Metadata of the chapter that will be visualized in SpringerLink

Book Title	ICT Innovations 2016	
Series Title		
Chapter Title	An Automatic Tracking System for Natural Hazard Events with Satellite Remote Sensing	
Copyright Year	2018	
Copyright HolderName	Springer International Publishing AG	
Corresponding Author	Family Name	**Tchorbadjieff**
	Particle	
	Given Name	**Assen**
	Prefix	
	Suffix	
	Division	
	Organization	Institute of Mathematics and Informatics
	Address	Acad. Georgi Bonchev Str., Block 8, 1113, Sofia, Bulgaria
	Email	atchorbadjieff@math.bas.bg

Abstract

The atmosphere satellite data for atmosphere parameters are the most important source of information for monitoring of areas without or with very rare environment research facilities. With growing dynamic of Climate change, the detailed observation, research and risk management is with a vital importance for nations in regions as the South-East Europe. Due to insufficient ground based research infrastructure and qualified personal, the satellites are main source of reliable data of atmosphere process. The presented paper describes the basic available functionalities of a system for automatic atmosphere events location and transport prediction based on available open data form NASA satellites.

Keywords (separated by '-')

Satellites - Computational physics - Natural events - Spatial-temporal data

An Automatic Tracking System for Natural Hazard Events with Satellite Remote Sensing

Assen Tchorbadjieff(✉)

Institute of Mathematics and Informatics,
Acad. Georgi Bonchev Str., Block 8, 1113 Sofia, Bulgaria
atchorbadjieff@math.bas.bg

Abstract. The atmosphere satellite data for atmosphere parameters are the most important source of information for monitoring of areas without or with very rare environment research facilities. With growing dynamic of Climate change, the detailed observation, research and risk management is with a vital importance for nations in regions as the South-East Europe. Due to insufficient ground based research infrastructure and qualified personal, the satellites are main source of reliable data of atmosphere process. The presented paper describes the basic available functionalities of a system for automatic atmosphere events location and transport prediction based on available open data form NASA satellites.

Keywords: Satellites · Computational physics · Natural events · Spatial-temporal data

1 Introduction

The different developments due to climate change cause very strong impact on mankind living behavior. Most of them are connected to atmosphere processes, which are the most imminent and dynamic in impact. The range of their influence is from brief to long term climate changes and from short haze through causing serious damage floods to long term drought and desert spread. The first big serious event that was observed in contemporary history is The Year Without a Summer in 1816 when massive eruption of Mount Tambora a year earlier caused global cooling. Since then the interest in study of atmosphere and its impact have been growing. As a part of this efforts, with technical revolution many new scientific instruments was involved in the ground and space observations of atmosphere.

The most perspective and very important sources of scientific data in last decades came from space satellites. Moreover, with their fast technical improvements the possibilities for monitoring are expanded. In addition to classical meteorology observations, the monitoring of air quality, aerosols particles and volcano eruptions, wildlife fires and Dust storms are enabled. With acquired data and its appropriately proceeding with computer programs it is possible to detect and

G. Stojanov and A. Kulakov (eds.), *ICT Innovations 2016*,
Advances in Intelligent Systems and Computing 665,
https://doi.org/10.1007/978-3-319-68855-8_24

distinguish different sources of pollution. A very good demonstration of effectiveness and usefulness of such type of analysis is the latest advanced research of regional SO2 and NO2 pollution changes from 2005 to 2015 with data from AURA OMI satellite [1].

The importance and usefulness of data from satellites are even more extended in areas and territories where ground based observatories are very rare and not well equipped. The similar region is Balkan peninsula, where only one permanent observation facility is available. It is Basic Environmental observatory at Moussala, the highest place in the Balkans [2]. However, the region is an area that is very strongly exposed to environmental changes, caused either from anthropogenic and natural factors. The region is exposed on the path of Sahara dust air masses transport [3,4], there are many local pollutants, the devastating wildfires are usual events during the late summer. Despite the fact that satellites data are not enough detailed mainly because of lack of permanent measurements [4], the available satellite data are important for region monitoring.

2 Automatic Detection System

2.1 Software Configuration

The specially dedicated software for observation of pollutants and extreme events detection and location is under development. For development are selected the fully functional tools for statistical computing based on R [5]. The R environment enables the usage of all possible resources of contemporary computational science and meteorology multiprocess and cluster computations, automatic data acquisition, parsing and segmentation, air transport simulation and finally- the machine learning based on all possible available statistical tools. For current version, the system is only used for training based on historical data observations. However, the development of R and technical infrastructure allows further functionality expansion of current system to real time analytical tool.

