Urban planning for Radio Communications

SP2

Measurements of propagation and cohabitation in Ile-de-France

Task 2.2

RF Measurements and analysis

Deliverable 2.2.4

“Proposed Models Consolidation”

Edition 1.0   01/10/2009

Editions

Edition	Date	Authors	Comments
0.1	18/03/09	M. ABDENNEBI (PRiSM)	Initial version Deliverable content
0.2	12/06/09	D. DEPIERRE (Thales)	Traffic occupancy BTS by BTS Spatial correlation of the traffic
0.3	31/08/09	L. GUEGUEN (OrangeLabs) P. TORTELIER (OrangeLabs) M. ABDENNEBI (PRiSM) S. TOHME (PRiSM)	Traffic characterization per BTS Traffic model Clustering Introduction Conclusion
0.4	01/09/09	D. DEPIERRE (Thales) M. ABDENNEBI (PRiSM)	Some additions and corrections
1.0	01/10/09	D. DEPIERRE (Thales)	Last corrections. Delivered version

Contents

1 Introduction

This deliverable D2.2.4 is related to the sub-project 2, SP2. The URC project has been supported by the French Ministry of Industry, the "Ile de France" region, the departments of Yvelines, Hauts de Seine and Val d'Oise under the frame of the SYSTEMS@TIC cluster (www.systematic-paris-region.org)

SP 2 is dedicated to the general conception of the spectrum management, taking into account traffic characterising and radio measurements. Task D2.2 defines radio measurements and analysis for traffic characterising.

This task D2.2.4 follows the task 2.2.3 and aims to consolidate the proposed models using refined and appropriate measurements.

The consolidation will especially consist in refining the measurements in the space scale, by taking measurements in several areas in Paris. Moreover, those measurements will differentiate the BTSs detected and hence will give the opportunity to analyze the traffic for each BTS, in each area and for each operator, instead of the averaged traffic corresponding to a bigger region.

After refining the measurements, we propose to refine the corresponding analyzing by estimating the correlation in the space scale between each BTS. This estimation will be done using clustering, which consists in merging BTSs with similar models and then by finding correlation between measured traffic. The obtained results will show the correlation between each BTS and a map could be given to illustrate the traffic similarities in the studied areas.

The deliverable will be organized as follow. First, in section 2, we give a method to correct measurements in case of the use of DTX/VAD. The third section gives the radio and traffic measurements obtained from the last campaign with the differentiation of BTSs. The 4th section will then analyze and find the appropriate model for those new measurements and finally cluster the similar models. The 5th section gives the correlation between the differents measurements and finally we end by a conclusion in the last section.

2 Traffic Measurements Correction

For telephony, if both Voice Activity Detection (VAD, in the mobile terminal) and Disrupted Transmission (DTX, in the radio interface) are activated, the measurements should be corrected by the activity factor of the voice, as the number of TCH detected as "free" could be reserved for an active call but with an inactive voice period.

A voice source is characterized by two alternating periods, the active period (ON period) and the inactive period (OFF period) where the source remains silent. In 2G/3G, the coders used in digitizing the voice are usually equipped with a Voice Activity Detector (VAD), which controls the frames coming out of the coder. The «silence» frames are prevented from being emitted through the network to reduce interferences in the radio interface.

The common model is a two-state Markov process, known as the ON/OFF model. The duration of the OFF and ON periods are independent exponentials with respective parameters μ and λ. In the ON period, packets are generated at a constant rate D.

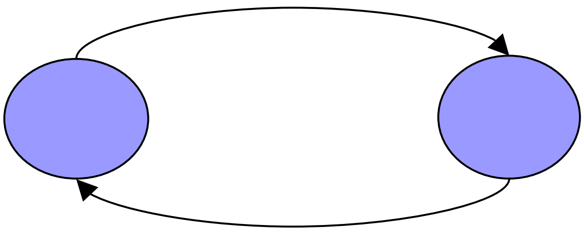


Figure 1: ON/OFF process

1/ λ **(ms)**	**1/ μ** **(ms)**	**1/D** **(ms)**	Packet size (bits)
650	350	20	**244**
1350	**1000**	**20**	**244**

Table 1: ON/OFF model parameters

Table 1: ON/OFF model parametersshows the most commonly used values of the ON/OFF parameters. These values were adopted by ETSI in the UMTS standards. It is worthy to note that the parameters of the model depend on the coder in use (G.711, G.26, G.28, etc.). In fact, the VAD has a great impact on the model; the more sensible the VAD, the longer the OFF periods.

Note, that it seems that the DTX is not allowed in the GSM French network. However, the correction factor to be used if DTX/VAD is activated should be about : $\frac{\lambda}{\lambda+\mu}$.

3 GSM traffic occupancy BTS by BTS

3.1 Introduction

During the measurement campaigns, several standards on various frequency bands were measured and analyzed. Among these standards, GSM (GSM 900, GSM 1800 and E-GSM) and UMTS were studied. GSM and UMTS bands were acquired and analyzed using respectively the ESMERALDA/CASTLE and the SMART AIR 3G tools. To see the complete description of these tools, the metrics it is able to estimate or decode and the measurement characteristics one should refer to [1].

In [2], measurement analysis on mobile cell networks (GSM and UMTS) was presented. This analysis concerned, among others, traffic occupancy (TCH occupancy for GSM, power and spreading factor occupancy for UMTS). It was “macroscopic” in the sense that it presented global results averaged on all the BTS detected by the analysis tool at the measurement point. In the GSM case, the presented results could be averaged on about ten BTS. It showed that this global occupancy matched quite accurately the double Gaussian model and therefore that this model was well suited for areas containing several BTS. In this document, we are going to analyze the occupancy in a more microscopic way, BTS by BTS.

This analysis will be carried out on GSM networks only. Indeed, on the measures performed on UMTS networks, there was, in most cases, one single BTS on the active set. As the traffic channels detection capacity of the used tool is not that reliable for weak BTS (BTS with low E_c/I_0) the analysis BTS per BTS on UMTS network would have been very similar to the one obtained in [2].

Let us remind how the GSM traffic occupancy is computed based on the RF measurements performed on GSM bands. Once the signal files are processed and the metrics computed (Cf. [1]), they are sorted in order to give an overall picture. In the following we are dealing with the traffic occupancy on the frequency channels that are available, at the measurement point. The idea is to see the load of the network as if the measurement equipment was a handset. It means that we are only taking into account the frequency channels that a handset could see and use and not the ones corresponding to further stations. The sorting of the metrics consists in the following steps:

- First, the frequency channels used for broadcasting, at the measurement location, are detected. A channel is considered as a broadcast channel if the CASTLE tool detects a BCH on this frequency channel on 100% of the measures. Each detected BCH corresponds a given BTS (uniquely identified by its MCC, MNC, LAC and CI)
- On the identified broadcast channels, the BCH message is decoded in order to output the “Frequency used” message. This message gives the list of the frequency channels that this BTS can use from traffic.
- The GSM occupancy as defined in [2] is computed on the identified traffic channels. Then it is averaged on all the TCH of a given BTS to obtain the traffic occupancy for each detected BTS.

In the case where two BTS have, in their “Frequency used” message, the same frequency channel for TCH (It can happen depending on how the operator planes its network, especially if the measurement tool detected BTS quite far apart), the TCH is linked to the BTS the closest from the measurement equipment. Indeed it is more likely the detected TCH.

As showed in [2], Bouygues Télécom uses generalized frequency hopping. It means that each BTS has the possibility to use all the available frequency channels they have as TCH. Therefore, it is impossible to link the traffic on a TCH to a given BTS. That is why, the analysis of the traffic occupancy on the Bouygues Télécom network, BTS by BTS, will not be carried out.

For the measurement campaigns in Colombes, Issy-les-Moulineaux and Paris (on five locations) and for the SFR and the Orange networks, a set of three curves is plotted:

- The traffic occupancy averaged on the available TCH versus time, BTS by BTS. This metric, comprised between 0 and 1, represents the load, in terms of traffic, of the BTS. In order to have smoother curves, the plotted metric at a given time is the mean metric on a 60-minutes window, centered on the time of interest. It corresponds to the subplot (a) of the following figures.

- The traffic occupancy $\times$ the number of available TCH versus time, BTS by BTS. This metric is proportional to the number of ongoing communications on the BTS. In order to have smoother curves, the plotted metric at a given time is the mean metric on a 60-minutes window, centered on the time of interest. It corresponds to the subplot (b) of the following figures.
- The mean traffic occupancy along the day versus the mean C/I of all the TCH of the corresponding BTS. The goal of this curve is to verify that null traffic occupancy is not due to a traffic channel with low C/I. Indeed, from the measures there is no way to differentiate a TCH with no ongoing communication from a TCH with ongoing communications but a C/I so low that it cannot be detected. It corresponds to the subplot (c) of the following figures.

On each subplot, each color corresponds to a given BTS. The color code is the same on each of three subplots.

The correlation of the traffic occupancy between each possible set of two BTS is also computed and statistics (mean, standard deviation, maximum value, minimum value) are given. The spatial aspect of these correlations (i.e. how these correlations vary depending on the geographic location of the BTS) will be studied in details in section 5.

3.2 Measurement campaign in Colombes

The analysis of the GSM traffic occupancy in Colombes, on Orange and SFR networks is presented respectively in Figure 2 and Figure 3. The characteristics of the traffic occupancy correlation on both networks are summarized in Table 2.

Orange network

First of all we observe that the traffic occupancy highly varies from one BTS to another even if these BTS are located in the same area. The shape of the curves completely differs from one BTS to another, not only in terms of BTS load (for example, the light green station is 10 times less loaded than the red one) but also concerning the time of the traffic peaks. All the BTS present a peak around 10am but the peaks around 6pm is present on only few stations. These important differences in shape are quantified by the inter BTS traffic occupancy correlations. These correlations are very low (0.34 in average) with only one correlation above 0.75 among 36 (as 9 BTS are detected, 36 correlations are computed) and 15 under 0.25.

Compared to the macroscopic results presented in [2], an important difference can be seen. Indeed, the curves are less steady. The curves do not match the double Gaussian model anymore. For example, if we look at the more loaded BTS (red curve), we can see that it presents four peaks. The number of peaks varies from on BTS to another (between one to five). It seems that at a more microscopic level, the traffic model must be adjusted and more Gaussians must be used.

For this measurement campaign, there is not a high difference between subplot (a) and (b) i.e. between the load of the BTS and the use of the GSM resource. It means that the number of available traffic channels for each BTS is similar. Nevertheless, we observe that the black BTS which is one of the most loaded one becomes one of the lowest in terms of ongoing number of communications. It is due to the fact that this BTS has one single available TCH.

On subplot (c), we can see the traffic occupancy (each star corresponds to one BTS, with the same color code than on subplot (a) and (b)) versus the mean C/I of the traffic channels of each BTS. This subplot shows that there is correlation between the TCH C/I and the traffic occupancy. Indeed, we can observe that the traffic occupancy is lower for BTS with low C/I. It means that in some cases the analysis tool can have difficulties to detect the traffic channel and therefore we can observe null traffic occupancy not because the TCH is not used but because the tool is unable to detect it. It means that we have to be very careful when concluding about traffic on TCH with low C/I.

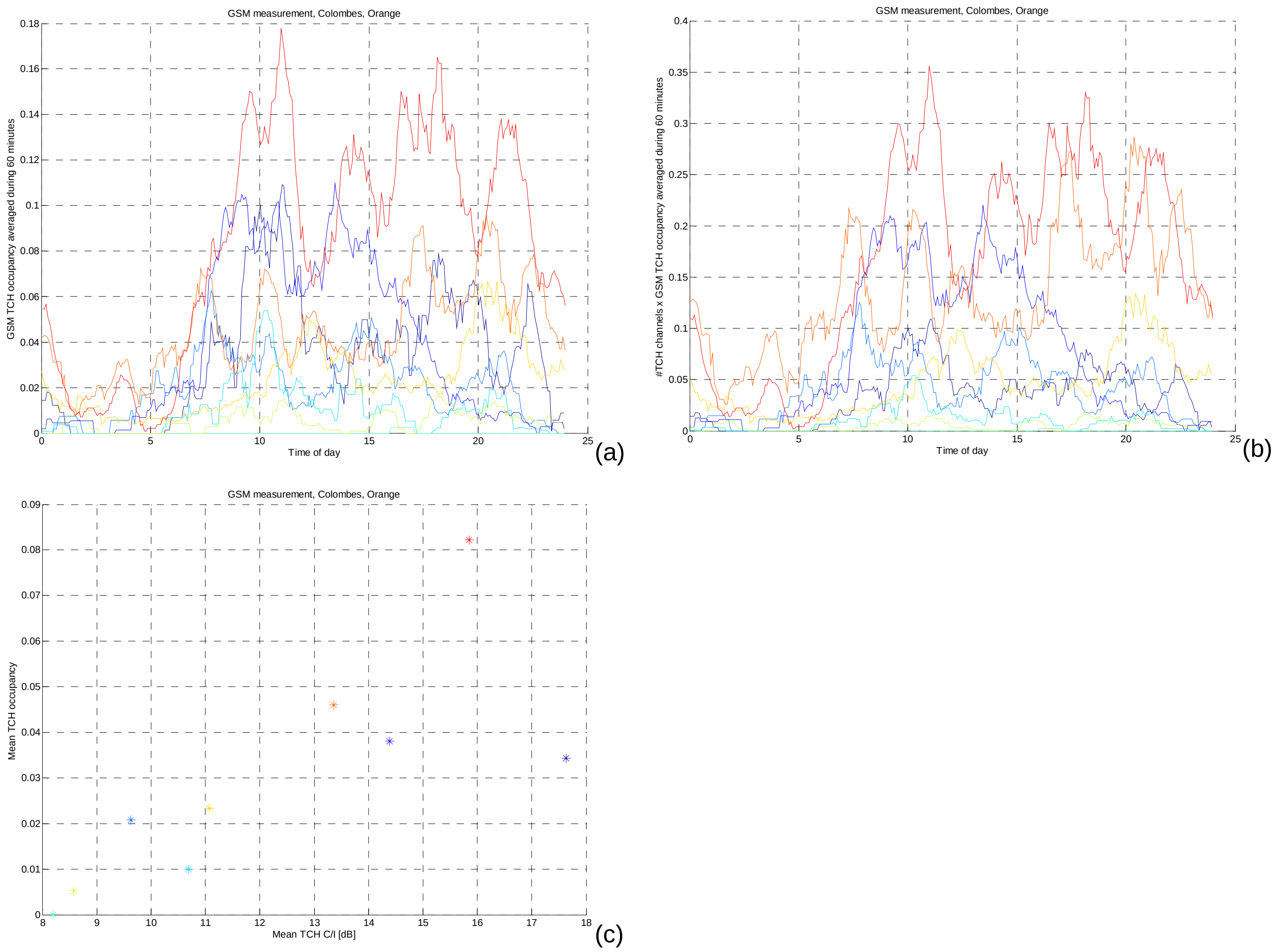


Figure 2 : BTS by BTS GSM traffic channel occupancy characteristics in Colombes, on Orange network

SFR network

On the SFR network, the traffic occupancy BTS by BTS seems more steady than on the Orange one. Nevertheless the double Gaussian model does not seem to be optimal either in this case (two to five peaks are noticeable on the curves).

