

# A geoprocessing model for the selection of populations most affected by diffuse industrial contamination: the case of oil refinery plants

Roberto Pasetto and Marco De Santis

Dipartimento di Ambiente e connessa Prevenzione Primaria, Istituto Superiore di Sanità, Rome, Italy

## Abstract

**Introduction.** A method to select populations living in areas affected by diffuse environmental contamination is presented, with particular regard to oil refineries, in the Italian context. The reasons to use municipality instead of census tract populations for environment and health small-area studies of contaminated sites are discussed.

**Methods.** Populations most affected by diffuse environmental contamination are identified through a geoprocessing model. Data from the national census 2001 were used to estimate census tract level populations. A geodatabase was developed using the municipality and census tract layers provided by the Italian National Bureau of Statistics (ISTAT). The orthophotos of the Italian territory – year 2006 – available on the geographic information systems (GIS) of the National Cartographic Portal, were considered. The area within 2 km from the plant border was used as an operational definition to identify the area at major contamination.

**Results.** The geoprocessing model architecture is presented. The results of its application to the selection of municipality populations in a case study are shown.

**Conclusions.** The application of the proposed geoprocessing model, the availability of long time series of mortality and morbidity data, and a quali-quantitative estimate of contamination over time, could allow an appraisal of the health status of populations affected by oil refinery emissions.

## Key words

- small-area studies
- geographic information systems
- petroleum
- environmental exposure

## INTRODUCTION

In Italy it is possible to plan ecological small-area studies to evaluate environmental risks using routinely collected data at municipality or census tract level. Municipality data are used in several ecological studies [1-3], while census tract data are adopted in small-area analysis in urban areas of some big cities [4-6]; studies at census tract level are suggested to overcome, at least in part, the limits of studies based on municipality data [7].

Environmental contamination from large industrial plants is extremely variable in both qualitative and quantitative terms. The amount of emissions from plants changes in time depending on the amount of raw materials and products, and on the characteristics of work processes (*i.e.* technologies applied to the production cycles and to the contaminants abatement).

Oil refineries (OR) are industrial plants for transformation of crude oil in several petrochemical products; they can be included in complexes together with other petrochemical plants. OR plants have extremely vari-