The model is designed to run on distributed computing infrastructure due to design requirements of large memory and computational power. The current results are computed on high-performance grid computing infrastructure at Institute of Information and Communication Technologies BAS in Sofia [6]. The raw satellite data, available in HDF format, are acquired by standard wget procedure. Then, the specially developed script in R processes and extracts the required data subsets from HDF format to more convenient for processing in R data formats. For advanced computing the High-Performance and Parallel Computing functionalities in R are used [7]. With them the access to every different satellite parameter is autonomous and could be processed next to parametric dependent analytic tools. The general view of the infrastructure is shown in (Fig. 1).

The data analysis system relies mainly on quality of NASA published data. The first step in data processing is data extraction based on available metadata for validity and measurement errors. This process is run by different multiprocessing data drivers, each one of them enables autonomy of different parame-

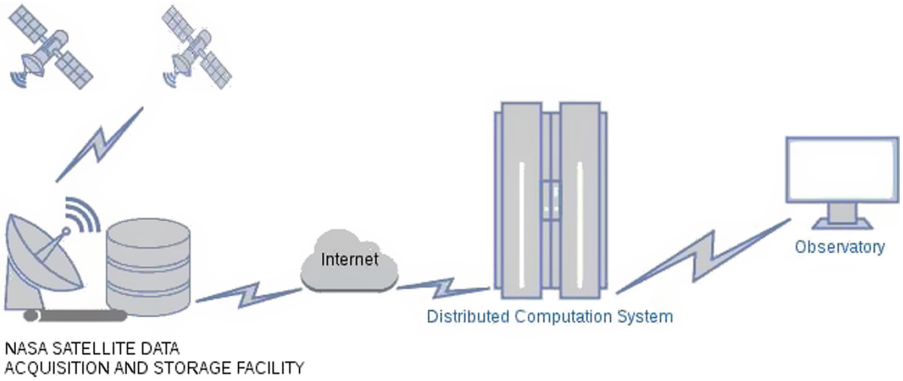

Fig. 1. The system relies mainly on NASA satellite open data system. The required data is downloaded via Internet to high-performance grid computing system for computation. The computations are observable and produced results are accessible through direct link.

ters observation. The R computational environment enhances the customization of quick and reliable statistical analysis with acquired data. The most possible outcomes from these analyses are air pollutant detection and large scientific information for research. Then, for the further data analysis of air mass transport, the most important information could be acquired from air mass backward and forward in time trajectories. For modeling air mass transport the system relies on NOAA HYSPLIT model dedicated for atmospheric trajectory and dispersion calculations [8]. For automatic computations of required trajectories and their plot in R are selected the functionalities of Openair R library [9]. The software library is distributed with a number of R scripts, which allows complete automatic raw meteo data acquisition from NOAA and proceeding of obtained data to trajectory computations with custom selected parameters in both directions, backward and forward in time. The acquired results are preprocessed and stored in shared system directory and could be accessed by different systems analytical drivers (Fig. 2).

2.2 Spatio-Temporal Design

The preprocessed trajectory records enable discovering of number of possible sources or predict the destination of air pollution immediately after its detection from satellite based on HYSPLIT trajectories. For this purpose, the statistical analysis of merged information of satellite data and computed particle transport trajectories is assumed to locate the most probable source and/or destination of pollutant. This combination of remote sensing data with trajectory creates 4-dimensional spatial-temporal space. The 2-dimensional spatial domain S (latitudes and longitudes) and temporal domain T (time of measurements) define 3-dimensional spatio-temporal sample $z = (z(s_1, t_1), ..., z(s_n, t_n))$, with locations