The global shapes of the curves, from one BTS to another, are more similar. If we look at the correlation between traffic occupancy, we see that it has a mean of 0.52 with peaks higher than 0.9. Among the 28 correlations computed (8 BTS are detected) only 3 are lower than 0.25 and 5 are above 0.75 (4 above 0.85). It means that the way the GSM resource is used along the day from one BTS to another is very similar.

In this case, as for the Orange network, the number of available traffic channels does not change the order of the curves between subploy (a) and (b). It is due, once again, to the fact that all the BTS have similar number of available traffic channels.

We also observe on subplot (c) a correlation between the TCH C/I and the traffic occupancy : BTS with high traffic occupancy tends to be the ones with high C/I. Nevertheless, it does not mean that a low occupancy is always due to a bad traffic estimation on channels with low C/I. Indeed, the black BTS has very low traffic occupancy even if its C/I is similar to the ones of the blue and light blue BTS that have much higher traffic occupancy.

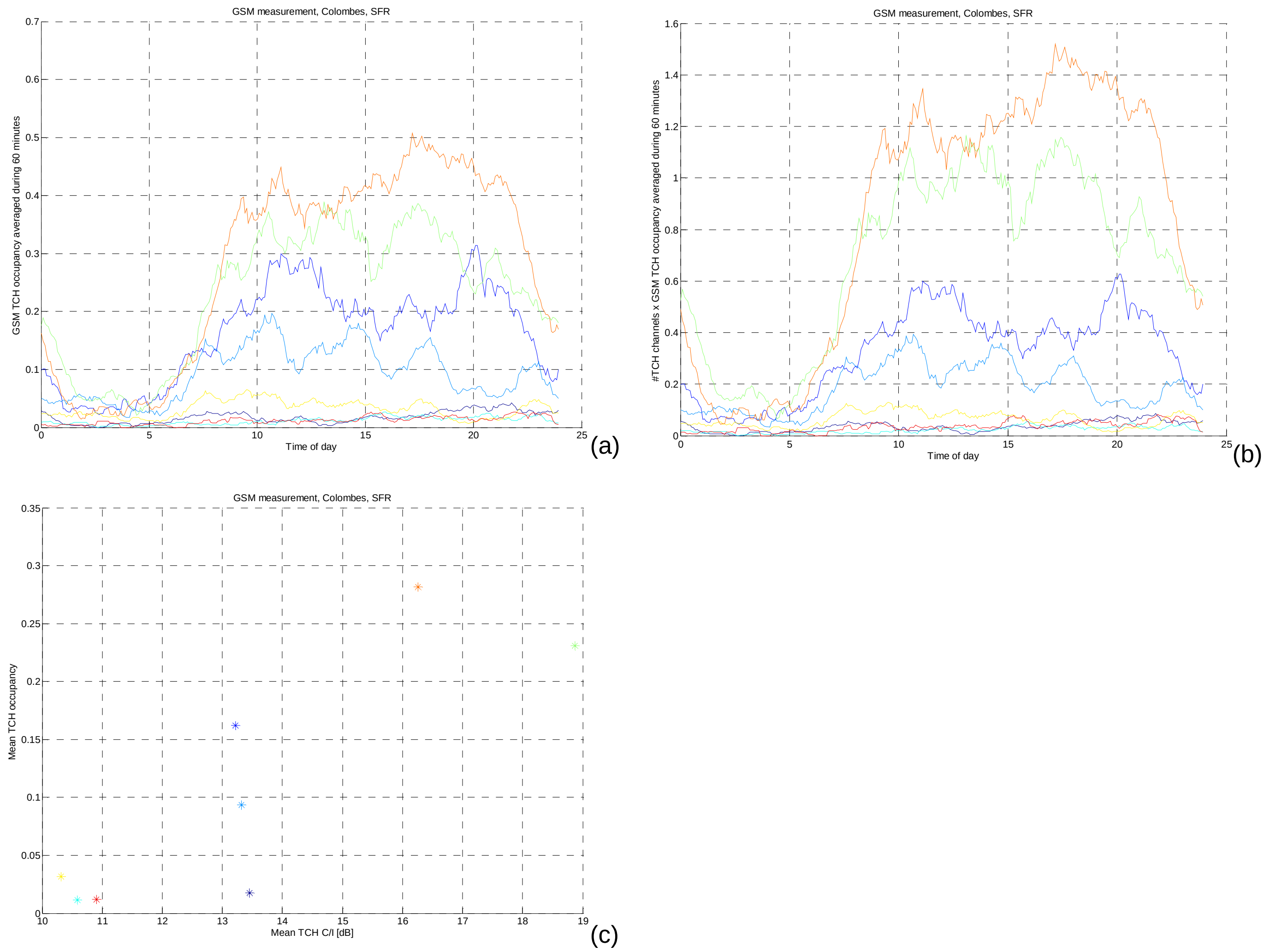


Figure 3 : BTS by BTS GSM traffic channel occupancy characteristics in Colombes, on SFR network

		Orange	SFR
Traffic occupancy correlation between BTS	Mean	0.34	0.52
	Std	0.23	0.25
	Max	0.81	0.94
	Min	0.01	0.02

Table 2 : Characteristics of the traffic occupancy correlation between BTS, in Colombes on Orange and SFR networks

3.3 Measurement campaign in Issy-les-Moulineaux

The analysis of the GSM traffic occupancy in Issy-les-Moulineaux, on Orange and SFR network is presented respectively in Figure 4 and Figure 5. The characteristics of the traffic occupancy correlation on both networks are summarized in Table 3.

Orange network

First of all, compared to measurements performed on the Orange network in Colombes, we can see that the traffic occupancy for any given BTS is steadier. The number of peaks along the day is less important than in Colombes. This time, the double Gaussian model seems more accurate even if three or four Gaussian should be necessary to model the traffic.

From one BTS to another, there is a high correlation in terms of traffic occupancy. On the 28 computed correlations (8 BTS are detected), 10 are above 0.75 and none below 0.25. The mean of the total correlations is equal to 0.65. In this area, on this network, the way the GSM resource is used by the BTS along the day is therefore highly similar.

Looking at subplot (c), we observe once again, that the BTS with lowest traffic occupancy are the ones with lowest C/I. Therefore, we have to pay attention when drawing conclusions on low BTS. We can consider that BTS with C/I under 10 dB should not be taken into account (the minimum working point is set to 9 dB by the GSM standard).

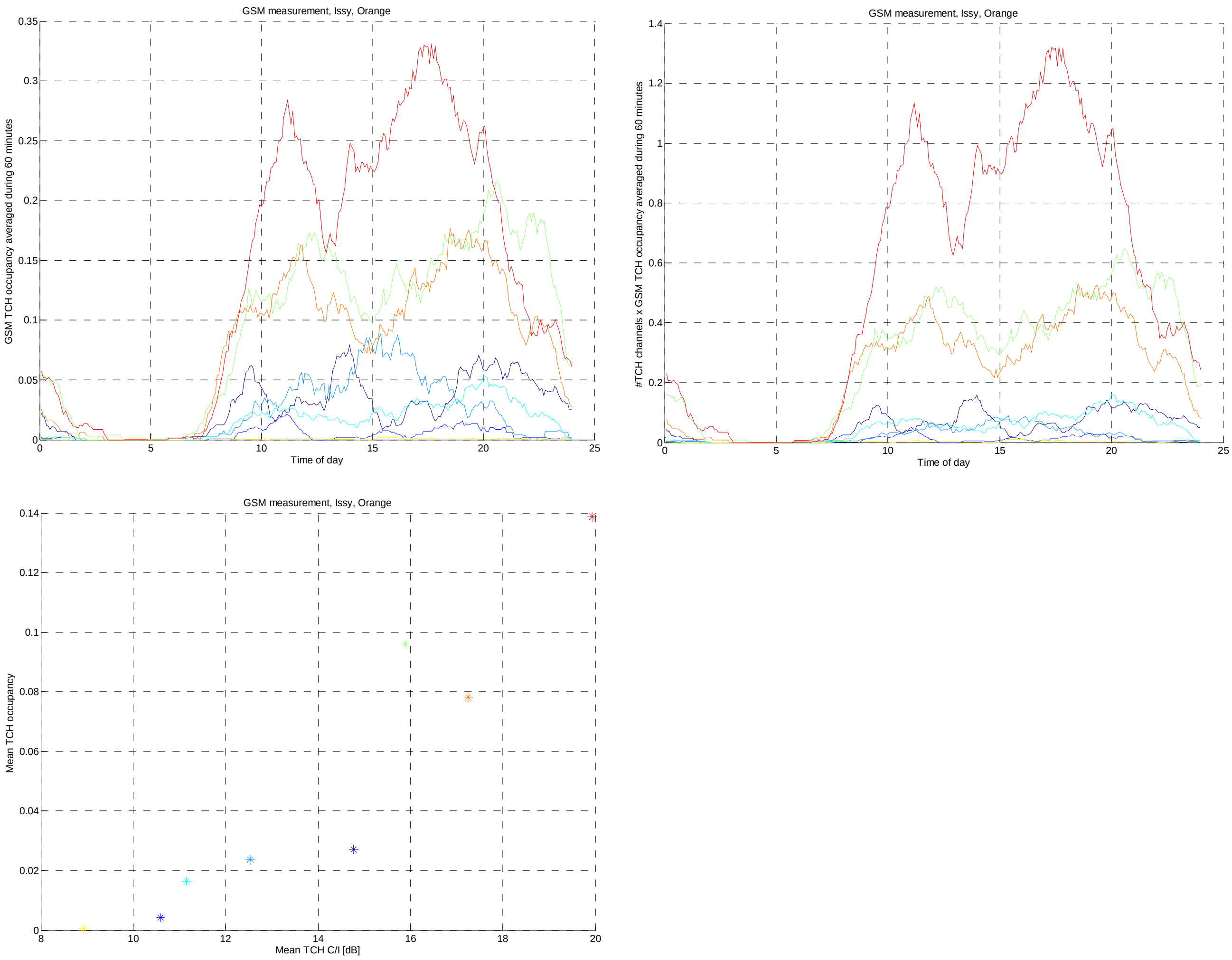


Figure 4 : BTS by BTS GSM traffic channel occupancy characteristics in Issy-les-Moulineaux, on Orange network

SFR network

During this campaign in Issy-les-Moulineaux, the traffic occupancy on the SFR networks seems quite similar to the one on Orange network, as far as the steadiness of the curves is concerned. Moreover, the traffic occupancy from one BTS to another is pretty highly correlated. On the 28 computed correlations (8 BTS are detected), 11 are above 0.75 (and 9 above 0.85) and only 2 below 0.25. The mean of the total correlations is equal to 0.64. As on the Orange network, in this area, the way the GSM resource is used by the BTS along the day is very similar.

Once again, there is a correlation between the traffic occupancy of the BTS and the C/I on their TCH, meaning that the occupancy could be underestimated on BTS with low C/I.

