



## Evaluation of roofing materials emissions at the city scale: Statistical approach for computing roofing area distribution

Emna Sellami-Kaaniche, Bernard De Gouvello, Marie-Christine Gromaire, Ghassan Chebbo

### ► To cite this version:

Emna Sellami-Kaaniche, Bernard De Gouvello, Marie-Christine Gromaire, Ghassan Chebbo. Evaluation of roofing materials emissions at the city scale: Statistical approach for computing roofing area distribution. NOVATECH, Jun 2013, Lyon, France. 2013. <hal-00965712>

**HAL Id: hal-00965712**

**<https://hal-enpc.archives-ouvertes.fr/hal-00965712>**

Submitted on 25 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Evaluation of roofing materials emissions at the city scale: Statistical approach for computing roofing area distribution

Evaluation des émissions des matériaux de couverture à l'échelle de la ville : approche statistique de calcul de la répartition des surfaces des toitures

Emna SELLAMI-KAANICHE<sup>1, 2</sup>, Bernard DE GOUELLO<sup>1, 2</sup>, Marie-Christine GROMAIRE<sup>2</sup> et Ghassan CHEBBO<sup>2</sup>

<sup>1</sup> Centre Scientifique et Technique du Bâtiment (CSTB), 11 rue Henri Picherit, BP 82341 44323 Nantes Cedex, France

<sup>2</sup> Université Paris-Est, LEESU (UMR MA 102), UPEC, UPEMLV, ENPC, AGROPARISTECH, F77455 ou F94010, 6 - 8 avenue Blaise Pascal, Cité Descartes - Champs sur Marne, MARNE LA VALLEE CEDEX 2, France  
e-mail: [sellamie@leesu.enpc.fr](mailto:sellamie@leesu.enpc.fr); [bernard.degouvello@leesu.enpc.fr](mailto:bernard.degouvello@leesu.enpc.fr); [gromaire@leesu.enpc.fr](mailto:gromaire@leesu.enpc.fr); [chebbo@leesu.enpc.fr](mailto:chebbo@leesu.enpc.fr)

### RÉSUMÉ

Aujourd'hui, les matériaux de couverture sont considérés comme une source importante de contamination des eaux de ruissellement urbaines. Dans le cadre de la directive européenne sur l'eau (2000/60 CE), une évaluation des flux de contaminants issus des ruissellements des toitures est donc nécessaire à l'échelle de la ville. Or, en passant à l'échelle de la ville, on se confronte à la réalité d'une diversité très importante et difficile à quantifier les matériaux de couverture. Le but de ce papier est de résoudre la problématique « comment quantifier les matériaux de couverture à l'échelle de la ville ? ». Pour ce fait, nous avons développé une approche statistique basée sur deux étapes. La première étape utilise la théorie de l'échantillonnage qui a pour but de définir la bonne taille de l'échantillon à tirer qui représente la proportion  $R$  (%) de la surface totale de toute zone urbaine à étudier. Pour déterminer la valeur de  $R$  (%), nous avons calculé et analysé les erreurs absolues  $E_a$ . Cette analyse montre qu'un tirage avec remise de  $R=4\%$  de la zone d'étude donne de bons résultats. La deuxième étape consiste à appliquer la théorie d'estimation. En effet, après avoir déterminé la proportion  $R$  (%) d'échantillonnage, nous allons estimer les caractéristiques inconnues d'une autre zone en tirant un échantillon représentant  $R(\%)$  de la surface totale de la zone étudiée. La validation de la taille d'échantillonnage a été effectuée sur deux zones de caractéristiques différentes.

### ABSTRACT

Roofing materials are considered as a major source of urban runoff contamination. Today, in the context of the European Water Directive (2000/60 CE), an accurate evaluation of contaminant flows from roofs is thus required at the city scale, and therefore the development of assessment tools is needed. However, at the city scale, the important diversity of roofing materials represents a difficult task. In fact, the most restrictive step when estimating contaminant flows for large-scale territories is the estimation of the roofing materials surfaces area, for which no direct data exists. This paper aims to describe the proposal statistical approach for computing roofing materials area distribution. This method is based on two steps. The first one is the sampling theory which aims to define the required size of the selected sample which represents the proportion  $R$  (%) from any total urban area surface. The second step is based on applying the estimation theory. In fact, after determining the proportion  $R$  (%), we estimate the unknown distribution of roofing materials in an urban area by performing a sample representing  $R$  (%) from the total urban area surface. The reliability of the methodology has been proved for tow different urban areas with known roofing materials distribution. Results obtained showed that a sample representing  $R=4\%$  of the zone area should be randomly generated.