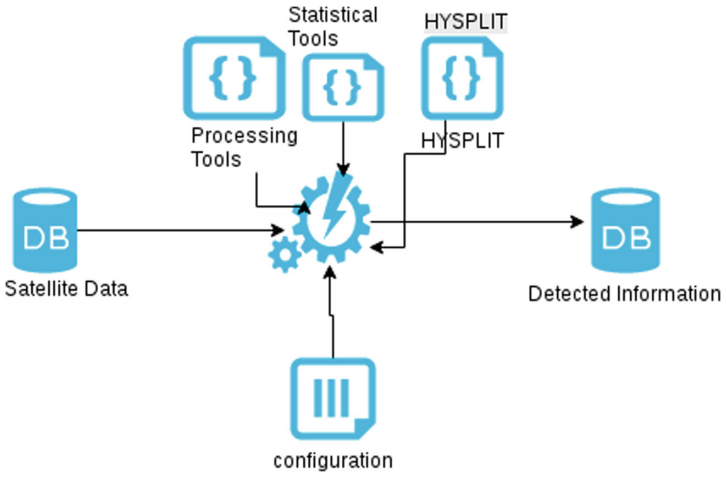

Fig. 2. The automatic system consists of assembly of preprocessing, statistical and trajectory computation software scripts and tools. The system management and synchronization between different parts are relied on configuration file.

$(s_1, t_1), ..., (s_n, t_n) \subseteq R^2 \times R$ [10,11]. For statistical analysis of spatial and temporal dispersion of pollutant particles are used spatial-temporal covariance for spatial distance h and temporal distance u: [10,11]:

$$C_{st}(h, u) = cov(Z(s, t), Z(\bar{s}, \bar{t})) \tag{1}$$

for $(s, t), (\bar{s}, \bar{t}) \in SxT$ and $||s - \bar{s}|| < h; ||t - \bar{t}|| < u$. Respectively, spatial-temporal variograms are equal to [11]:

$$\gamma_{st}(h, u) = C_{s,t}(0, 0) - C_{s,t}(h, u) \tag{2}$$

Having computed in parallel of multiple different in initial altitude trajectories from the same spatial-temporal location, the vertical 3-rd spatial dimension is yielded in comparison between them.

The implemented solution is based on output results from particle transport trajectories. They are with time resolution of 1 h. However, the satellites could not scan observed place from same stationarity less than once a day. Due to this constrains the different implementations are assumed for approximation of distances between trajectories and coordinates of remotely scanned cites. The coincident between trajectory and satellite data is assumed to occur when the region where both, the coordinates of the latitude and longitude of trajectory $W_t(\Delta_i)$ and these of the remote scan $W_s(\Delta_i)$, are in the same location window during equal time interval Δ_i:

$$\sqrt{|W_t(\Delta_i) - W_s(\Delta_i)|} <= \epsilon \tag{3}$$

With ϵ is denoted the maximum allowed distance. For practical analysis three different equal values for location with $\epsilon = (0.5, 1, 1.5)$ and time with $\Delta_i = 3; 6; 12$ h are assumed.

3 Results and Demonstrations

For numerical demonstration the explanatory information from Sahara dust events (SDE) over Balkan region during the Spring 2013 is used. The data are acquired from NASA Aqua/AIRS satellite as dust score index, which is computed from several different infrared channels on board of satellite [12]. The score values above 360 indicate presence of mineral dust in atmosphere. However, this score should not be assumed as being a quantitative measure since this satellite is not trained on the Balkan Peninsula territory. Thus, the dust scores with flag for land uncertainties equal to 1 over observed territories may be considered only as a qualitative confirmation of SDE detection, rather than a quantitative measure. The wrongness due to clouds were excluded because of very high error [12].

3.1 Quantitative Results

The selected SDEs are three mineral outbursts which are detected above Balkans on May 28th, May 19–20th and on May 16–17th [4]. The observed coordinates are for the area above the territory of Republic of Macedonia. The dust scores confirm mineral dust arrival with relatively low values in between 360 and 370 during the midnight on May 18th. The next available observations for SDE are from midnight of May 20th, where the score range is between 438 and 502. The last SDE on May 29th and May 30th are confirmed with day and night observations with values between 406 and 511. The visual maps available on NASA Worldview [14], as well as raw data directly extracted from processing software, show that the main area for the first event is mainly on South-East direction and detected particles most probably are dispersed from the main front. However, the two bigger events which occurred later completely covers the territory of Republic of Macedonia.

For simplicity, the computed trajectories begin from place with averaged coordinates for multiple observations during the same hour above territory of Republic of Macedonia. But the software originally is designed to obtain coordinates directly from satellite data output. The number of selected levels about ground levels are also optional. Due to atmosphere models, the higher altitude yields the longer directories and investigated travel distances are longer. The demonstrative trajectories with starting point from Skopje with 4000 and 100 m above ground level (m AGL) with time of beginning during the SDE of May 19–20th are generated and shown in Fig. 3.