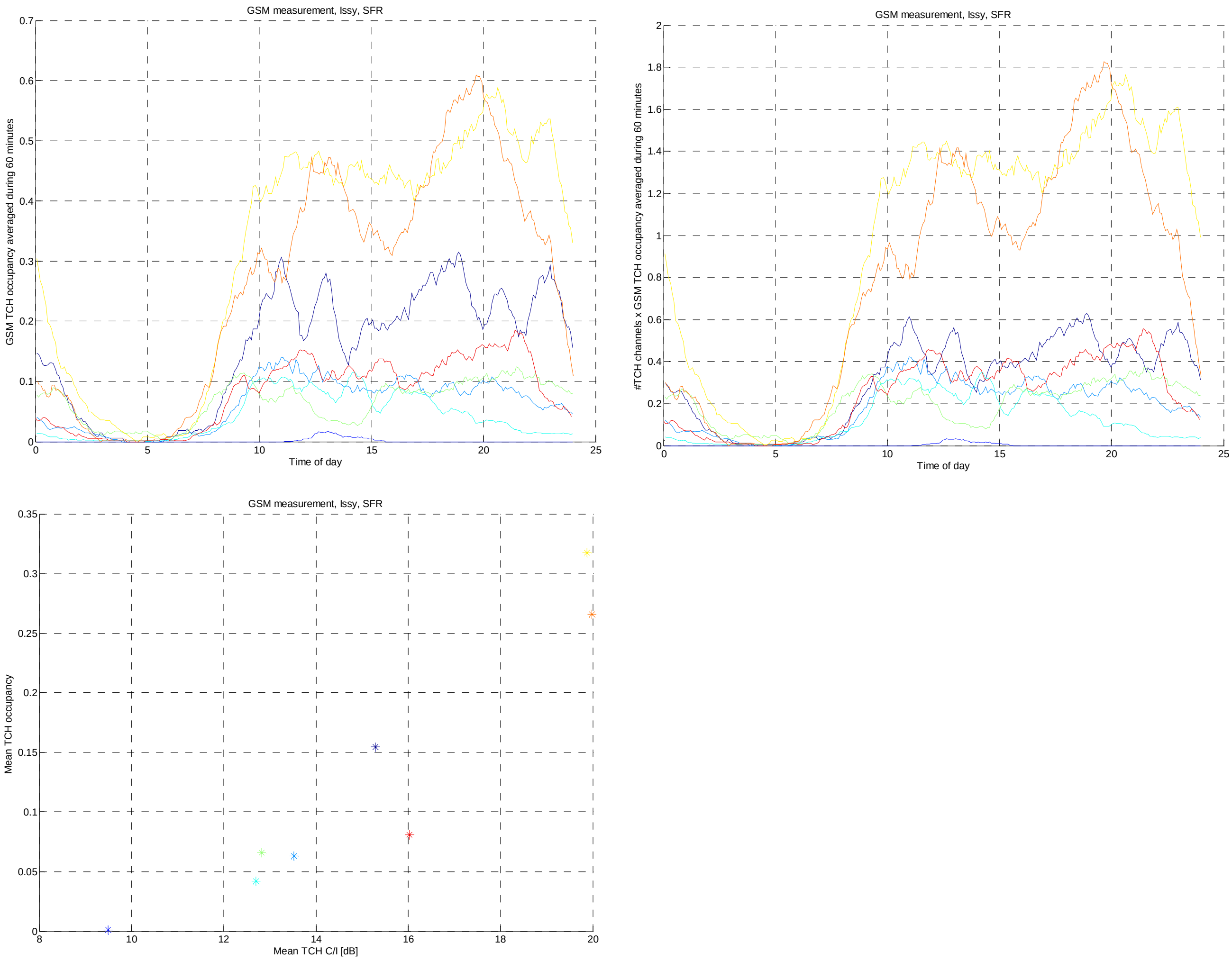


Figure 5 : BTS by BTS GSM traffic channel occupancy characteristics in Issy-les-Moulineaux, on SFR network

		Orange	SFR
Traffic occupancy correlation between BTS	Mean	0.65	0.64
	Std	0.18	0.25
	Max	0.91	0.94
	Min	0.26	0.21

Table 3 : Characteristics of the traffic occupancy correlation between BTS, in Issy-les-Moulineaux, on Orange and SFR networks

3.4 Measurement campaign in five different locations in Paris

The analysis of the GSM traffic occupancy in the five locations in Paris on Orange and SFR network are presented respectively in Figure 6, Figure 7, Figure 8, Table 4 and in Figure 9, Figure 10, Figure 11, Table 5.

Orange network

First of all, compared to the measurements in the suburbs of Paris (in Colombes and in Issy-les-Moulineaux), we can see that the number of detected BTS is significantly lower in Paris (about 2 times lower). It is due to the fact that in urban areas as Paris, the size of the cells is smaller and therefore the power emitted by the BTS lower. So at a given location, the measurement equipment is able to detect less BTS.

Looking at all the traffic occupancies on the five locations, one can observe a very important diversity from one location to another. On a given location, this diversity is less important and the inter BTS occupancy correlation can reach important values (0.82 in Montparnasse, 0.87 in Nation at the most). Nevertheless, the correlations are less important than in the suburbs (in Issy-les-Moulineaux, the average correlation is equal to 0.65 with one third of the point above 0.85), with an average correlation between 0.34 and 0.58.

As for the measurements in the suburbs the double Gaussian model does not apply here. More Gaussians are necessary to model the number of peaks that can be seen along the day (for example three peaks for the light blue BTS in Montparnasse, also three peaks for the dark blue BTS in Louvre).

The number of traffic channels per BTS being similar in a given area, there is not an important difference in the order of the curves between the traffic occupancy and the amount of communications (Cf. Figure 7 and Figure 8). It means that the network is correctly planned and that the most loaded BTS are the ones on which the number of ongoing communications is higher

Due to the network planning density, the detected TCH have higher C/I than in the suburbs (most of the time above 16 dB) and rarely under 10 dB (3 cases over 18). Therefore, except for these three low C/I cases, the estimated TCH occupancies are reliable.

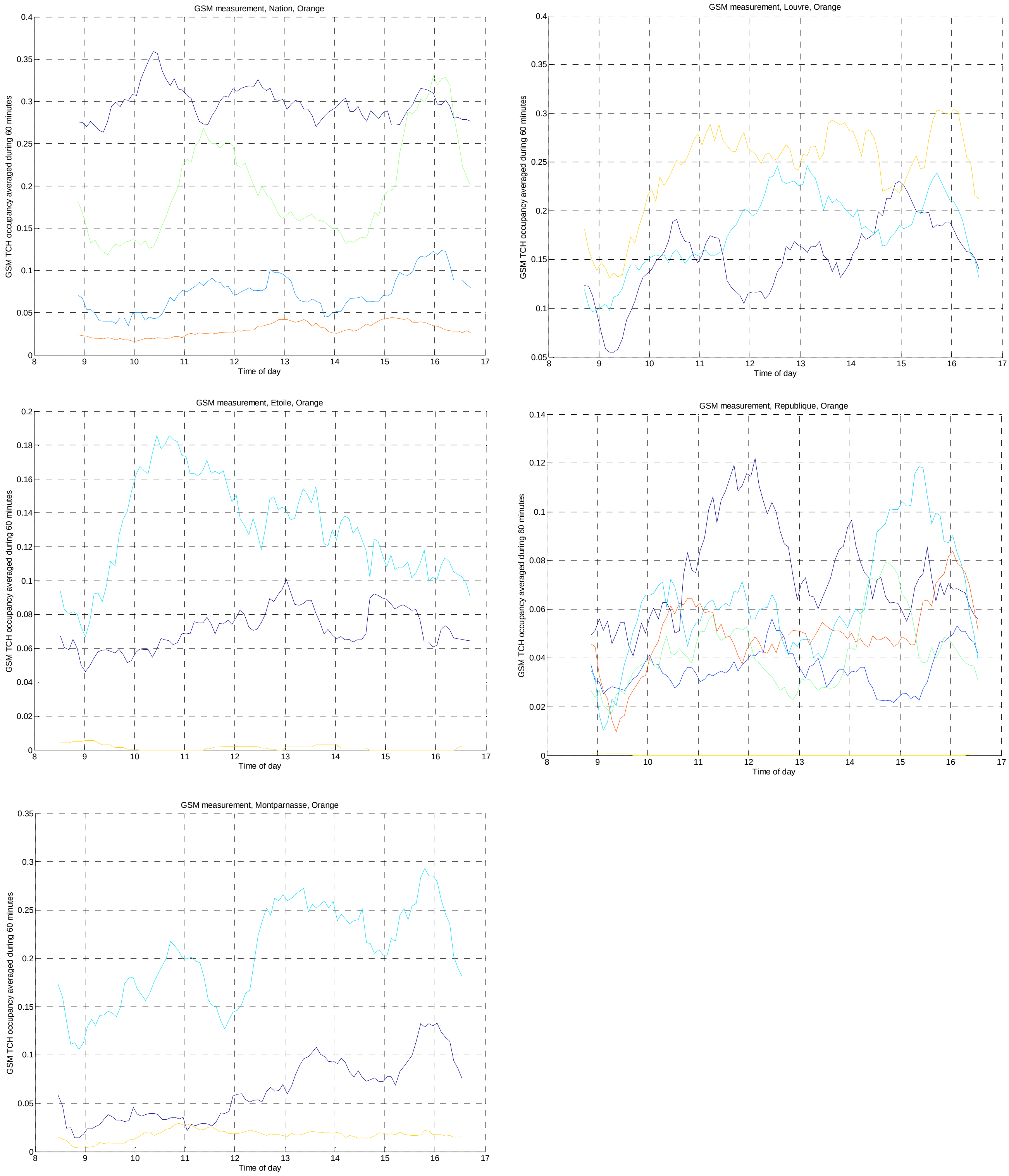


Figure 6 : Traffic occupancy averaged on the available TCH versus time, BTS by BTS, in Nation, Louvre, Etoile, Republique and Montparnasse, on Orange network

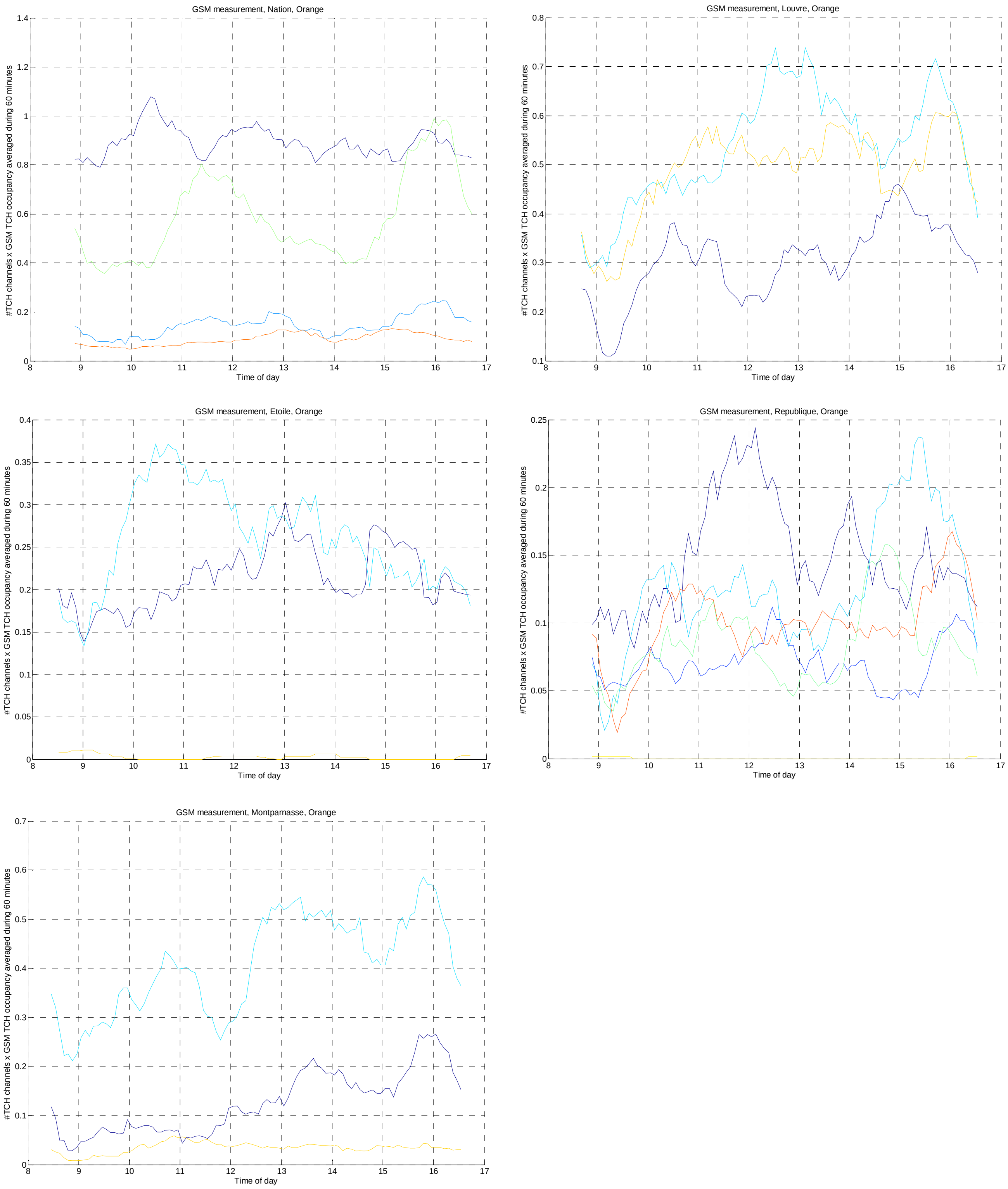


Figure 7 : Traffic occupancy × the number of available TCH versus time, BTS by BTS, in Nation, Louvre, Etoile, Republique and Montparnasse, on Orange network

