. https://doi.org/10.3390/ijerph16111992.
Zewdie, G.K., Liu, X., Wu, D.J., Lary, D.J., Levetin, E., 2019b. Applying machine learning to forecast daily *Ambrosia* pollen using environmental and NEXRAD parameters. Environ. Monit. Assess. 191 (S2), 261, 1-11. https://doi.org/10.1007/s10661-019-7428-x.
Zhang, Y., Bielory, L., Cai, T., Mi, Z.Y., Georgopoulos, P., 2015. Predicting onset and duration of allergenic pollen season in the United States. Atmos. Environ. 103, 297–306. https://doi.org/10.1016/j.atmosenv.2014.12.019.
Zhang, Y., Thorburn, P.J., Xiang, W., Fitch, P., 2019. SSIM – a deep learning approach for recovering missing time series sensor data. IEEE Internet Things J. 6 (4), 6618–6628. https://doi.org/10.1109/JIOT.2019.2909038.
Zhao, J.H., Nie, Y.W., Ni, S.J., Sun, X.K., 2020. Traffic data imputation and prediction: an efficient realization of deep learning. IEEE Access, Spec. Sect. Big Data Technol. Appli. Intell. Transp. 8, 46713–46722. https://doi.org/10.1109/ACCESS.2020.2978530.
Zink, K., Vogel, H., Vogel, B., Magyar, D., Kottmeier, C., 2012. Modeling the dispersion of *Ambrosia artemisiifolia* L. pollen with the model system COSMO-ART. Int. J. Biometeorol. 56, 669–680. https://doi.org/10.1007/s00484-011-0468-8.
Ziska, L.H., Beggs, P.J., 2012. Anthropogenic climate change and allergen exposure: the role of plant biology. J. Allergy Clin. Immunol. 129 (1), 27–32. https://doi.org/10.1016/j.jaci.2011.10.032.
Ziska, L., Knowlton, K., Rogers, C., Dalan, D., Tierney, N., Elder, M.A., Filley, W., Shropshire, J., Ford, L.B., Hedberg, C., Fleetwood, P., Hovanky, K.T., Kavanaugh, T., Fulford, G., Vrtis, R.F., Patz, J.A., Portnoy, J., Coates, F., Bielory, L., Frenz, D., 2011. Recent warming by latitude associated with increased length of ragweed pollen season in central North America. PNAS 108 (10), 4248–4251. https://doi.org/10.1073/pnas.1014107108.
Ziska, L.H., Makra, L., Harry, S.K., Bruffaerts, N., Hendrickx, M., Coates, F., Saarto, A., Thibaudon, M., Oliver, G., Damialis, A., Charalampopoulos, A., Vokou, D., Heiðmarsson, S., Guðjohnsen, E., Bonini, M., Oh, J.-W., Sullivan, K., Ford, L., Brooks, G.D., Myszkowska, D., Severova, E., Gehrig, R., Ramón, G.D., Beggs, P.J., Knowlton, K., Crimmins, A.R., 2019. Temperature-related changes in airborne allergenic pollen abundance and seasonality across the northern hemisphere: a retrospective data analysis. Lancet Planet. Health 3 (3), e124–e131. https://doi.org/10.1016/S2542-5196(19)30015-4.