The results from comparison between the events show correlation between intensity and number of coincidences in the trajectories and satellite data over all different distance intervals. Because the detected particles at 18 May 2013 are from SDE which is not directly associated with the territory of Republic of Macedonia, the number of coincidences are very lower than the direct SDE events from May 19–20th and May 29. The magnitude of differences for these numbers could be more than 7 times for 36 h backward trajectories. This may be explained with the nature of particle transport. Because the event from May 18th

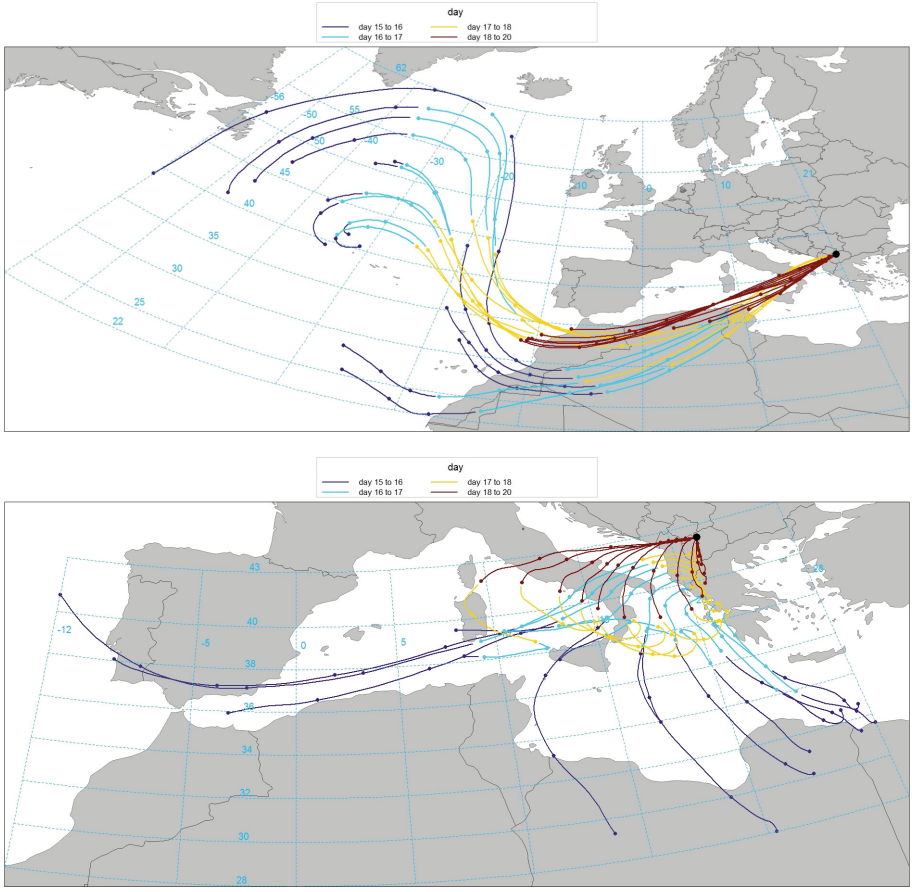

Fig. 3. The graphics show automatically generated backward trajectories with OpenAir library [9] in R for HYSPLIT model. The starting point is the region of Skopje in between 0000UTC at May 19th to 0000UTC at May 21st. The starting altitude for upper graphic is 4000 m agl and 100 m agl for lower. The backward paths are 96 h long.

is mainly associated with dispersed local transport from different source of mineral plume, the coincidences are missed during backward trajectories of air-masses. The only detected reliable coincidences except those above mainly observed territory are at nearby region above Albania and their Mediterranean coasts, which is more likely to show availability of similar dispersed particles. However, the coincidences during the next two events are available over the backward part of air-masses, which confirm hypothesis of direct transport.

Besides, the comparison between number of coincidences in trajectories and satellite data confirms the expectations of tradeoffs between lower precision in spatial-temporal domains and the area of covered territory. The data from tests show that when the model is run with larger window for coincidences the

observed area is also larger, conversely to the case with more narrow observations sample space (Fig. 4). Moreover, as data from both figures show that lack of results during backward transport is possible outcome in case of small spatial-temporal window of coincidences. The most acceptable explanation for this are uncertainties in trajectory models and time synchronization between them and satellite measurements.