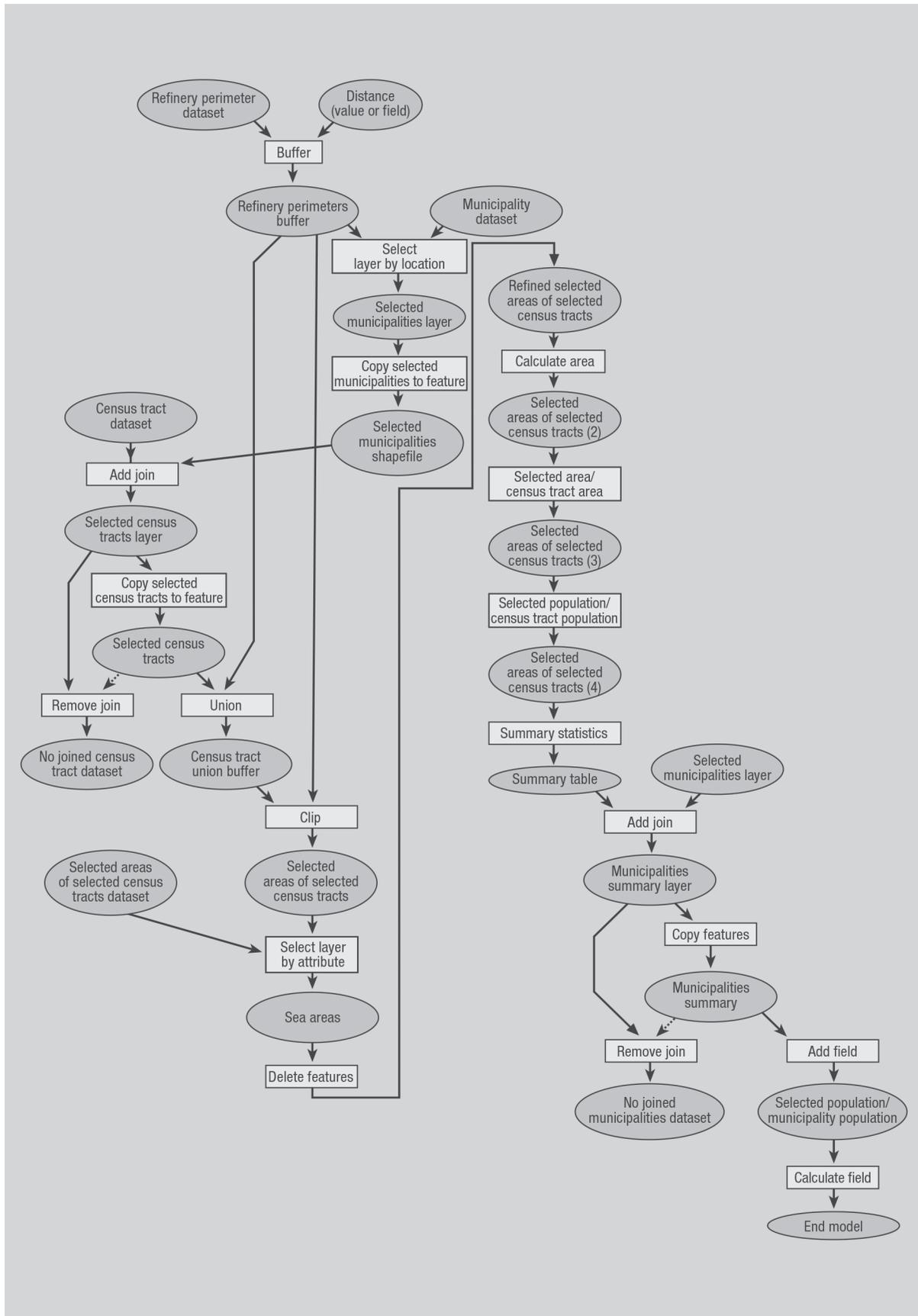
able production cycles and their production can be partly tailored to the market needs [8].

The main environmental issues in the OR sector are linked to the emission of pollutants into the atmosphere, to the production of industrial waste, and to soil and groundwater contamination. Some major problems such as noise, light, smoke (flaring) and odour emissions, are caused by the vicinity of some plants to residential areas [9].

Emissions into the atmosphere are the main contamination from OR [10]. These emissions are continuous in the plants lifetime, except for short interruptions due to some necessary activities, such as maintenance operations. The soil and groundwater contamination is difficult to be estimated, especially its evolution in time.