### KEYWORDS

City scale, Representative sample, Roofing material, Statistical approach

## 1 INTRODUCTION

Roofing materials are considered as a major source of urban runoff metal contamination. This observation was revealed by several research programs conducted since the 1990s (Förster, 1996; Gromaire-Mertz *et al.*, 1999; Odnevall Wallinder, 1998). Further undergoing research aims at extending runoff rates at the city scale (Robert-Sainte, 2009; Gromaire *et al.*, 2011). The roof area estimation is a fundamental input for evaluating roofing materials emissions. However, at the city scale, the information concerning roofing materials and their surfaces on an urban district does not exist currently in urban data banks. Some methods have already been developed for the evaluation of roof surface on a large-scale, in order to study building-integrated solar-energy applications (Izquierdo *et al.*, 2008; Bergamasco and Asinari, 2011) or to explore the potential of green infrastructure in adapting cities for climate change (Gill *et al.*, 2006). But, in these studies, the roofing materials were not taken into account. Other approaches (Le Bris *et al.*, 2009; Gromaire *et al.* 2002) have evaluated roofing materials surfaces using data obtained from aerial photographs and image classification software. The classification method based on aerial images was applied to an urban catchment with 2.25 km<sup>2</sup> of surface (Le Bris *et al.*, 2009). The obtained results show about 75% of well classified roofing surfaces. Nevertheless, this classification method presents some limitations especially in terms of confusion between different classes (eg: zinc in the shadow and slates in the sun). On the other hand, in order to apply this method, we need a high resolution image which requires an expensive cost and it is not usually available in every study site. Therefore, a method must be developed to estimate roofing materials distribution at the city scale. Ideally, a method to calculate the roofing materials surfaces should: (a) be accurate; (b) be reliable, with the possibility of computing or bounding the error of roofing materials areas estimation; (c) be inexpensive (low cost); (d) require few, global, available and standard input data; (f) produce geo-referenced results; (g) be scalable from local to global scales.

Given the size of the city, a complete census of the roofing material surfaces was unfeasible. Instead, a stratified random sampling technique was employed in conjunction with aerial photograph interpretation of surface roofing materials types. It is preferable to divide the city into distinctive homogenous strata in order to accurately characterize it. Stratified random sampling ensures that results can be extrapolated to the city level and is suitable for a heterogeneous population. Dividing a city into fairly homogeneous units reduces the variance of the estimates, thereby leading to more precise results (Nowak *et al.*, 2003).

Several considerations must be made in order to compute the roofing materials surfaces area. In fact, the choice of roofing material is extremely dependent on the history of the building (age, renewal), the urban planning (land use, building typology...), the social characteristics, the town regulation framework... For instance, in a previous work (Sellami-Kaaniche *et al.*, 2012), a set of rules have been developed for identifying the use of a roofing material at the city scale and have applied these rules to Créteil city. New maps of this city were then obtained from which homogeneous strata have been defined for identifying the distribution of roofing materials using a statistical approach.

In this work, a methodology based on a statistical approach to compute an estimation of roofing materials area distribution for a stratum is described. The methodology is applied to Créteil city (France). The data are processed with the aid of a Geographical Information System (QGIS) and Matlab was used to implement the statistical approach.

The remainder of this paper is organized as follows. In section 2, we describe the study site and the databases used. Then the developed methodology for computing roofing materials data is addressed in section 3. In section 4, experimental results are presented and discussed, and finally some conclusions are drawn in section 5.

## 2 CASE STUDY SITE

### 2.1 Urban Characterization

The goal of our research work is to develop a general method to quantify roofing material surfaces at the city scale. In this context we need to choose a case study to achieve two objectives. The first one is to validate the methodology. The second one is to make it possible to apply the methodology to other cities. Therefore we need to choose a complex site in which we try to have the different aspects of the city. In fact, the city should present a sufficient urban diversity. The selected city is Créteil (Department 94) located about 10 km from Paris (France). Créteil has 89 304 inhabitants (INSEE,

2008) distributed over 11.5 km<sup>2</sup> which represents a reasonable size for our research. This city is divided into four major historical urban areas: Old center, Mont Mesly, New Créteil I, New Créteil II. Each area presents a specific urban organization which depends especially on the period of construction (PLU, 2010). The recent zone is New Créteil II. Figure 1.b shows the interesting urban and functional diversity of Créteil. In fact, we can see different land use represented by different color: industrial activities located nearby, health equipment, different types of building (apartment, individual houses). The land use is not homogeneous in the different urban areas of the city. For example the Old center mainly consists of individual habitat. However Mont Mesly is composed of collective habitat. This distribution is mainly due to historical factors. Finally, Créteil is easily accessible from the laboratory.

## 2.2 Databases

Créteil city presents a big diversity and a large number of buildings which represents an area of 6.18 km<sup>2</sup>, about 54% of the city area. To quantify the different roofing materials, urban data banks (cadastre), land use database *MOS-IAU*<sup>1</sup> (IAU-IDF, 2008) have been consulted and interviews have been conducted with various actors and stakeholders (eg: master of work, contracting authority, architect). These contacts showed that informations relating roofing materials and their surfaces on an urban district does not exist currently in urban data banks. Therefore, in a previous work (Sellami-Kaaniche et al, 2012), we have developed a set of rules for identifying the use of a roofing material at the city scale. These rules have been applied to Créteil city. In this work, we have associated different databases for Créteil: land use database *MOS-IAU*, images database *BD-topo*<sup>2</sup> and *BD-ortho*<sup>3</sup> (IGN, 2008). The obtained results correspond to two maps of Créteil city (see Figure 1).