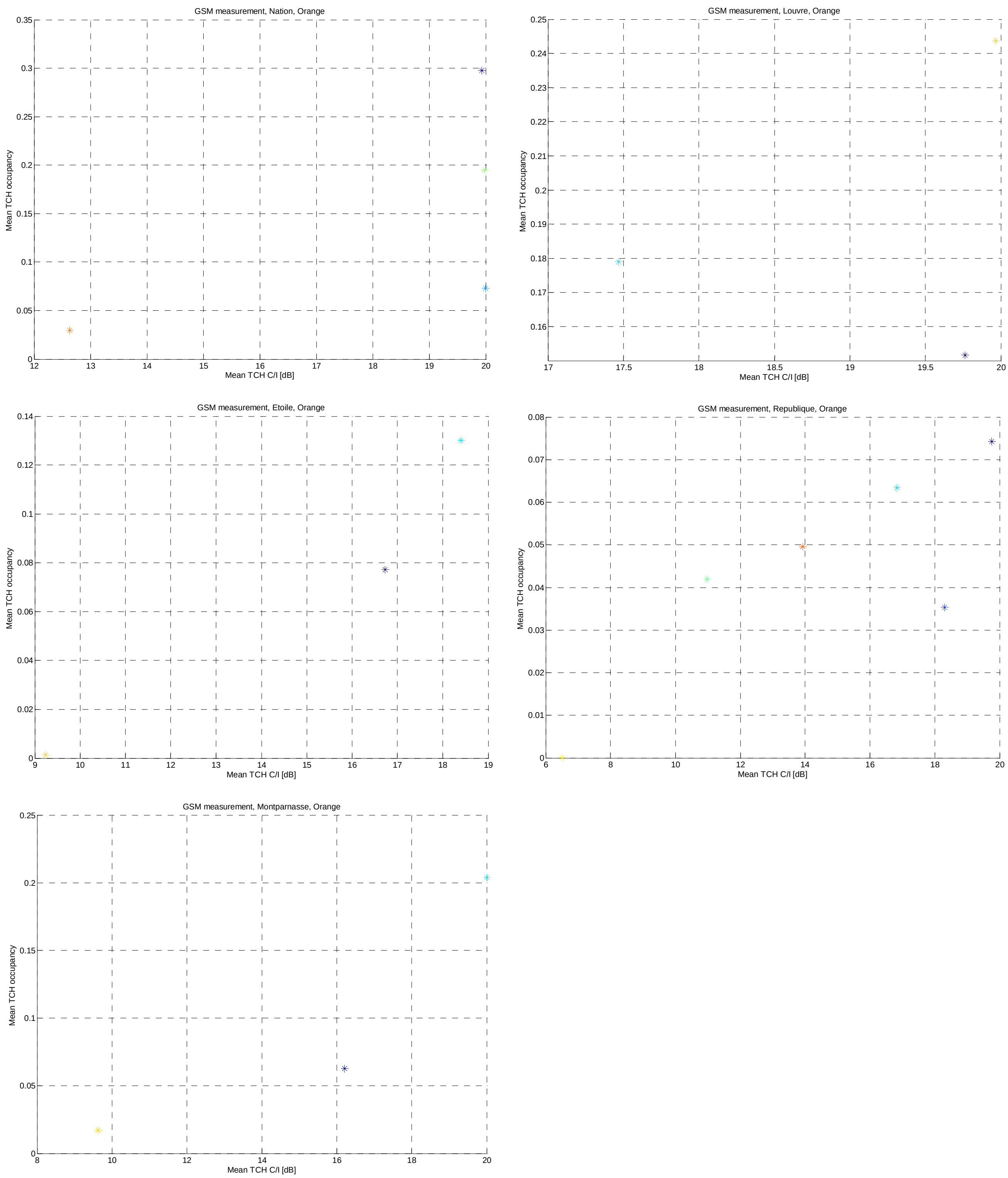


Figure 8 : Mean traffic occupancy along the day versus the mean C/I of all the TCH of the corresponding BTS, in Nation, Louvre, Etoile, Republique and Montparnasse, on Orange network

		Orange				
		Nation	Louvre	Etoile	Republique	Montparnasse
Traffic occupancy correlation between BTS	Mean	0.34	0.58	0.43	0.33	0.48
	Std	0.33	0.16	0.21	0.18	0.31
	Max	0.87	0.75	0.60	0.68	0.82
	Min	0.01	0.42	0.19	0.08	0.20

Table 4 : Characteristics of the traffic occupancy correlation between BTS, in the five locations in Paris, on Orange network

SFR network

Comparing the results, one can notice several similarities between Orange and SFR networks, for these campaigns in Paris.

First of all, due to network planning, the number of detected BTS is significantly lower in Paris than in the suburbs (about 2 times lower).

The traffic occupancy from one location to another importantly differs. There is a high diversity in the curves shape on the five locations in Paris. On a given location, this diversity is less important and the inter BTS occupancy correlation can reach important values (0.99 in Republique). Nevertheless, the correlations are less important than in the suburbs (in Issy-les-Moulineaux, the average correlation is equal to 0.65 with one third of the point above 0.85), with an average correlation between 0.31 and 0.59.

Looking at Table 4 and Table 5, one can see similarities about the correlation characteristics for the five locations, between Orange and SFR networks : highest mean is in Louvre and the lowest in Nation.

Once again, the double Gaussian model does not apply here. More Gaussians are necessary to model the number of peaks that can be seen along the day (for example three peaks for the green BTS in Republique, also three for the dark blue BTS in Nation).

The number of traffic channels per BTS being similar in a given area, there is not an important difference in the order of the curves between the traffic occupancy and the amount of communications (Cf. Figure 9 and Figure 10). It means that the network is correctly planned and that the most loaded BTS are the ones on which the number of ongoing communications is higher.

Contrary to the Orange network, the detected TCH do not have high C/I. The observed C/I values are closer to the ones observed in the suburbs of Paris. Therefore a great care must be taken when concluding about the lowest BTS.

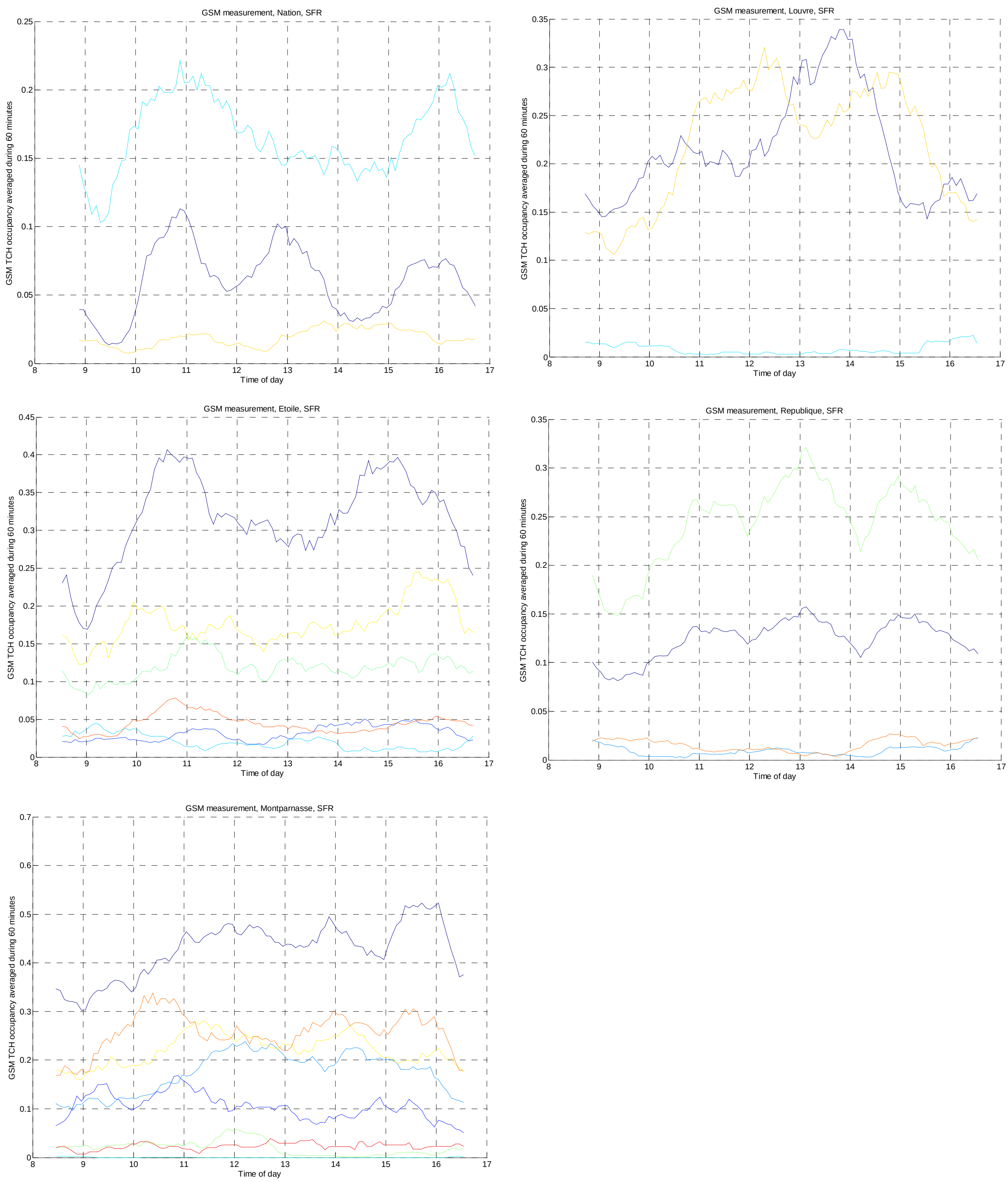


Figure 9 : Traffic occupancy averaged on the available TCH versus time, BTS by BTS, in Nation, Louvre, Etoile, Republique and Montparnasse, on SFR network

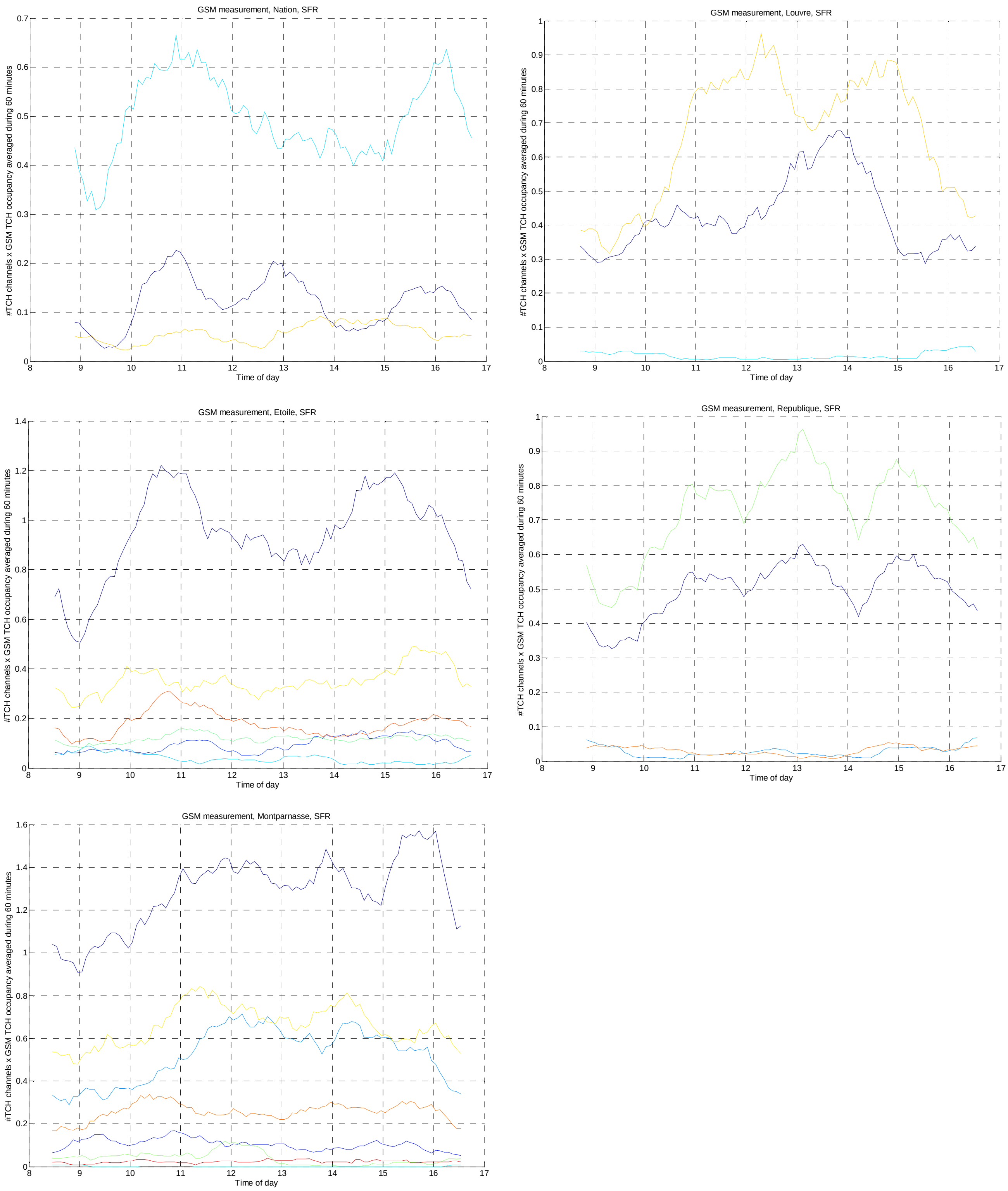


Figure 10 : Traffic occupancy × the number of available TCH versus time, BTS by BTS, , in Nation, Louvre, Etoile, Republique and Montparnasse, on SFR network