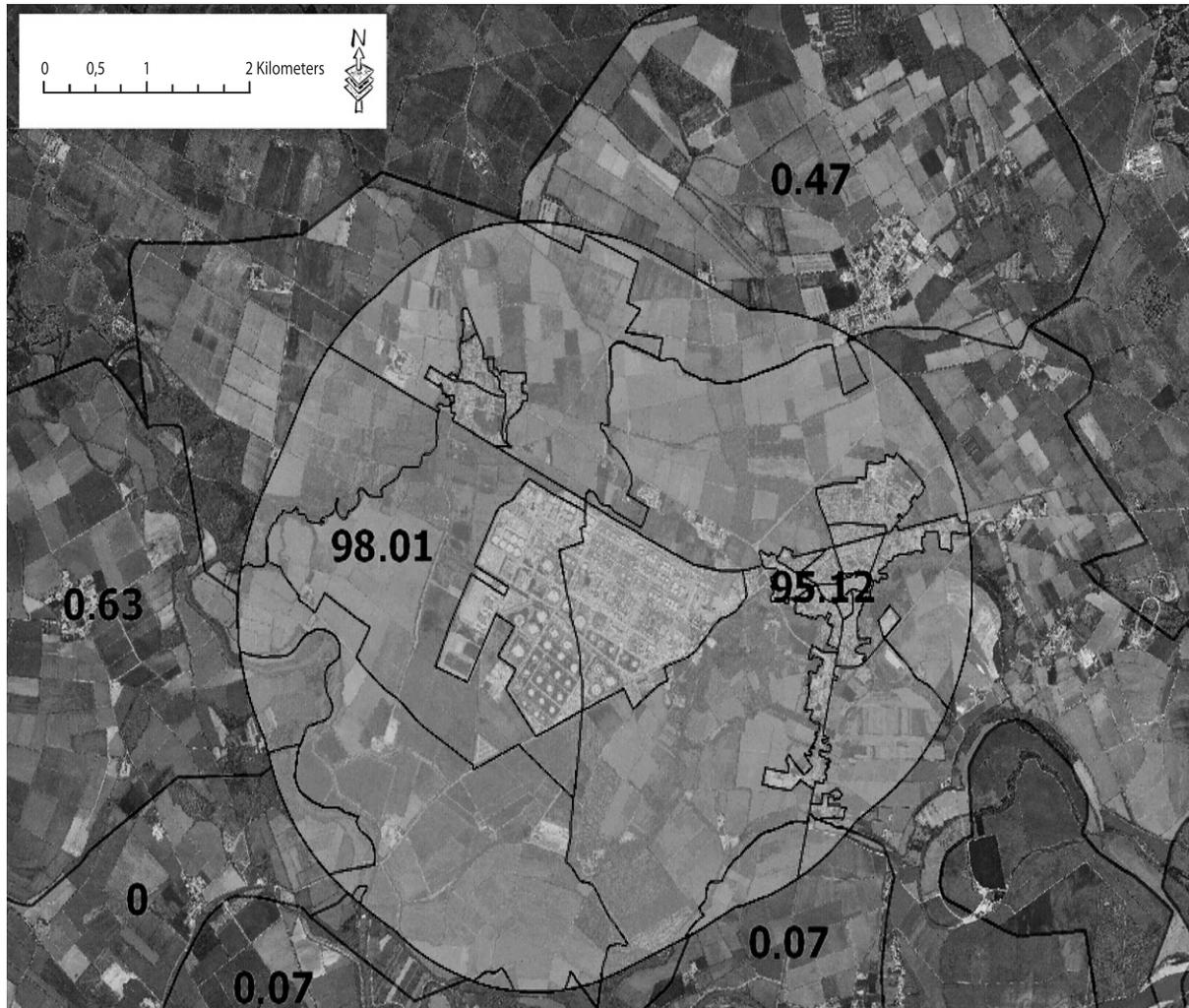
Emissions from OR into the atmosphere can be distinguished in: emissions from chimneys, fugitive emissions, and accidental or maintenance losses. Main air pollutants are: carbon monoxide (CO), carbon dioxide (CO<sub>2</sub>), nitrogen oxides (NO<sub>x</sub>), particulates including metals, sulfur oxides (SO<sub>x</sub>), volatile organic compounds

**Figure 1**  
The geoprocessing model to calculate the proportion of municipality populations within the contaminated area using population data at census tract level.



**Figure 2**

Estimate of municipality populations within the buffer of 2 km from the border of the Sannazaro de' Burgondi oil refinery.



(VOCs) [10]. Exposure to most of these pollutants is associated to severe health effects [11] (<http://www.atsdr.cdc.gov>).

In 2010, the worldwide total number of OR was about 650 and 200 plants were located in Europe [12]. In 2009, Italy was ranked the 14<sup>th</sup> for oil consumption (<http://www.energy.eu/>).

Despite the relevant number of OR, few studies have been performed in order to evaluate their possible health impact in populations residing in their neighborhood. These studies have several limitations, mainly in defining the most affected populations and in the evaluation of exposure over time [13]. The latter issues are the main weak points in evaluating the health risk associated with long term environmental contamination from different industrial sources [7].