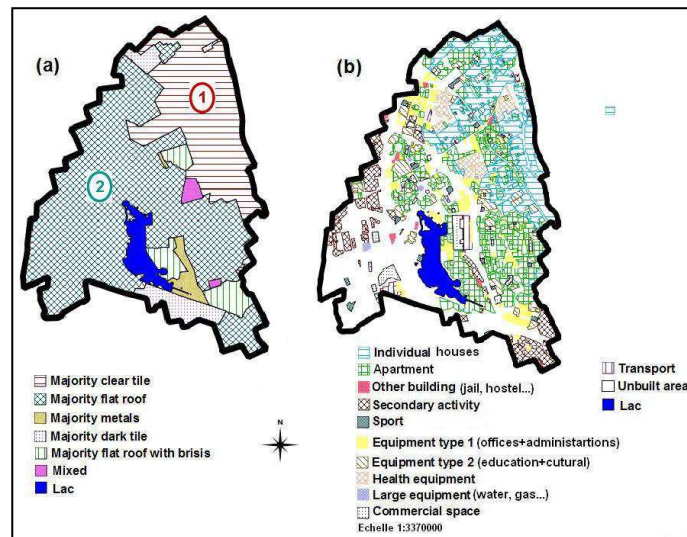


Figure 1 : (a) Material distribution map (b) Building class map

The first map represents the roofing material distribution (see Figure 1.a). Each color represents an area with the same material distribution. The second one is a building class map (see Figure 2.b) established from reorganizing the *MOS-IAU* database. Each color represents a building class. These maps are the basis of the statistical approach.

## 2.3 Issue

By overlaying these two maps (Figure 1), we consider that each building class in Figure 1.b located in an area of the Figure 1.a is a homogeneous urban area in terms of building types and the historical location. We define homogeneous urban area as an area representing specific probabilities of roofing

<sup>1</sup> IAU: Institut d'Aménagement et d'Urbanisme d'Ile-de-France: institute of planning and development for the Greater Paris region.

<sup>2</sup> BD-Topo is a database for Topographic Information IGN roughly corresponding to the traditional content of the map at 1:25,000, with a much higher precision of the order of 1:10000 or 1:5000, made from an aerial image. These images define only the built by masking inbuilt areas (roads, gardens ...).

<sup>3</sup> BD-ortho images come from IGN's (Institut Géographique National) which contains digital colour ortho-photos with three or four (red-green-blue-near infrared) bands and with a 50 cm ground resolution

material distribution. In this context, the question that arises is: "How to evaluate the roofing materials distribution for each homogeneous urban area?"

### 3 METHOD FOR ROOFING MATERIALS AREA ESTIMATION, AND APPLICATION TO CRÉTEIL

In this section, we focus on describing the method of evaluating the roofing materials distribution for a stratum. We define a stratum as a building class in Figure 1.b located in an area in Figure 1.a. For each stratum we are looking to evaluate the proportion  $p_i$  of each material  $i$ . The proposed methodology is explained through its application to Créteil. To solve the previous problem, we propose to apply an inferential statistic which aims to extend the observed sample properties to the entire population and to validate or invalidate hypotheses made *a priori* or after an exploratory phase (Saporta, 1990).

The method developed is based on two steps. The first step is the sampling theory which aims at determining the number of samples  $n_r$  which is required for all existing materials in a stratum. This step involves taking a stratum with known characteristics area ( $N$  (m<sup>2</sup>)), proportion of different existing materials ( $p_i$  (%)), which allows us to make the sampling. To make a choice of an optimized sample size, absolute error  $E_a$  has been computed. The second step is the application of the estimation theory. Indeed, after determining the number of required samples  $n_r$ , we determine the proportion  $R = n_r/N$  and then we estimate the unknown characteristics of a population by selecting a sample representing  $R$  (%) from the total population area. To validate this proportion  $R$  (%) we have tested two other stratum with different characteristics.

In what follows, we will use the different symbols defined in table 1.