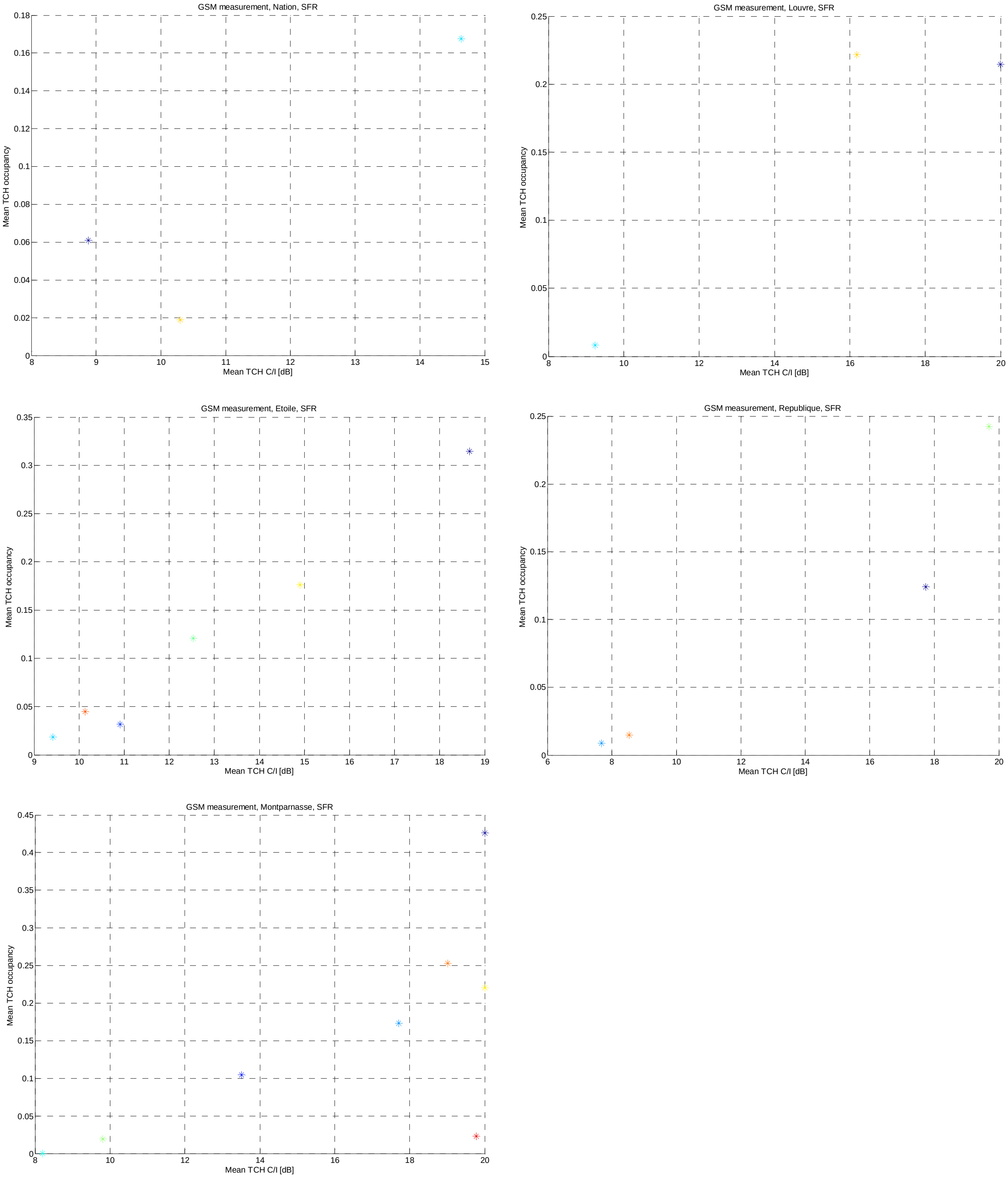


Figure 11 : Mean traffic occupancy along the day versus the mean C/I of all the TCH of the corresponding BTS, in Nation, Louvre, Etoile, Republique and Montparnasse, on SFR network

		SFR				
		Nation	Louvre	Etoile	Republique	Montparnasse
Traffic occupancy correlation between BTS	Mean	0.31	0.59	0.47	0.44	0.33
	Std	0.34	0.17	0.17	0.31	0.23
	Max	0.67	0.78	0.64	0.99	0.74
	Min	0.00	0.48	0.03	0.10	0.02

Table 5 : Characteristics of the traffic occupancy correlation between BTS, in the five locations in Paris, on SFR network

3.5 Conclusion

In [2], we showed that the double Gaussian model fitted quite accurately macroscopic GSM traffic occupancy. By macroscopic, we mean on a geographic zone covering several BTS corresponding to the BTS that a handset is able to see at the measurement location.

In this section we pointed out, based on the measurements performed in Paris and in the suburbs (Colombes and Issy-les-Moulineaux) that this model does not fit the microscopic GSM traffic occupancy, i.e. the traffic occupancy on a single BTS. Indeed the microscopic GSM traffic occupancy is less steady than the macroscopic one. It presents more than two peaks (up to five). Moreover, from one BTS to another, even in the same geographic area, the number of peaks and their respective time can vary significantly. New models have to be developed if microscopic traffic has to be modeled. That what will be studied in the following sections.

Nevertheless the computation of the correlation of the traffic occupancy between sets of two BTS showed that in the suburbs they could reach high values (0.94), with important means (0.64 and 0.65, in Issy-les-Moulineaux, in respectively SFR and Orange networks). In Paris, it is harder to find traffic correlations between BTS. The means, in the five measurement locations are lower (around 0.45). However we have to remain careful when concluding about the measures of Paris. Indeed for these measurement campaigns, the number of detected BTS is much lower than in the suburbs (sometimes only two against more than ten). It is due to network planning in dense urban areas where the cells are small and the emitted powers low. The number of available data is therefore less significant.

4 BTS traffic modeling

In this section, the traffic will be characterised with a refined space scale. Obtained results will be used to model the traffic per base station and also to analyse the differences between each studied area/region.

Hence, we will first find the appropriate model for each BTS, using the measurements, and then analyse the obtained results to obtain the space correlation.

4.1 Traffic modeling and parameterization

In this sub section, we propose to extend the large time scale modelling from the double Gaussian model to a model based on multiple Gaussian model. This model should fit correctly to small areas, especially for single base station traffic. Indeed, in section 3, we saw that when looking at a single BTS traffic occupancy, the double Gaussian model did not fit anymore and that more Gaussians were needed to adjust the model.

As each occupation profile will be approximated by a mixture of Gaussians, the corresponding model parameters should be characterised by the appropriate Gaussians variables to fit with the traffic B(t).

$$B(t) \approx \sum_k \pi_k N(\mu_k, \sigma_k)$$

Where N is the gaussian function and μ_k, σ_k and π_k are respectively the mean, the standard deviation and the multiplicative factor of each component.

$$B(t) \approx \sum_k \pi_k . \frac{1}{\sigma_k{}^2 \sqrt{2\pi}} e^{\frac{-(t-\mu_k)^2}{\sigma_k{}^2}}$$

From a real observation we should be able to approximate the parameters with sub indexes k.

The estimation of each traffic model should follow the following steps:

- First, we transform B(t) into a distribution b(t) such that:

$$b(t) = \alpha(B(t) - C)$$

$$\int_0^1 b(t)dt = 1$$

By doing this transform, we focus on the shape of the occupation more than the absolute value.

- Then, we generate n (n=10000 for example) random sample {t1,… , tn} following the empirical distribution b(t). These samples will be used to generate a Gaussian function with random parameters which will be the starting point to estimate the values corresponding to parameters of the model of each area.
- After that, we fix the number of Gaussians to be K. From the set {t1,… , tn}, we estimate the parameters of the mixture of Gaussian by finding the maximum likelihood estimation of those values. The algorithm used is Expectation-Maximisation (EM algorithm); this algorithm for Gaussian distribution is gmdistribution.fit in MATLAB.

We then obtain the following parameters:

$$P = \{\pi_k, \mu_k, \sigma_k\}_{k=1}^{K}$$

Note that this estimate is done on the filtered measurements. The filtering operates on measurements with a scale of 1 hour.

This operation is repeated until finding the optimal number of Gaussian mixtures. We use the Baysian Information Criterion (BIC) to determine this optimal number K. We perform the estimate from increasing number K and select the one which gives the minimum BIC.

Below the example of the obtained result for the BTS# 2 measured in *Etoile* area on the SFR radio network. We also give the standard error *se* (or root mean squared error) between measurements and estimation.

Note that even if 5 gaussian components were identified, the plot will show 4 main Gaussians as the 2nd and the 3rd ones are very close to each other.

Component 1:
Mixing proportion: 0.128460
Mean: 9.2523

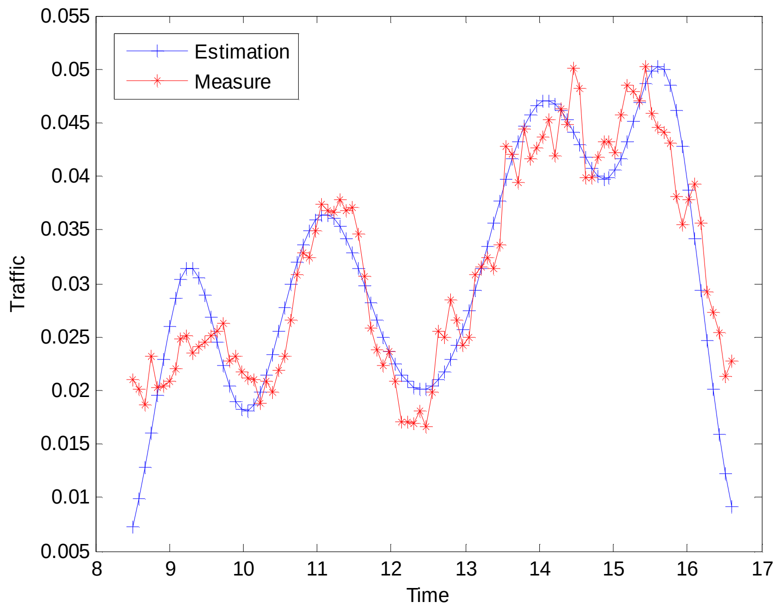
Component 2:
Mixing proportion: 0.138749
Mean: 10.9867

Component 3:
Mixing proportion: 0.159955
Mean: 11.6752

Component 4:
Mixing proportion: 0.198195
Mean: 15.7308

Component 5:
Mixing proportion: 0.374641
Mean: 14.1132

MSE = 0.0014



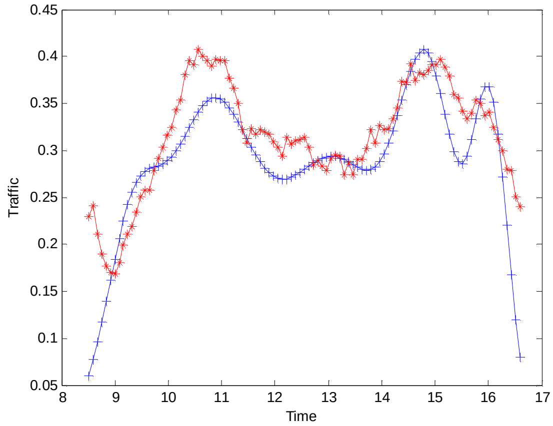
Model parameters

Traffic, estimation (blue) and measurements (red)

Figure 12 : Obtained traffic model for SFR Etoile, BTS #2

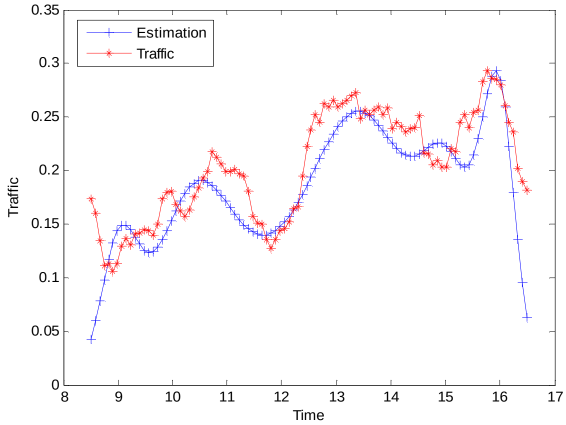
This operation is also repeated for the other measurements and we hence estimate the parameter for each concerned regions/areas.

The results are given below for some of BTSs detected.