The main objective of this contribution is to describe a method to identify municipality populations most affected by diffuse environmental contamination, referring to the case of OR emissions. Furthermore, strengths and weaknesses of a municipality approach vs an approach based on smaller areas are discussed.

## METHODS

At first, a geoprocessing model was implemented, then it was applied to select the populations supposed to be affected by contamination from the Sannazaro de' Burgondi oil refinery (Lombardy Region), that was used as a case study.

Data from the Italian National Census 2001 were used to estimate census tract level populations. The following procedure was implemented in order to define the proportion of municipality populations most affected by OR emissions.

A geodatabase was developed using the municipality and census tract layers provided by the Italian National Bureau of Statistics (ISTAT). The orthophotos of the Italian territory – year 2006 – available on the geographic information systems (GIS) of the National Cartographic Portal, ([http://www.pcn.minambiente.it/GN/progetto\\_scc.php?lan=en](http://www.pcn.minambiente.it/GN/progetto_scc.php?lan=en)) were considered. The distance from the refinery was used as a proxy of area's contamination level; the area within 2 km from the plant border was defined as an operational definition of the area at major contamination.

The geoprocessing model was implemented using the

Model Builder application of the ArcGIS ArcInfo software®. This software enables to: i) model input and output processes in a visual modality; ii) reiterate the developed model; iii) use parametric variables. The model was developed using the main geoprocessing tools, *i.e.* buffer, join, clip, union, spatial and attribute queries.

## RESULTS

The architecture of the GIS model adopted to select municipality populations is shown in *Figure 1*. The model is based on parameters used to identify the spatial diffusion of pollutants (in the case study a buffer of 2 km from the plant border), and the following steps are taken: 1) the census tracts even partly included in the buffer are identified; 2) the proportion of census tract areas included in the buffer is calculated; 3) the total population of each census tract is multiplied by the proportion calculated in the step 2 to estimate the potentially most exposed population for each census tract identified in the step 1; 4) the populations of all census tracts within the same municipality calculated in the step 3 are summed up for each municipality at least in part included in the buffer; 5) for each municipality at least partly included in the buffer, the proportion of the potentially most exposed population is estimated dividing the population obtained in the step 4 by the total population.

The results for the case study of the Sannazaro de' Burgondi refinery are shown in *Figure 2*. Seven municipalities have at least part of their territory included in the buffer. The proportion of population in the area with major contamination is more than 95% for two municipalities, while for the other five municipalities is less than 1%.

## DISCUSSION

The geoprocessing model adopted in the case study, considered the distance from oil refinery as a *proxy* for the area's contamination level, and indirectly for exposure of residing populations. This approach is, however, particularly problematic in the case of a complex contamination, such as the one resulting from emissions into the atmosphere from OR. In fact, in the case of oil refineries, the pollutants dispersion is influenced by multiple factors. The proposed geoprocessing model can be adopted to define the populations affected by contamination, independently from the model of pollutants diffusion/dispersion. On the other hand, the definition of areas most affected by contamination, and the consequent identification of populations most affected by the emissions, depends on the accuracy of the model of pollutants diffusion/dispersion. There are several models used to evaluate the areas affected by the emissions, their implementation and improvement depend on the available information on several parameters, *i.e.* characteristics of emission sources – *e.g.* height, flow rate, composition of emissions, exit temperature; local orography and meteorological conditions [14].

In small-area environment and health studies, the geographical unit of observation (area level) should be selected so that the contamination under study, for example the atmospheric concentration of a given pol-

lutant, equally contributes to the exposure of subjects within the same area/population [15]. It should be underlined that choosing the smallest area level, could not necessarily result in a better capacity of representing exposure-disease associations. Moreover, exposure misclassification could not be greater when using larger rather than smaller areas (*i.e.* larger rather than smaller populations). These considerations apply when the emissions are particularly considerable and the contamination is diffuse over large areas, as in the case of emissions into the atmosphere from OR.