Table 1 : Symbols definition

Symbol	Definition	Symbol	Definition
$N$ (m <sup>2</sup> )	zone 1 surface area	$N_b$	Number of building in zone 1
$n_r$ (m <sup>2</sup> )	Sample size required in zone 1	$R = n_r/N$ (%)	proportion of the required sample
$n$ (m <sup>2</sup> )	sample size tested	$p_i$ (%)	real proportion of material $i$ in zone 1
$k$	the number of draws from a population	$p_{e,i}$ (%)	proportion of material $i$ in zone 1 in the sample $e$
$C_b$ (m <sup>2</sup> )	cumulative area	CI (%)	confidence interval
$E_a$ (%)	absolute error	$r$	a value selected in an uniform random variable between 1 and $C_b$

#### 3.1 Sampling Strategy-Monte Carlo simulation

The chosen stratum is zone 1: the individual houses class (Figure 1.b) located in area (1) (Figure 1.a). For this zone we have associated, with QGIS, the different previous maps with BD-topo and BD-ortho images and we have obtained a database  $T$  of existing buildings (numbered  $1 \dots N_b$ ) and their roofing materials (see Table 2). For each building, roofing material has been identified by looking to the ortho-photo. From this database we have identified the roofing materials distribution for zone 1 (see Table 3).

To study the impact of sampling on the roofing materials proportions uncertainty, we have used a Monte Carlo simulation. To perform this study, samples should be representative. In fact, it is necessary that each element of the zone has an equal chance of belonging to this sample. Thus, a random sampling technique of independent individuals of 1 m<sup>2</sup> from a population composed of the total roofs surfaces areas ( $N$  (m<sup>2</sup>)) was employed. In this case, the observations become random variables following a probability law that will be identified (Saporta, 1990). The random variable is the roofing material. Each sample has a statistical parameter: the proportion  $p_{e,i}$  of each material  $i$  for which related statistical law has to be defined.

To proceed, a number of samples  $n$  (individuals of 1 m<sup>2</sup>) (where  $n \leq N$ ) was randomly generated among the surface  $N$  (m<sup>2</sup>). To generate each sample  $n$ , a variable  $r$  between 1 and  $C_b$  (see Table 1) has been generated (with and without replacement, the final choice of the draw strategy is defined in

section 4.3). Then, in table 1, for each value  $r$  we seek (in the column of cumulative area) for the first surface greater or equal to  $r$  and we identify the roofing materials corresponding to that surface. Finally, we obtain the proportion for each material for each sample. The process is repeated  $k$  times for each value of  $n$  (the choice of  $k$  value is defined in section 4.1).

For each material, we plot the different statistics of its distribution (the mean, maximum and minimum values, confidence interval at level of 95% computed using the method of percentiles) and for the two draws (repeated  $k$  times) with and without replacement. We also plot the actual distribution for each material to be compared with the simulation results. In addition, we have verified, using the Shapiro-Wilk normality test at the 5% threshold, that for  $n \geq 2$ , the normal distribution of each material is readily accepted. Finally, we have determined from the different value of  $n$  the number of samples  $n_r$  which is required for all roofing materials. This was being done for each material.

Table 2: Urban database for zone 1

	Building	Area (m <sup>2</sup> )	Cumulative area (m <sup>2</sup> )	Roofing material ( $i$ )
I	1	$S_1=100$	$C_1=100$	Clear tile
	2	$S_2=40$	$C_2=S_1+ S_2=140$	Slates
	...	...	...	...
	$N_b$	$S_{N_b}$	$C_b = N$	

Table 3 :

roofing materials distribution for zone 1

The real

Material	Area (m <sup>2</sup> )	Proportion $p_i$ (%)
Clear tile	2107.05	33.68
Dark tile	1228.71	19.64
Slates	889.67	14.22
Zinc	833.95	13.33
Flat roof	983.38	15.72
Steel	116.58	1.86
Other material	97.50	1.56
<b>Total (N)</b>	<b>6256.84</b>	<b>100.00</b>

### 3.2 Estimation

The objective of this section is to describe the method of estimating roofing materials distribution ( $p_i$ : proportion of material  $i$ ) in a given zone by sampling (see Figure 2). For this purpose, we need to apply the Theorem for Proportions (Saporta, 1990) in which the sampling distribution for samples (with size  $n \geq 30$  and  $p_{e,i}$ : proportion of material  $i$  in the sample  $e$ ) is approximately normal, so :

$$p_{e,i} \rightarrow N\left(p_i; \sigma_{p_i} = \sqrt{\frac{p_i(1-p_i)}{n}}\right)$$

The point estimate of  $p_i$  is  $p_{e,i}$  and an estimation of standard error is given by:

$$\sigma_{p_i} = \sqrt{\frac{p_{e,i}(1-p_{e,i})}{n}}$$

A confidence-level coefficient of  $\alpha = 0.05$  has been chosen and therefore:

$$p_i \in [p_{e,i} - 1,96 \cdot \sigma_{p,i}, p_{e,i} + 1,96 \cdot \sigma_{p,i}]$$

Before applying this theorem in a given zone to estimate its roofing material distribution, we need to

choose the number of samples  $n_r$  and the draw strategy. Thus, we have compared the different statistics computed for each draw (with and without replacement) with statistics related to the Theorem for Proportions. In fact, we have applied this theorem for zone 1 with the same previous conditions.