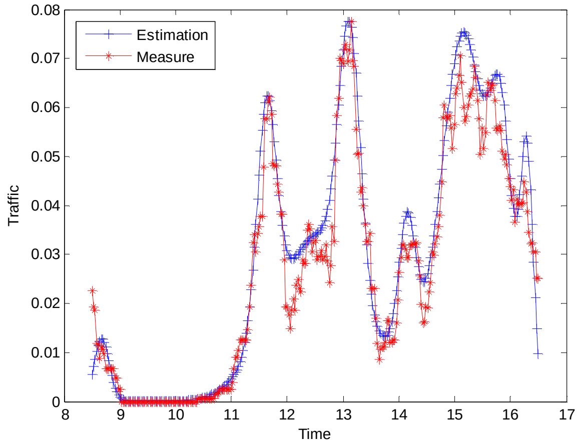
MSE = 0.0021

SFR etoile, BTS #1



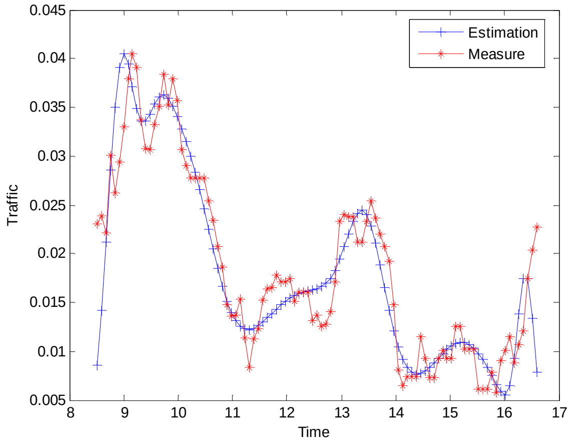
MSE = 0.0016

Orange Montparnasse – BTS # 2



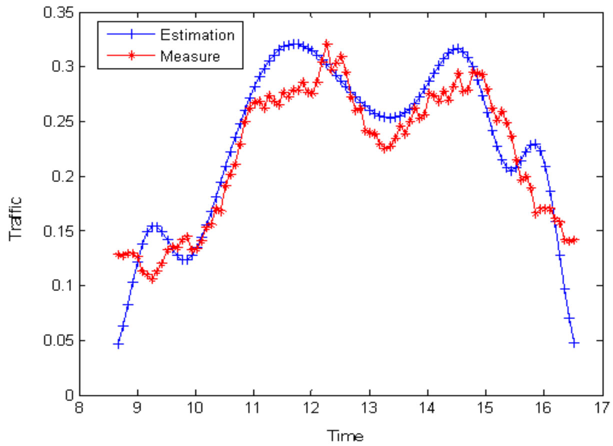
MSE = 0.0008

Orange Issy – BTS # 1



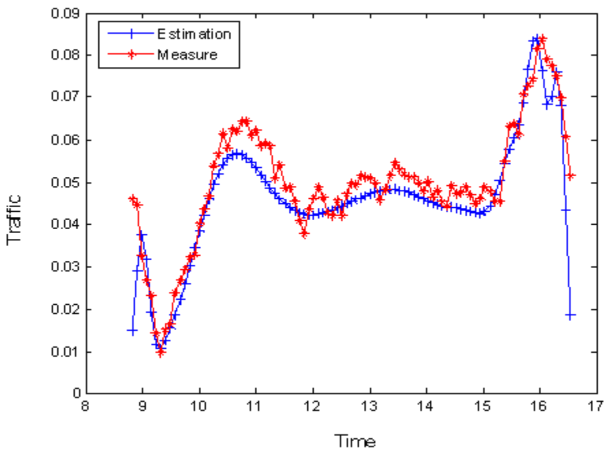
MSE = 0.0012

SFR etoile, BTS #3



MSE = 0.0013

SFR Louvre, BTS #3



MSE = 0.002

Orange République, BTS #6

Figure 13 : Obtained traffic model for SFR and Orange, In various locations

We can observe from those plots that the K-Gaussian model fit fairly well with obtained measurements, whatever is the number of Gaussians measured.

4.2 Traffic analysis in space

In this sub-section we propose to analyze the parameters in space through a clustering of the obtained parameters. Let {b1(t),…, bm(t)} be m occupations profiles from which we have extracted m parameters sets {P1,…,Pm}

We want to cluster the parameters sets. As the parameters have various lengths, we need an adapted method to analyze them. We propose information derived metric to measure similarity between parameters.

Let us suppose that measured curve b(t) is conditionally distributed with respect to a parameter P. Let c(t;P) the occupation curve generated taking the parameter P. Thus, we can compute the error between b(t) and c(t;P). We assume that these differences are Gaussian i.i.d, such that if we have j temporal measurements, the conditional log distribution is given by:

$$\log\left(p(b|P)\right) = -0.5\, j\log(2\pi) - j\log(\sigma) - 0.5\sum_{t=1}^{j}\frac{(b(t)-c(t;P))^2}{\sigma}$$

Assuming that we determine P1 from b1, we define the deviation as the mean standard deviation between the measurements:

$$\sigma_1 = \sqrt{\frac{1}{j-1}\sum_{t=1}^{j}(b_1(t)-c(t;P_1))^2}$$

And the conditional distribution is expressed as in the previous equation. We normalize the conditional distributions by taking 1/j log(p(b|P)).

For any parameters Pk and occupation curve bq(t), we can compute p(bq|Pk). Therefore we define the distance between two parameters Pk, Pl as the Kullback Leibler distance with their respective conditional distribution on the observed data:

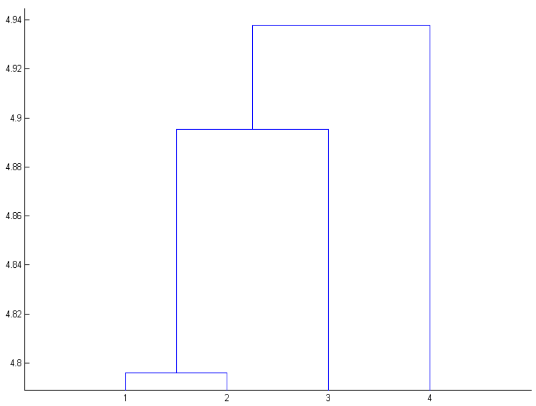
$$d(P_k,P_l) = d_{KL}(p(B|P_k);p(B|P_l)) + d_{KL}(p(B|P_l);p(B|P_k))$$

$$d_{KL}(p(B|P_k);p(B|P_l)) = \sum_{i=1}^{m} p(b_i|P_k)\log\frac{p(b_i|P_k)}{p(b_i|P_l)}$$

Having the distance between each pair of parameters, we can perform a hierarchical clustering taking the complete distance. You can use 'linkage' function in MATLAB.

We performed the analyses for GSM traffic measurements and we analysed especially the correlation between BTSs in the same area/region.

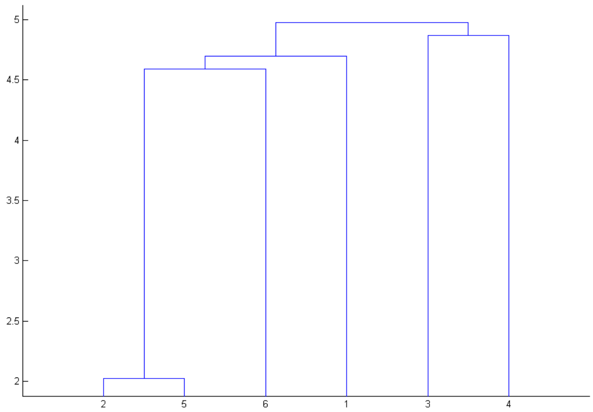
We illustrate below with two examples of traffic space correlations obtained for two areas, Nation and Republique (Orange operator). The figures below give the clustering result, which means that the BTSs could eventually be merged together in the same area model.



BSs coordinate

BS#1	604223	2428184
BS#2	604223	2428184
BS#3	604223	2428184
BS#4	604800	2428200

Nation, Orange network



BSs coordinate

BS#1	601592	2429344
BS#2	602047	2429313
BS#3	601592	2429344
BS#4	602433	2429328
BS#5	603018	2428439
BS#6	601780	2429640

Republique, Orange network

Figure 14 : Example of clustering for Nation and Republique areas (Orange)

From the obtained results, we can see that the correlation in Republique area is especially between the BTSs #2 and BTS #5, while it is between BTSs #1 and BTS #2 in Nation area.

This correlation could be illustrated by the correlation of the shapes of the two plots regardless of their absolute value.

In the figure below, we illustrate on a map, the obtained cluster in the area of Republique (the colour indicates the cluster precision).

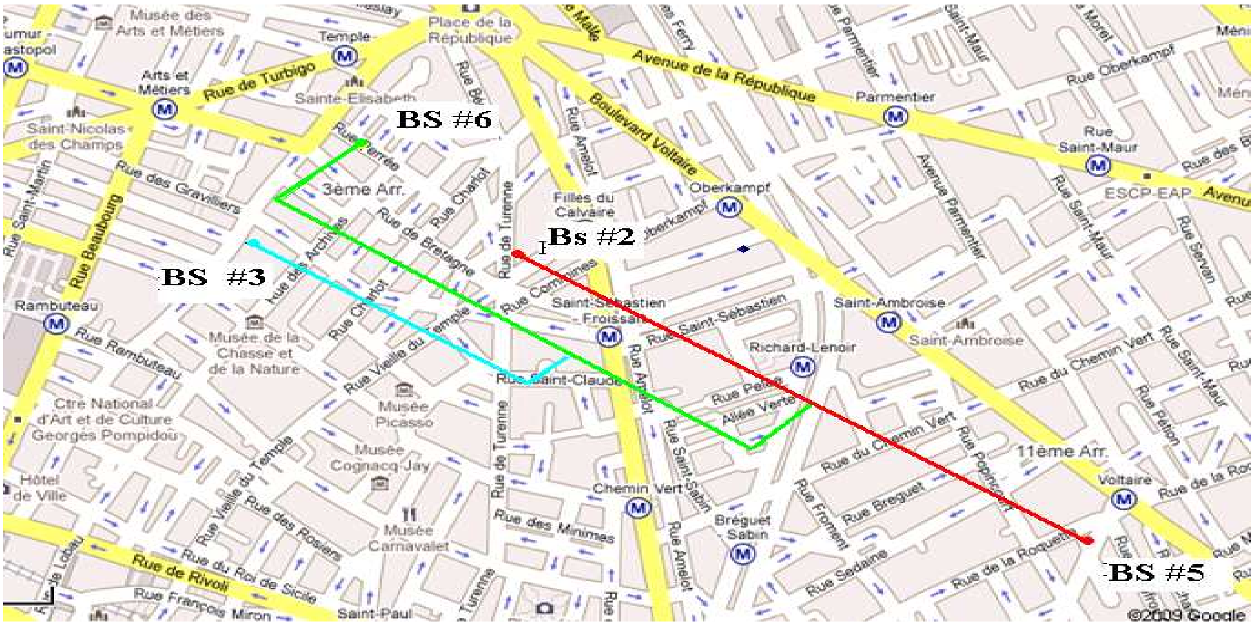


Figure 15 : Republique (Orange operator) obtained clusters

We also propose to illustrate the similarity between studied areas and how the clustering could be obtained from two different models. We use Nation results as an example to illustrate the procedure.

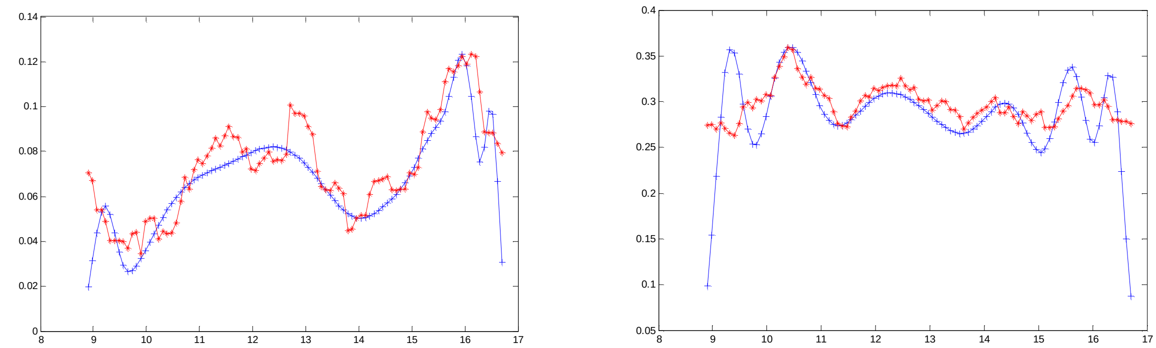


Figure 16 : BTS #1 and BTS #2 traffic (red = measure, blue = estimate) in Nation (Orange operator)

As we can see, the clustering algorithm gives the opportunity to merge the concerned BTS model as one single traffic model for the corresponding areas.

For example, BTSs #1 and BTS #2 can be represented by one single model as a mixture from the two different models.

It is important to note that in the case below, only the shapes of the plots were taken into account. The figure below illustrates the obtained model after the clustering.