In small-area studies, it should be considered that the possibility to calculate risk estimates depend on the availability of data for both denominators (*i.e.* populations of each area) and numerators (*i.e.* cases that arise from each population). Populations and cases are attributed to a given area on the basis of residence. In the case of a diffuse air pollution, the exposure contribution due to the contamination, not only depends on exposure profile of residential area (*i.e.* house location), but also on those of the locations of daily activity patterns, for example the ones of work/study, or those of recreational and leisure activities. Furthermore, in case of complex industrial contamination, as the one from OR, populations can experience several route of exposure, mainly through inhalation of pollutants emitted into the atmosphere, also through ingestion when contaminants fall out and are accumulated in soil, water and in the food chain [7].

In Italy, the municipality level can be the most appropriate to study, with a small-area approach, the associations between the diffuse contamination resulting from emissions from OR and the health profiles of populations residing in their neighborhood. In fact, municipalities are administrative and social life units, where the main daily activities usually occur and the majority of working and study daily flows take place [16]. Furthermore, routinely collected data are available at municipality level for the whole Italian country since 1980 for mortality, and for the last ten years for morbidity (*i.e.* hospital discharges).

The selected municipalities should be small or medium sized (small-medium populations) to curb the within municipality heterogeneity of risk factors other than the considered environmental contamination. In fact, it was shown that the within municipality variability of health determinants at small-area level, as for example the socioeconomic conditions, increases with the size of the municipality population [17].

Census tracts are the other geographic level used in small-area environment and health studies in Italy. At that level, residential exposure could be more homogeneous, but exposure profiles linked to daily activities are worse represented. Furthermore, even if census tract denominators can be estimated for the whole Italian country, the cases needed to estimate the numerators are available or can be retrieved only for some locations.

It should also be underlined that migration flows, which can distort risk estimates at area level [18], in particular for long period evaluations, are more considerable using census tracts instead of municipalities.

In fact, using census tracts, the bias of migration flows within municipalities is added to that of migration flows between municipalities.

Only a largely diffuse and homogeneous environmental contamination can allow an efficient evaluation at small-area level of the association between a source of contamination and the health profile of populations living in the neighborhood. In order to evaluate the specificity of the associations, it is important that other sources of contamination are not located in the area at study, or that they are adequately accounted for.

Ecological small-area approaches have several well known limitations [15, 19], though in some cases they are the only feasible investigations. Some studies at individual level are applied to retrospectively evaluate the health risk for populations living in contaminated areas [20, 21]. In those cases individual residence history are reconstructed using registry office records. However, carrying out such studies may prove too costly in time and financial resources, and it depends on the availability of individual residential information in an electronic format.

## CONCLUSIONS

In Italy small-area studies at municipality level are used to define health profiles of populations living in the neighborhood of contamination sources. The most recent example is the SENTIERI study (mortality study of residents in Italian polluted sites) [1]. In that study municipalities at major risk were defined on an administrative basis, as municipalities somehow involved in remediation activities for each polluted site were included. That criterion did not necessarily identified the exposed municipalities, which would have been identified if a model of diffusion/dispersion of contaminants was used.

The geoprocessing model proposed in the present paper allows the identification of populations most af-

ected by environmental contamination from OR; it is applicable to other sources of contamination. A novel approach to study OR using small-area data has been recently proposed [22]. It has the potential to provide a quali-quantitative estimate of the different contaminants in time, using data on crude oil consumption, information on refinery cycle, and applying specific emission factors. However, there are some limitations that might hinder a quali-quantitative assessment of the emissions, such as the variability of the combustibles used and the resulting products over time within the same plant and between different plants. For this reason, the quali-quantitative emissions estimate performed on the basis of this approach, should be validated in settings where it is possible to compare results from the predictive models with direct assessments made by environmental monitoring.