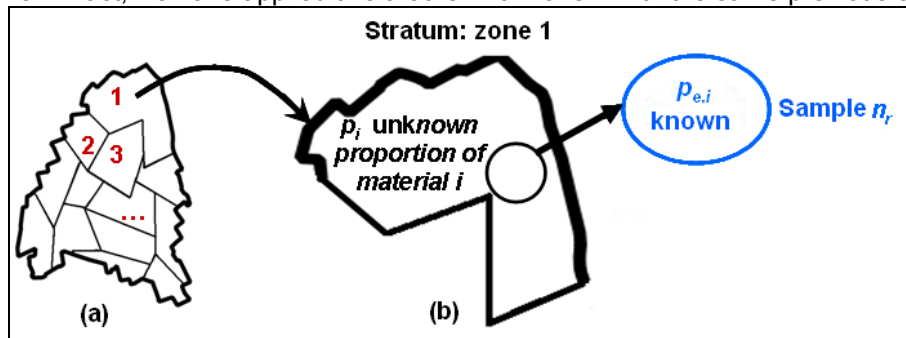


Figure 2 : (a) Créteil city divided into homogeneous strata; (b) Estimation strategy for a stratum

## 4 RESULTS AND DISCUSSIONS

### 4.1 The number of draws

To choose the required  $k$  value for each value  $n$ , we have tested some arbitrary values of  $k$  for different values of  $n$ . The illustration was made for clear tile material as an example (see Figure 3). In this Figure, from  $k=1000$  upwards, the different computed statistics (mean value, confidence interval) become stable. In what follows, we have kept  $k=1000$ .

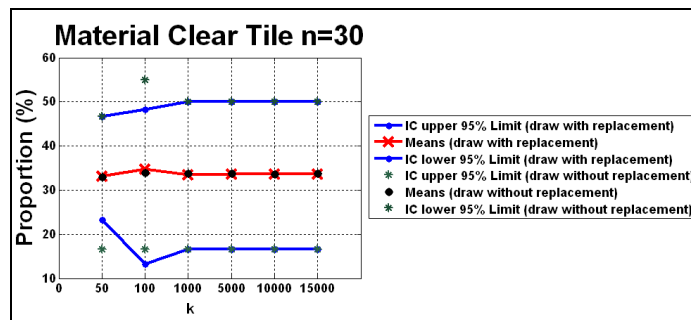


Figure 3 :  $k$  value variability- case of Clear tile and for  $n = 30$

### 4.2 Analysis of sampling simulations

Sampling simulations with and without replacement were applied. In Figure 4, confidence intervals, the mean, maximum and minimum value and the actual proportion for clear tile material are illustrated.

As we can see in Figure 4, the mean computed over 1000 draws (with and draw without replacement) for each sample  $n$  coincides with the actual proportion of the clear tile. This observation confirms our choice of  $k=1000$  and shows that the sampling strategy represents the reality of the roofing material distribution in zone1. This result was validated for all roofing materials (slates, steel, dark tile...).

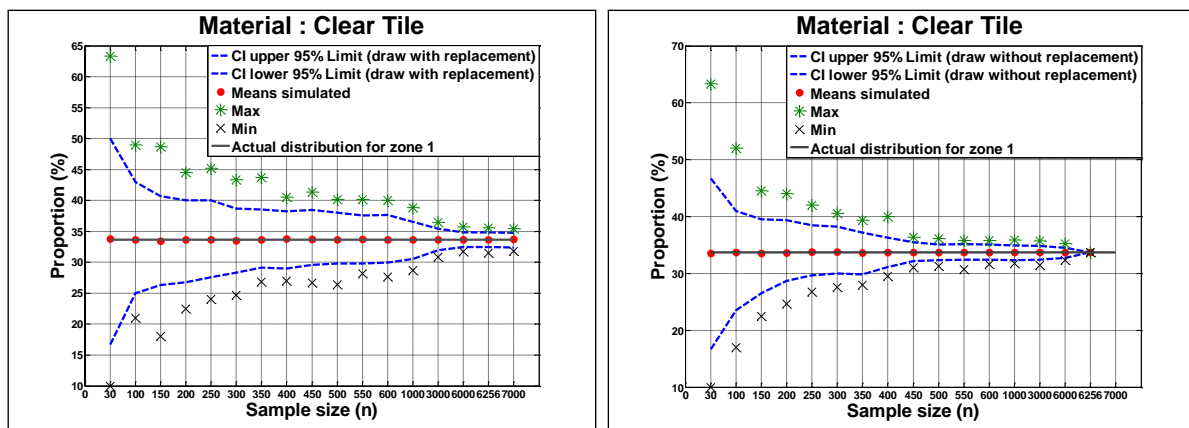


Figure 4 : Computing statistics for the clear tile material for the two draws strategies -Zone1-

### 4.3 Sampling strategy choice

The comparison of the different statistics computed for each draw (with and without replacement) with statistics related to the Theorem for Proportions, shows that the mean values computed from the theorem coincide with those calculated for each draw. The illustration was made for the slates material as an example (see Figure 5). Theoretically, the draw with replacement is used to ensure that each independent element of the population has an equal chance of belonging to the sample. Figure 5 shows that the sampling simulation results obtained from the draw with replacement (dotted red line (-)) coincide with those of the Theorem for Proportions (solid blue line). In addition, for  $n \leq 250$ , limits of the confidence interval for the two draws (with and without replacement) coincide with those of the Theorem for Proportion. However, for  $n > 250$  the limits of the confidence interval for the draw without replacement (green point in Figure 5) do not coincide with those of the theorem. Therefore, draw without replacement could be used only for  $n \ll N$ .