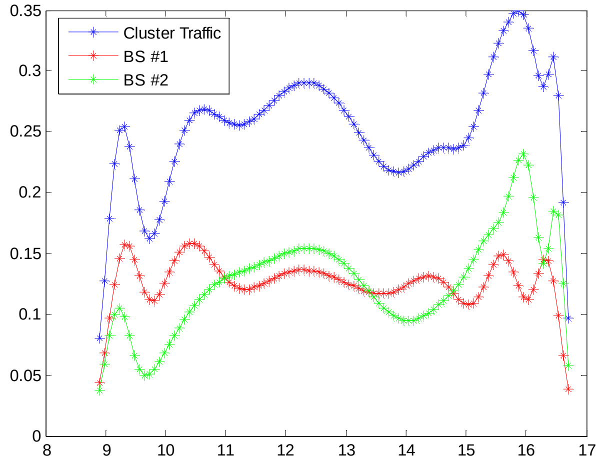


Figure 17 : BTS #1 and BTS #2 and merged model for Nation area (Orange operator)

5 Spatial correlation of GSM traffic occupancy

5.1 Introduction

In section 3, the GSM traffic occupancy BTS per BTS were studied. We are now going to analyze if the traffic occupancy of BTS of a same geographic area is correlated and if this correlation is linked to the geographic positioning of the BTS. We will try to answer the following question: Is there a high correlation between the traffic occupancy of the three collocated BTS of a tri-sector site or between neighbor BTS?

In order to study the traffic occupancy correlations between BTS depending on their location it is mandatory to link the detected BTS (uniquely identified by their MNC, MCC, LAC and CI) and their GPS location. On the Internet, some databases are available with GPS coordinates of BTS (www.realtimeblog.free.fr, www.opencellid.org, etc.) but these databases are not complete (only 10% of the BTS detected during our measurement campaigns were listed in these databases) and the GPS coordinates are not very accurate (the locations are not the real one but are calculated based on some algorithms). Therefore, as far as spatial traffic correlation is concerned, we only focused on the Orange network. Indeed, Orange being a partner involved in URC, we had the exact locations of the detected GSM base stations.

In section 3, we saw that the traffic occupancy computed on BTS with low TCH C/I was not reliable, that is why, in this section, we only take into account the BTS whose mean C/I on the traffic channels they actually use is over 10 dB. We took 1dB margin compared to the minimum C/I imposed by the GSM standard (9dB).

In the following, cartographic representations of the traffic correlation between BTS are given. In the maps, the green dot corresponds to the location of the measurement bench, the red triangles to the detected BTS (even the ones with low C/I that are not taken into account when processing inter BTS traffic occupancy correlations) and the color of the lines between two BTS to the traffic occupancy correlation (red for high correlation, blue for low correlation). As up to three BTS can be collocated, several lines can be drawn between two locations. For correlation between two collocated BTS, a circle is drawn.

For the measurement campaigns in Colombes and Issy-les-Moulineaux where an important number of BTS is detected, the traffic occupancy correlations between BTS versus the inter-BTS distance is also plotted.

5.2 Measurement campaign in Colombes

Figure 18 shows the cartographic representation of the inter-BTS traffic occupancy for the measurement campaign in Colombes. First of all, one can remark that several remote BTS are detected, more than 5-km away from the measurement equipment. Even if low BTS are not taken into account, the inter BTS distances are quite high (between 600m and 3000m). As far as the correlation is concerned, it is low (in most cases below 0.6). The only case with high correlation (above 0.8) is seen between two collocated BTS.

Looking at Figure 19, it seems that, for this campaign, there is no clear connection between inter BTS traffic occupancy and distance between BTS. Indeed, we can see low correlation for close BTS (870m) and distant BTS (2240m) and correlation at 0.6 for several different distances (650m, 2250m, 3000m).

The two red triangles with no link with another triangle correspond to detected BTS that were not taken into account as the C/I on their traffic channel was too low.

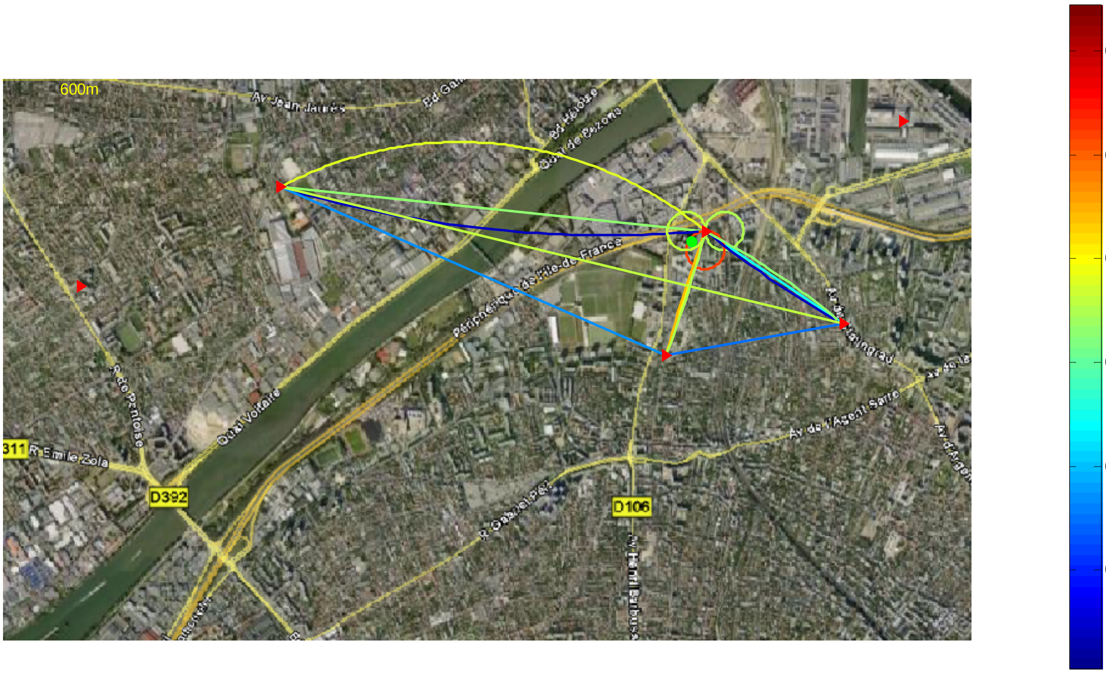


Figure 18 : Spatial correlation between BTS traffic occupancy, in Colombes, on Orange network

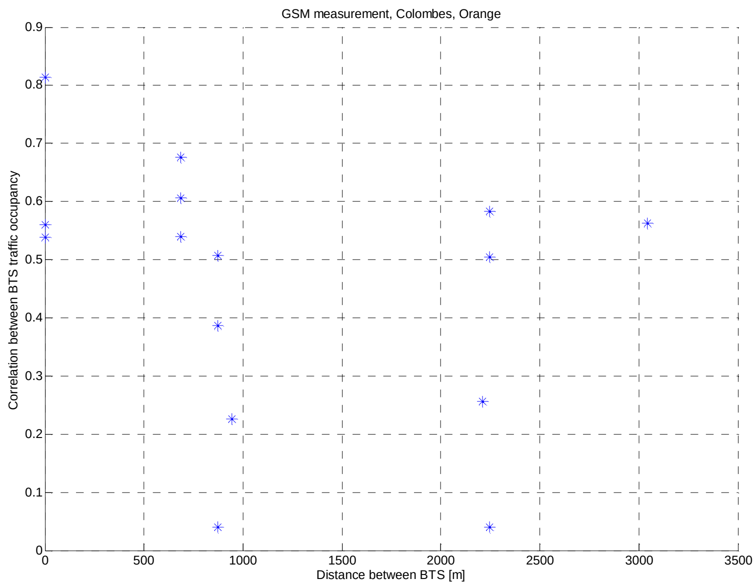


Figure 19 : Correlation between GSM BTS traffic occupancy versus inter-BTS distance, in Colombes, on Orange network

5.3 Measurement campaign in Issy-les-Moulineaux

Figure 20 shows the cartographic representation of the inter BTS traffic occupancy for the measurement campaign in Issy-les-Moulineaux.

Compared to the campaign in Colombes, the detection range is less important, the further BTS are 2000m away (against more than 3000m). It is due to the difference in network planning for both cities; Issy-les-Moulineaux being more urban than Colombes. Moreover, much more highly correlated cases can be seen (more than 30% of the points with correlation higher than 0.8).

In Figure 21, we see that very highly correlated BTS (with correlations above 0.89) are quite close (less than 600m apart) and that the highest correlation is seen for collocated BTS. We also notice that roughly, the correlation between two BTS decreases with their distance. It tends to prove that, for this campaign, on small geographic areas (of about one to two cells) the way the GSM resource is used is the same.

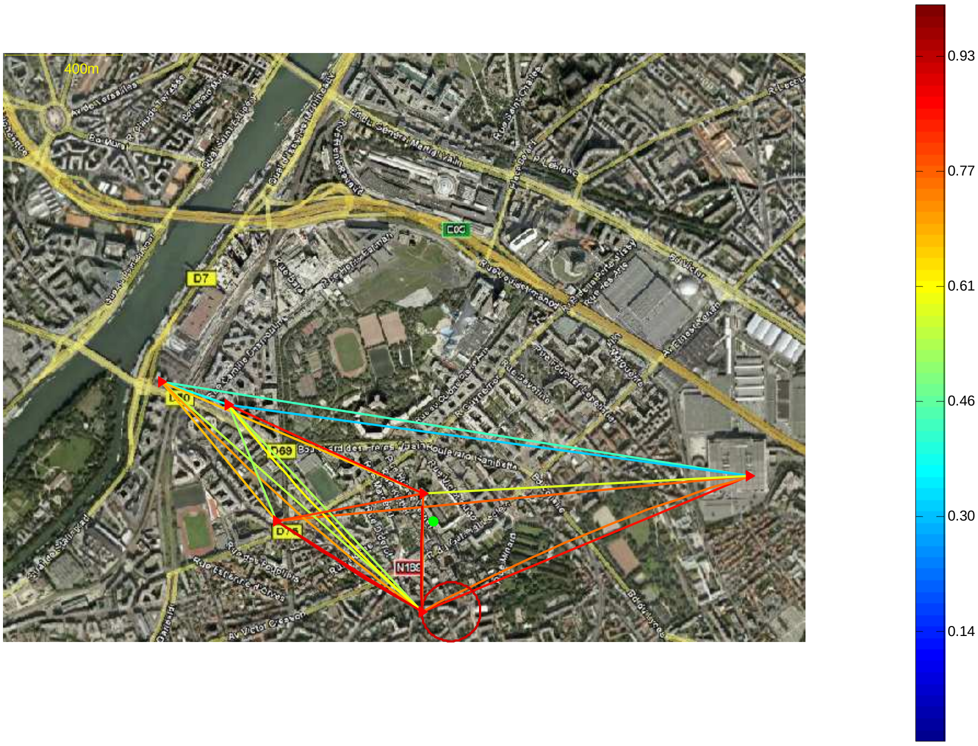


Figure 20 : Spatial correlation between BTS traffic occupancy, in Issy-les-Moulineaux, on Orange network

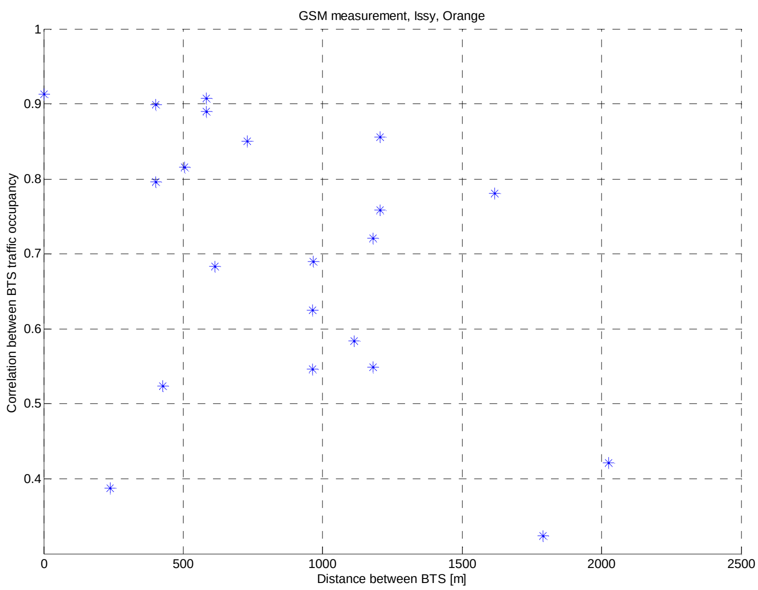


Figure 21 : Correlation between GSM BTS traffic occupancy versus inter-BTS distance, in Issy-les-Moulineaux, on Orange network

5.4 Measurement campaign in five different locations in Paris

Figure 23 shows the cartographic representations of the inter BTS traffic occupancy for the measurement campaigns in Paris, respectively in the areas of Nation, Le Louvre, Etoile, Republique and Montparnasse.

For these campaigns, as we are in a dense urban configuration, the number of detected stations as well as the area covered by these stations is significantly lower than for the two campaigns in the suburbs. The number of detected stations can be very small (2 stations in Etoile and Montparnasse, 3 in Louvre), therefore it is hard to draw reliable conclusions. Nevertheless some trends can be seen.

Figure 22 shows the correlation between GSM BTS traffic occupancy versus inter-BTS distance, for the 5 locations in Paris. The correlation is computed only between BTS from the same area. First of all, we notice that the values of the correlations are much lower than in the suburbs. Most of the points are comprised between 0.2 and 0.6. The three highest correlations are found for collocated BTS but their values (0.73, 0.80 and 0.87) are once again low compared to what we saw in Issy-les-Moulineaux (30% of the correlation above 0.8). It is as if, in Paris, the population in terms of GSM usage was very disparate. The way people use their mobile from one block to another seems to change importantly.