On the whole, the proposed approach could allow an appraisal of the health status of populations residing near OR, through: a) quali-quantitative reconstruction of contamination and consequent definition of areas at high environmental risk; b) selection of populations at major exposure applying the proposed geoprocessing model; c) definition of population health profiles using long term series of mortality and morbidity data. Finally, this approach has the potential for methanalysis of data from different OR at national and international level.

## Acknowledgements

The authors wish to thank Ivano Iavarone and Pietro Comba for manuscript revision and useful suggestions; Letizia Sampaolo for linguistic revision.

## Conflict of interest statement

The authors declare no potential conflict of interest.

Received on 30 March 2012.

Accepted on 14 September 2012.

## REFERENCES

1. Pirastu R, Iavarone I, Pasetto R, Zona A, Comba P (Eds). SENTIERI Project – Mortality study of residents in Italian polluted sites: Results. *Epidemiol Prev* 2011;35(Suppl. 4).
2. Marinaccio A, Scarselli A, Binazzi A, Altavista P, Belli S, Mastrantonio M, Pasetto R, Uccelli R, Comba P. Asbestos related diseases in Italy: an integrated approach to identify unexpected professional or environmental exposure risks at municipal level. *Int Arch Occup Environ Health* 2008;81:993-1001. <http://dx.doi.org/10.1007/s00420-007-0293-x>
3. Uccelli R, Binazzi A, Altavista P, Belli S, Comba P, Mastrantonio M, Vanacore N. Geographic distribution of amyotrophic lateral sclerosis through motor neuron disease mortality data. *Eur J Epidemiol* 2007;22:781-90. <http://dx.doi.org/10.1007/s10654-007-9173-7>
4. Federico M, Pirani M, Rashid I, Caranci N, Cirilli C. Cancer incidence in people with residential exposure to a municipal waste incinerator: an ecological study in Modena (Italy), 1991-2005. *Waste Manag* 2010;30(7):1362-70. <http://dx.doi.org/10.1016/j.wasman.2009.06.032>
5. Parodi S, Stagnaro E, Casella C, Puppo A, Daminelli E, Fontana V, Valerio F, Vercelli M. Lung cancer in an urban area in Northern Italy near a coke oven plant. *Lung cancer* 2005;47:155-64. <http://dx.doi.org/10.1016/j.lungcan.2004.06.010>
6. Chellini E, Cherubini M, Chetoni L, Costantini AS, Biggeri A, Vannucchi G. Risk of respiratory cancer around a sewage plant in Prato, Italy. *Arch Environ Health* 2002;57(6):548-53. <http://dx.doi.org/10.1080/00039890209602087>
7. Comba P, Bianchi F, Conti S, Forastiere F, Iavarone I, Martuzzi M, Musmeci L, Pasetto R, Zona A, Pirastu R. SENTIERI Project: discussion and conclusions. *Epidemiol Prev* 2011;35(Suppl. 4):163-71. Italian.
8. Giavarini C. Structures and schemes. In *Oil refining industry: general aspects. Encyclopedia of hydrocarbons. Volume II Refining and petrochemicals*. Roma: Treccani; 2006. p. 3-24. Available from: [http://www.treccani.it/export/sites/default/Portale/sito/altre\\_aree/Tecnologia\\_e\\_Scienze\\_applicate/enciclopedia/inglese/inglese\\_vol\\_2/001-24\\_ING3.pdf](http://www.treccani.it/export/sites/default/Portale/sito/altre_aree/Tecnologia_e_Scienze_applicate/enciclopedia/inglese/inglese_vol_2/001-24_ING3.pdf)
9. Iorio G. Environmental management in refineries. In: *Safety and environmental protection in the refining industry. Encyclopedia of hydrocarbons. Volume II Refining and petrochemicals*. Roma: Treccani; 2006. p. 393-403. Available from: [http://www.treccani.it/export/sites/default/Portale/sito/altre\\_aree/Tecnologia\\_e\\_Scienze\\_applicate/enciclopedia/inglese/inglese\\_vol\\_2/393-404\\_ING3.pdf](http://www.treccani.it/export/sites/default/Portale/sito/altre_aree/Tecnologia_e_Scienze_applicate/enciclopedia/inglese/inglese_vol_2/393-404_ING3.pdf)