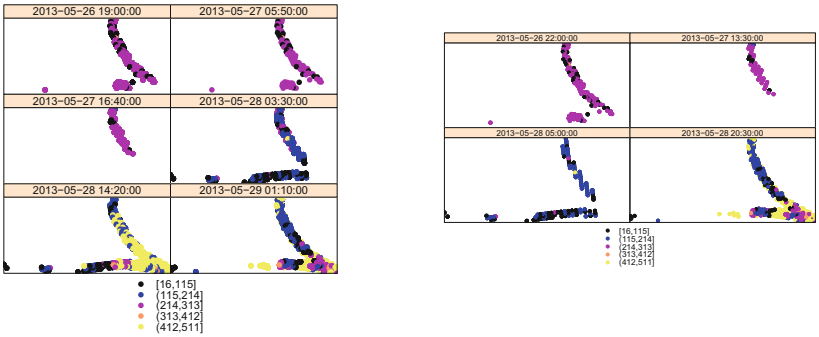

Fig. 4. The graphics show the number of registered satellite data and their Dust score with time window $\Delta = 6$ (left) and $\Delta = 3$ (right) hours for 72 h long trajectory with starting time at 29 May 2013 1200 UTC at 1000 m. AGL

The data from Table 1 show the different pattern of computational output as a result of the initial trajectory altitude. When the altitude is higher the longer distance range is covered but the counts of coincidence are smaller. This could be noticed at the Fig. 3 where long time backward trajectory paths reach the Atlantic Ocean, the place where mineral dust is not observed. Conversely, when the initial altitude is lower, the covered area for the same observed time window is smaller. This allow to investigate smaller area. Following this dependency the initial trajectory altitude is parameter dependent. In case of Sahara dust transport the higher altitudes yields better results. But when satellite data is used for local pollutants, as forest fires for example, the very small values of altitude, like 10 or 100 m agl for example, should be considered.

3.2 Spatial-Temporal Statistics

The geostatistical analysis confirms the intuitive expectations and quantitative results. As the variograms show, that computations with higher initial altitudes are more suitable for long range rural transport in comparison to these with lower ones. As the data from Table 1 show, the lower altitude trajectory missed the source and only intercept the plume transport above Greece and Mediterranean. These results yield random spatial-temporal distribution in diurnal basis (Fig. 5). Conversely, the variogram produced from data form transport in higher altitudes (2000m AGL starting point) shows clear trend in both domains - time and

Table 1. The most distant area of registered Sahara dust plume during the SDE event from May 29th 2013 with computed coincidences for 36 h long backward trajectories for altitudes of 1000 and 2000 m above ground level.

Trajectory start time dd-mm-yyyy hh	ϵ	Lat/Long	Initial altitude of 2000 m agl	Initial altitude of 1000 m agl
29-05 2013 0900	0.5	Lat	21.16087	21.37168
		Long	29.50668	36.076
29-05 2013 1200	0.5	Lat	20.86322	21.13143
		Long	31.58606	36.4857
29-05 2013 1500	0.5	Lat	20.2287	20.86723
		Long	31.558	36.68876
29-05 2013 1800	0.5	Lat	16.07134	20.4431
		Long	38.3999	37.0137

location (Fig. 6). This trend follows intensity decrease and dispersion of dust plume from event source place, measured with lower number and values of index score.

However, the variograms show very strong Nugget Effect - very high initial values of $\gamma_{st}(h = 0, u = 0)$ [13]. This result is due to noise from measurement error and microscale variation in used dust score data. They are produced mainly because dust score index records are used without any filtering procedure and their values are in the range between 1 to 511. Thus, for model improvements, the functionalities must be expanded further with methods for filtering, robust estimation of particle dispersion and spatial prediction (kriging) [13].

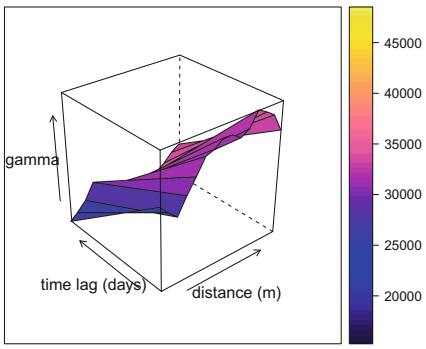

Fig. 5. 3D diurnal spatial-temporal variograms for a model with resolutions $\epsilon = 0.5$ and $\Delta_i = 6$ h for SDE registered on May 29th 2013 1200 UTC. The initial trajectory resolution is 1000 m AGL (left). Backward trajectories are computed for 48 h.