In what follows, for the estimation method, the draw with replacement will be elaborated.

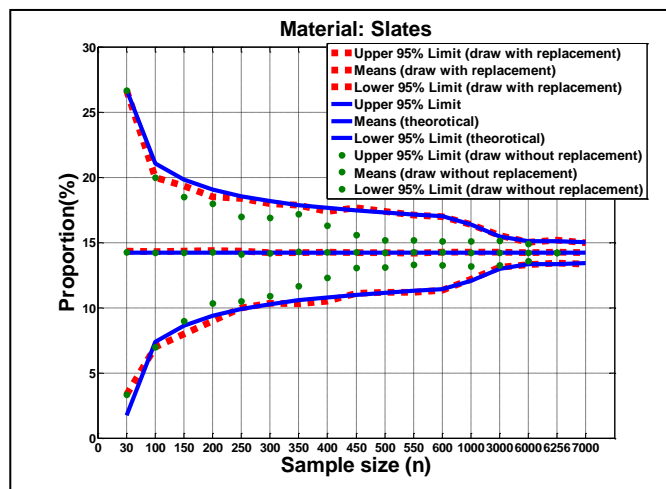


Figure 5 : Comparison computed statistics for draws with replacement and without replacement with computed statistics for the Theorem for Proportions - Slates case -

### 4.4 The sample number choice

The objective of this section is to determine the number of samples  $n_r$  which is required for all roofing materials. Table 4 shows that sample size greatly varies from one material to another for the same confidence interval. In fact, to obtain a confidence level of  $\pm 20\%$  for clear tile, we need a sample of 182 m<sup>2</sup>, whereas for steel, we need  $n=5250$  m<sup>2</sup>. This large difference can be explained by the actual roofing material distribution in zone 1. Indeed, table 3 shows that clear tile represents 33.7 % of the zone 1 area while steel represents only 2%.

In this context, absolute error  $E_a$  was computed for each material. This error is the result of the multiplication of confidence interval obtained for each  $n$  with the real proportion of the material  $i$ .

Table 4 : Computed confidence interval for each material for different sample size value

Material Sample size (n) (m <sup>2</sup> )	CI confidence interval ( $\pm$ ) (%)						
	Clear Tile	Dark Tile	Slates	Zinc	Flat roof	Steel	Other material
100	28	42	40	43	44	111	168
300	15	23	26	30	23	75	76
600	12	15	19	18	17	60	60
3000	5	6	9	8	9	26	28
6000	4	5	6	6	6	17	20



Results are obtained in Figure 6. Figure 6 shows the evolution of the absolute error  $E_a$  with the sample size for each roofing material. The analysis of the shape of the different curves shows that the zone Z is a critical zone for the  $E_a$  evolution. In fact, for the left hand side of the zone Z ( $n \leq n^-$ ),  $E_a$  fluctuations become much more important when fewer samples have been taken. However, for the right hand side of the zone Z ( $n \geq n^+$ ),  $E_a$  has progressively stabilized. We can observe that when  $E_a$  is very low, the roofing material proportions are closer to the real distribution. Therefore, the range of acceptable sample size is the range  $n^-$  for  $n^+$  corresponding to a range of absolute errors between 1% and 7% for all roofing materials. From this range, we choose  $n_r=250$  m<sup>2</sup> (a value in the middle of zone Z) as the required sampling size. In fact, for  $n$  values close to  $n^-$ , absolute error is very high and for  $n$  close to  $n^+$ , the sampling size becomes important. In addition, we look to optimize this size. Thus, we define  $n_r$ , representing  $R=4\%$  of the zone area as the size sampling needed. For this size we get a confidence level  $\leq 30\%$  for the roofing materials with high densities but absolute error low values (see Table 5).