10. European Commission. *Integrated Pollution Prevention Control (IPPC). Draft reference document on best available technique for mineral oil and gas refineries*. Draft 1 July 2010. Available from: [http://eippcb.jrc.es/reference/BREF/ref\\_d1\\_0710.pdf](http://eippcb.jrc.es/reference/BREF/ref_d1_0710.pdf).
11. WHO Europe. *Air quality guidelines*. Copenhagen: WHO Regional Office for Europe; 2005. Available from: [http://www.who.int/phe/health\\_topics/outdoorair\\_aqg/en](http://www.who.int/phe/health_topics/outdoorair_aqg/en).
12. True WR, Koottungal L. Global refining capacity advances: US industry faces uncertain future. *Oil & Gas Journal* 2009;107:46-53.
13. Pirastu R, Pasetto R. Review of epidemiological evidence on health effects of residence near petrochemical plants. In: Mudu P, Terracini B, Martuzzi M (Eds). *Human health in areas with industrial contamination*. Copenhagen: WHO Regional Office for Europe; 2012. In press.
14. Isakov V, Touma JS, Burke J, Lobdell DT, Palma T, Rosenbaum A, Ozkaynak H. Combining regional- and local-scale air quality models with exposure models for use in environmental health studies. *J Air Waste Manag Assoc* 2009;59:461-72. <http://dx.doi.org/10.3155/1047-3289.59.4.461>
15. Wakefield J. Ecologic studies revisited. *Ann Rev Public Health* 2008;29:75-90. <http://dx.doi.org/10.1146/annurev.publhealth.29.020907.090821>
16. ISTAT. *Gli spostamenti quotidiani e periodici*. Censimento 2001. Available from: [http://www3.istat.it/salastampa/comunicati/non\\_calendario/20050609\\_00/](http://www3.istat.it/salastampa/comunicati/non_calendario/20050609_00/).
17. Pasetto R, Caranci N, Pirastu R. Deprivation indices in small-area studies of environment and health in Italy. *Epidemiol Prev* 2011;35(Suppl. 4):174-80. Italian.
18. Tong S. Migration bias in ecologic studies. *Eur J Epidemiol* 2000;16:365-9. <http://dx.doi.org/10.1023/A:1007698700119>
19. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. *Annu Rev Public Health* 1995;16:61-81. <http://dx.doi.org/10.1146/annurev.publhealth.16.1.61>
20. Comba P, Bruno C, Fazzo L, Pasetto R, Zona A. Occupational and residential cohorts. In: Mudu P, Terracini B, Martuzzi M (Eds). *Human health in areas with industrial contamination*. Copenhagen: WHO Regional Office for Europe; 2012. In press.
21. Marinaccio A, Belli S, Binazzi A, Scarselli A, Massari S, Bruni A, Conversano M, Crosignani P, Minerba A, Zona A, Comba P. Residential proximity to industrial sites in the area of Taranto (Southern Italy). A case-control cancer incidence study. *Ann Ist Super Sanità* 2011;47(2):192-9. [http://dx.doi.org/10.4415/ANN\\_11\\_02\\_11](http://dx.doi.org/10.4415/ANN_11_02_11)
22. Pasetto R, De Santis M, Sampaolo L, Settimo G. *An approach to study health effects of environmental contamination from oil refineries in a longitudinal perspective*. Abstracts of the 23<sup>rd</sup> Annual Conference of the International Society of Environmental Epidemiology (ISEE). Barcelona, Spain: Environ Health Perspect; 2011 <http://dx.doi.org/10.1289/ehp.isee2011>.