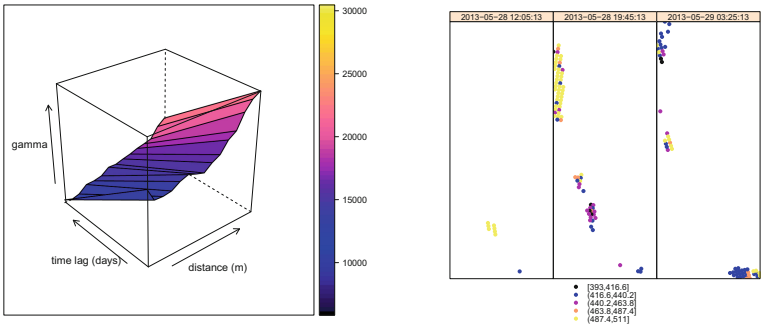

Fig. 6. 3D diurnal spatial-temporal variogram (left) and spatial-temporal distribution of Dust Index (right) for a model with resolutions $\epsilon = 0.5$ and $\Delta_i = 6\,\mathrm{h}$ for SDE registered on May 29th 2013 1200 UTC. The initial trajectory resolution is 2000 m. AGL. Backward trajectories are computed for 48 h.

4 Conclusions

The results in this paper show the basic available functionalities of working model of satellite observations of atmospheric processes. The very simple tests are shown and its validity is confirmed with very well registered Sahara Dust Events ground measurements from BEO observatory in Moussala. However, firstly the model requires more detailed calibration for its complete functionality with all possible other parameters. For this purpose the detailed statistical verification and training with large data, obtained from satellites and ground measured quality data from different transnational sources are needed. Secondly, the model requires additional development for almost real-time functionality and possible prediction features. Having prediction enabled, the system could fix data gaps of irregular satellite scans and enable possible event expectation in area even before its detection.

References

1. Krotkov, N., et al.: Aura OMI observations of regional SO2 and NO2 pollution changes from 2005 to 2015. Atmos. Chem. Phys. **16**, 4605–4629 (2016)
2. Angelov, C., et al.: BEO Moussala a new facility for complex environment studies. In: Sustainable Development in Mountain Regions, pp. 123–139 (2011)
3. Pey, J., et al.: African dust outbreaks over the Mediterranean Basin during 2001–2011: PM10 concentrations, phenomenology and trends, and its relation with synoptic and mesoscale meteorology. Atmos. Chem. Phys. **13**, 1395–1410 (2013)
4. Tchorbadjieff, A., et al.: Sahara dust events over south-western Bulgaria during the late spring of 2013. Comptes rendus de lAcadmie bulgare des Sciences **68**(10), 1229–1234 (2013)
5. R-Project. https://www.r-project.org/
6. Atanassov, E., Gurov, T., Karaivanova, A., Ivanovska, S., Durchova, M., Georgiev, D., Dimitrov, D.: Tuning for scalability on hybrid HPC cluster. In: Mathematics in Industry, pp. 64–77. Cambridge Scholar Publishing (2014)

7. Ostrouchov, G., Chen, W.-C., Schmidt, D., Patel, P.: Programming with Big Data in R (2012). http://r-pbd.org/
8. Stein, A., Draxler, R., Rolph, G., Stunder, J., Cohen, M., Ngan, F.: NOAAs HYSPLIT atmospheric transport and dispersion modeling system. Bull. Am. Meteorol. Soc. **96**(12), 2059–2077 (2015)
9. Carslaw, D.C., Ropkins, K.: Openair-an R package for air quality data analysis. Environ. Model Softw. **27–28**, 52–61 (2012)
10. De Iaco, S., Myers, D.E., Posa, D.: Space-time analysis using a general product-sum model. Stat. Probab. Lett. **52**(1), 21–28 (2001)
11. Graler, B., Pebesma, E., Heuvelink, G.: Spatio-temporal interpolation using gstat. R J. **8**(1), 204–218 (2016)
12. DeSouza-Machado, S., et al.: Infrared retrievals of dust using AIRS: comparisons of optical depths and heights derived for a North African dust storm to other collocated EOS A-Train and surface observations. J. Geophys. Res. **115** (2010)
13. Cressie, N.: Statistics for Spatial Data. Wiley, New York (1993)
14. NASA Earth Data. https://earthdata.nasa.gov

Author Queries

Chapter 24

Query Refs.	Details Required	Author's response
AQ1	Please check and confirm if the inserted details are correct in Refs. [8, 11].	
AQ2	Kindly provide page range for Ref. [12].	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	