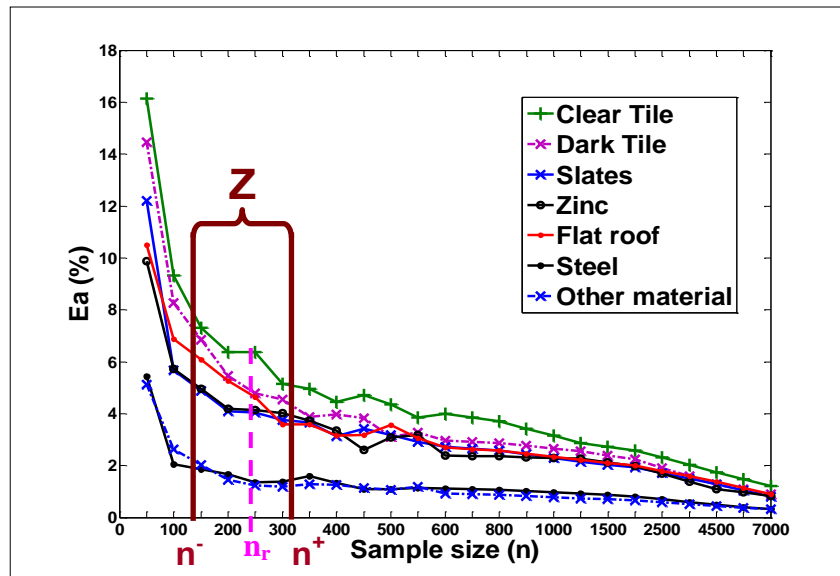


Figure 6 :  $E_a$  variability for different samples sizes for each material

Table 5 :  $E_a$  value for  $n_r=250$

$n=250$ m <sup>2</sup>	Clear Tile	Dark Tile	Slates	Zinc	Flat roof	Steel	Other material
$E_a$ (%)	6.37	4.79	4.05	4.14	4.65	1.36	1.24

#### 4.5 Estimation and validation of the sampling size

In this section, we firstly describe the operational method for roofing material distribution estimation in an unknown area. Then, we will validate the choice of sampling size by estimating roofing material distribution in two new stratum (zones 2 and 3). In each stratum, we have identified the real roofing material distribution as described in section 3.1 for zone 1. For each zone, a random sample representing  $R=4\%$  of the zone area was generated with replacement. The sampling strategy is based on randomly generating independent individuals of 1 m<sup>2</sup> representing 4% of the stratum roofs area ( $N$  (m<sup>2</sup>)). Each 1 m<sup>2</sup> belongs to a specific building with specific roofing material. So, we obtain two databases, the first one is a database of different buildings (with their roof surface) existing in the stratum, and the second one is composed of the selected surfaces for each building in the sample. By associating BD-topo and BD-ortho we identify its roofing material. Therefore, we compute for each material  $i$ :

- in the stratum (area ( $N$  (m<sup>2</sup>)): surface area of each material  $i$  and the proportion of different existing materials ( $p_i$  (%)),
- in the sample  $e$  representing 4% of the stratum area ( $N$ ) : surface area of each material  $i$  and the proportion of materials ( $p_{e,i}$  (%))

Finally, we have computed the confidence interval, and the absolute error values (see table 6 and 7). The two zones 2 and 3 are described as follows:

- zone 2 ( $N_2=81100.11 \text{ m}^2$ ): apartment class (Figure1.b) located in area **(2)** (Figure1.a)
- zone 3 ( $N_3=105304.04 \text{ m}^2$ ): secondary activity class (Figure 1.b) located in area **(2)** (Figure1.a)

Table 6 and 7 show that  $E_a$  absolute error values are low for the two zones. It is also less than 1% for all roofing material. The computed confidence level is very low for roofing material distribution with high density, this is the case of flat roof with 92% of real proportion in zone 2 and 87% in zone 3, we get a confidence level  $\leq 1\%$  so a high precision. For material with confidence level  $\geq 30\%$ ,  $E_a$  is still low and their real proportions do not exceed 2% in the whole zone 2 or 3. This indicates that the results gained through this method are truly representative and can be used with confidence.

Table 6 : Results obtained for zone 2

Material	Real roofing material distribution (%)	Material proportion (%) for n = 3244 m <sup>2</sup> (4% of zone 2 area)	CI confidence interval ( $\pm$ ) (%)	$E_a$ (%)
Dark Tile	5,58	5,52	<b>11,48</b>	<b>0,641</b>
Slates	1,18	1,23	<b>33,28</b>	<b>0,393</b>
Flat roof	91,95	91,95	<b>0,92</b>	<b>0,842</b>
Steel	0,69	0,52	<b>8,16</b>	<b>0,056</b>
Other material	0,6	0,77	<b>73,30</b>	<b>0,440</b>

Table 7 : Results obtained for zone 3

Material	Real roofing material distribution (%)	Material proportion (%) for n = 4213 m <sup>2</sup> (4% of zone 3 area)	CI confidence interval ( $\pm$ ) (%)	$E_a$ (%)
Zinc	1,07	1,31	<b>38,427</b>	<b>0,4127</b>
Flat roof	86,50	86,31	<b>1,002</b>	<b>0,8668</b>
Steel	8,78	8,64	<b>12,859</b>	<b>1,1291</b>
Other material	3,65	3,74	<b>10,171</b>	<b>0,3709</b>

## 5 CONCLUSIONS

In this paper we have developed a statistical method to quantify roofing materials distribution at the city scale. Given the size of the city, a stratified random sampling technique was proposed. We have divided the city into distinctive homogenous strata. In this paper, we have focused on describing the statistical approach to evaluate roofing materials distribution for a stratum. This method is based on two steps. The first one is the sampling theory which aims to determine the number of samples  $n_r$  required for all roofing materials. The second step is the application of the estimation theory. Indeed, after determining the number of samples  $n_r$ , we determine the proportion  $R = n_r/N$  and then we estimate the unknown characteristics of a stratum by generating a sample representing  $R$  (%) from the total area of the new stratum. To validate this proportion  $R$  we have tested two other strata with different characteristics. The obtained results show that a random sample representing  $R=4\%$  of the zone area could be drawn from the available population with and without replacement. In fact, we consider that a sample representing 4% of the whole stratum area is very small, thus the draw without replacement could be used. Then, the estimated roofing material distribution of the zone will be established by applying theorem for proportions. Thus, we will obtain an estimation of each roofing material proportion by computing the confidence interval.

The object of our future study is to evaluate the roofing materials distribution for the entire city. Firstly, the city should be divided into fairly homogenous urban areas, and each urban area represents a stratum. Then, for each stratum (with known area) a sample representing  $R=4\%$  of the stratum area is randomly generated. After that, using GIS software we associate different maps BD-topo, BD-ortho and the map of strata to identify the roofing material sample generated for each stratum. Finally we obtain the roofing materials proportions and their confidence intervals.

## ACKNOWLEDGEMENTS

This study has been conducted as part of the OPUR research program. The authors gratefully acknowledge the Centre Scientifique et Technique du Bâtiment (CSTB) for their financial support and Ali Hannouche for his helpful suggestions in the statistical approach.

## LIST OF REFERENCES

- Bergamasco L., Asinari P. (2011) *Scalable methodology for the photovoltaic solar energy potential assessment based on available roof surface area: Application to Piedmont Region (Italy)*. Solar Energy 85:1041-1055
- Förster J., (1996). - *Patterns of roof runoff contamination and their potential implications on practice and regulation of treatment and local infiltration*. Water, Science and Technology, 33 (6): 39 - 48.
- Gill S.E., Handley J.F., Ennos A.R., Pauleit A.R. (2006) *Adapting Cities for Climate Change: The Role of the Green Infrastructure*. Built environment 33(1):115-133
- Gromaire-Mertz, M., Garnaud, S., Gonzalez, A. et Chebbo, G., (1999). - *Characterisation of urban runoff pollution in paris*. Water, Science and Technology, 39 (2): 1-8.
- Gromaire, M. C., Chebbo, G. et Constant, A. (2002). -*Incidence of zinc roofing on urban runoff pollutant loads: The case of Paris*. Water, Science and Technology 45(7): 113 - 122.
- Gromaire M.C., Robert-Sainte P., Bressy A., Saad M., de Gouvello B. et Chebbo G., (2011). - *Zn and Pb emissions from roofing materials-Modelling and mass balance attempt at the scale of a small urban catchment*. Water Science and Technology, 63 (11): 2590-2597.
- IAU-IDF. <http://www.iaurif.org/>
- INSEE (Institut national de la statistique et des études économiques) (2008): <http://www.recensement-2008.insee.fr/chiffresCles.action?zoneSearchField=CRETEIL&codeZone=94028-COM&idTheme=3>
- Izquierdo S., Rodrigues M., Fueyo N. (2008). *A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations*. Solar Energy 82:929-939
- Le Bris A. et Robert-Sainte P. (2009). - *Classification of roof materials for rainwater pollution modelization*.
- Nowak, D., J., Crane, D., E., Stevens, J., C. and Hoehn, R., E. (2003). *The Urban Forests Effects (UFORE) Model: Field Data Collection Manual*. USDA Forest Service, Syracuse.
- Odnevall Wallinder I., Verbiest P., He W. et Leygraf C., (1998). - *The influence of patina age and pollutant levels on the runoff rate of zinc from roofing materials*. Corrosion Science 40: 1977-1982
- PLU (Plan Local de l'Urbanisme) de Créteil (2010). <http://www.ville-creteil.fr/PDF/PDF-vm/PDF-urb/Rapport%20de%20presentation%20PLU%202010.pdf>
- Robert-Sainte P., (2009). - *Contribution des matériaux de couverture à la contamination métallique des eaux de ruissellement*. Ecole doctorale "Ville et environnement". Paris: Université PARIS-EST, 335.
- Saporta G. (1990) *Probabilités, analyse des données et statistiques*. Editions technip, Paris
- Sellami-Kaaniche E., De Gouvello B., Gromaire M.C. and Chebbo G., (2012). *Règles de choix des matériaux de couverture dans une ville, Cas de la ville de Créteil*. 5èmes Journées Doctorales en Hydrologie Urbaine Strasbourg (France)