Surprisingly, we also notice that the lowest correlations are found for collocated BTS. No explanation seems to justify this observation.

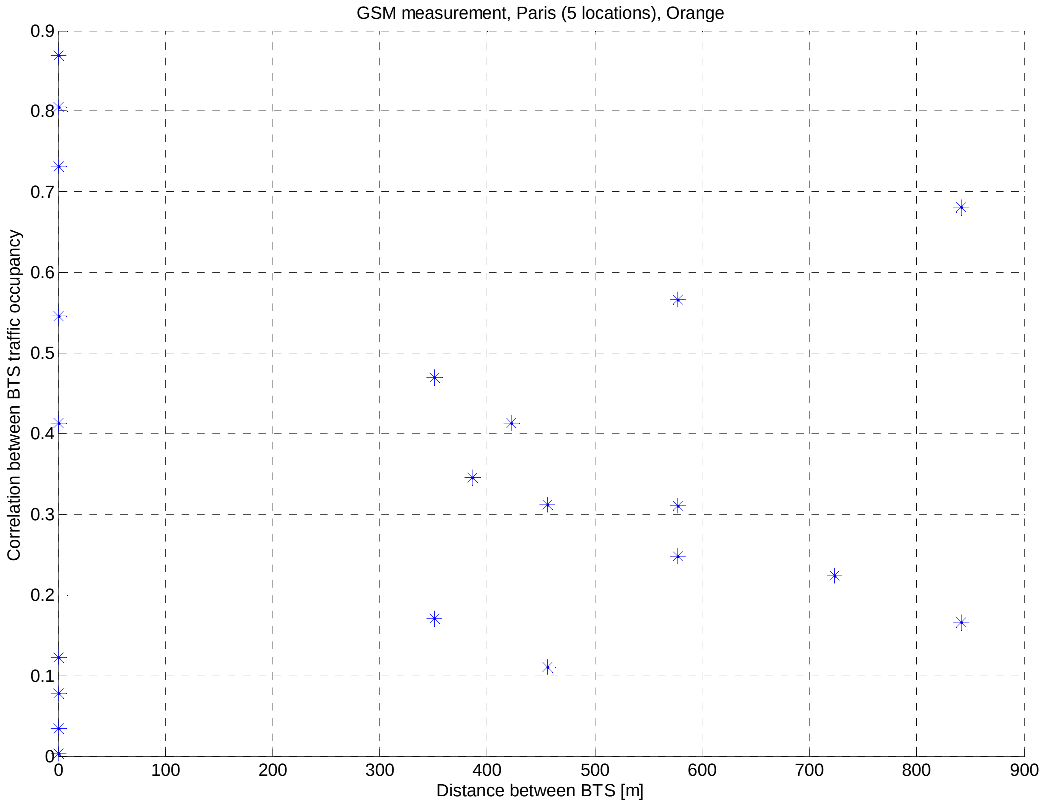


Figure 22 : Correlation between GSM BTS traffic occupancy versus inter-BTS distance, for the 5 locations in Paris, on Orange network

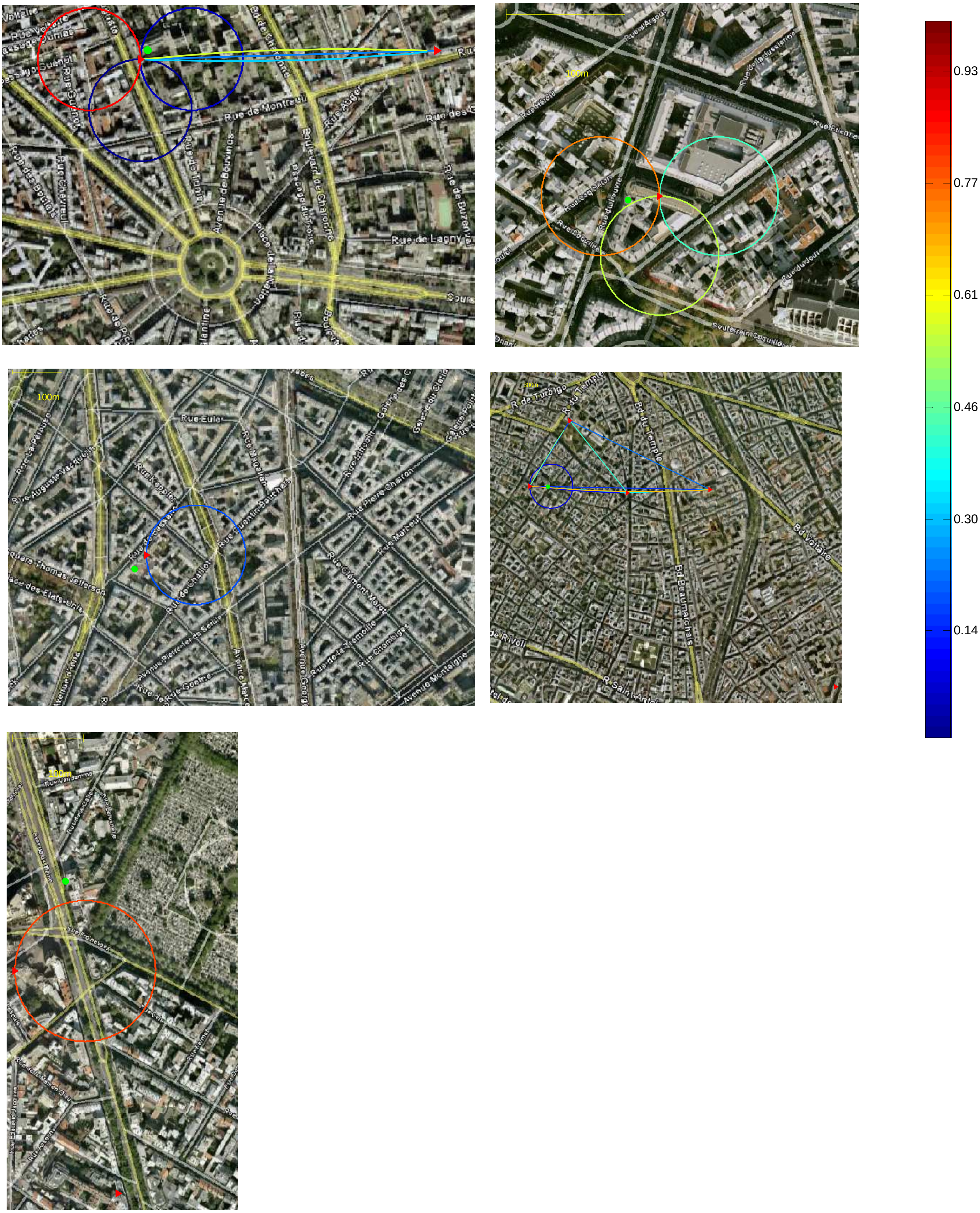


Figure 23 : Spatial correlation between BTS traffic occupancy, on Orange network, in Nation, Louvre, Etoile, Republique and Montparnasse

On the previous figures, we studied the correlation between BTS from the same district. In Figure 24, we are now interested by the correlation between BTS from different districts. It allows seeing if a correlation can exist between BTS from distant districts. In this figure different markers are used to differentiate the correlation between two given districts. The markers used are detailed in Table 6.

In Figure 24, the markers are gathered vertically. That is completely logical as the distance between two BTS from two given districts is quite constant. Except from this, no particular pattern pops up. The correlation values are quite low (mostly between 0.6 and 0.2). The mean correlation for each couple of districts is summarized in Table 7. These mean values are very low (roughly between 0.3 and 0.4).

For two given districts, the correlation values importantly vary, except between Louvre and Montparnasse where the correlations are more important and stay between 0.5 and 0.73. For the others couple of districts the range of the values is important.

In [2], we looked at the inter-district correlation i.e. at the correlation between the mean traffic occupancy of the BTS of a given district. We had much higher values (0.88 between Louvre and Republique, 0.83 between Nation and Republique, 0.82 between Louvre and Montparnasse). It means that, at a macroscopic level (the level of a district), the shape of the traffic occupancy from one district to another is quite constant. But at a microscopic level (the level of a BTS), the shape of the traffic from one BTS to another (from the same district or not) highly differs.

To have a reliable model for traffic occupancy, it is though necessary to be at a macroscopic level. The traffic occupancy at the BTS level varies in a too important way to be correctly modeled.

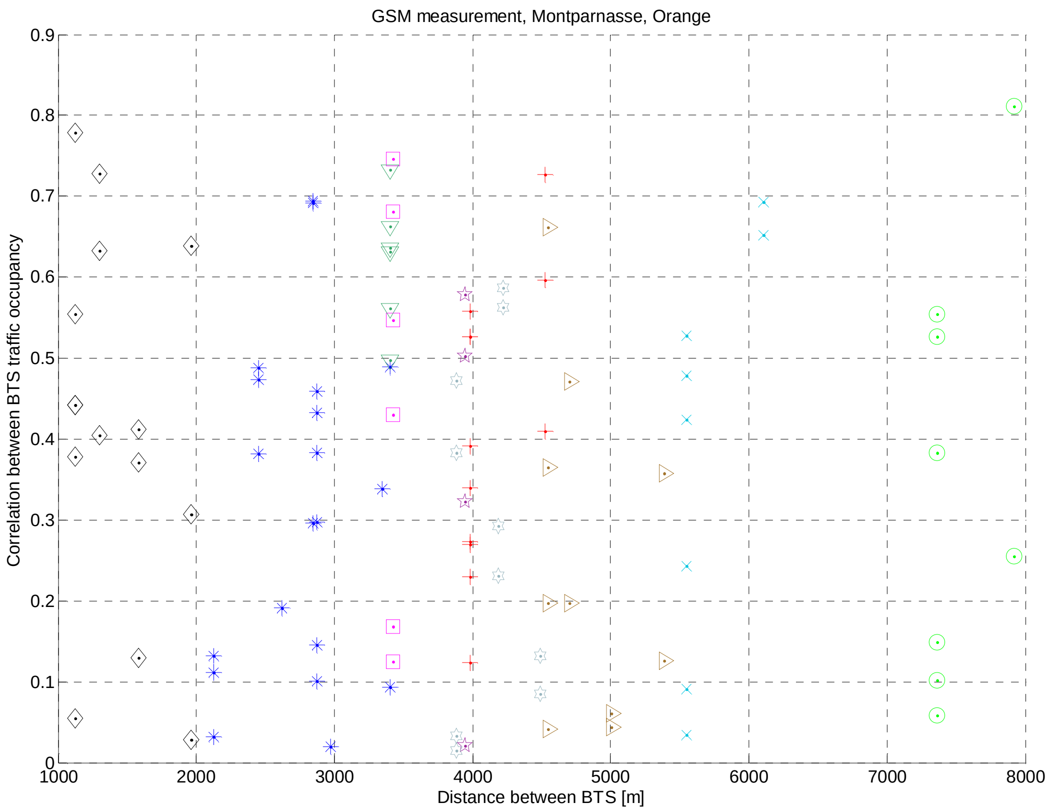


Figure 24 : Correlation between GSM BTS traffic occupancy versus inter-BTS distance, for BTS from different districts, on Orange network

	Nation	Louvre	Etoile	Republique	Montparnasse
Nation		+	O	*	x
Louvre			€	◊	∇
Etoile				▷	★
Republique					✶
Montparnasse					

Table 6 : Markers used in Figure 24

	Nation	Louvre	Etoile	Republique	Montparnasse
Nation		0.41	0.36	0.31	0.39
Louvre			0.45	0.42	0.62
Etoile				0.25	0.36
Republique					0.28
Montparnasse					

Table 7 : Mean correlation between GSM BTS from different districts, on Orange network

5.5 Conclusion

In this section, we studied the traffic occupancy correlation between BTS depending on their geographic location.

In Issy-les-Moulineaux, we pointed out that the highest correlations were seen for close BTS (the highest one for collocated BTS) and that, roughly, the correlation decreases when the inter-BTS distance increases. In tends to prove that, for this campaign, on areas covering one or two cells, the way GSM resource is used is similar. Nevertheless, these trends were not found in Colombes (except that the highest correlation is for collocated BTS) where it doesn't seem to have any link between the correlations values and the inter-BTS distance.

In Paris, no pattern seems to pop up. Looking at traffic correlations between BTS from the same district or from different districts, all kind of situations can be found (close BTS with low correlation, distant BTS with high correlation).

It is important to remind that, in [2], we computed the macroscopic traffic occupancy correlation between the different districts in Paris. We found high correlations (0.88 between Louvre and Republique, 0.83 between Nation and Republique, etc.). When computing the mean microscopic traffic occupancy correlations between districts the values are significantly lower (0.38 in average). It means that, in Paris, at a macroscopic level (the level of a district), the shape of the traffic occupancy from one district to another is quite constant. But at a microscopic level (the level of a BTS), the shape of the traffic from one BTS to another (from the same district or not) highly differs.

6 Conclusions

In this deliverable we proposed a consolidated model for characterizing traffic in Ile de France main Regions/areas.

The consolidated model is based on refined radio measurements which take into account each BTS traffic and hence offer the finest space scale granularity.

After refining the measurements, we proposed a model that takes into account the disparities between BTSs in the same region. The proposed model is a K-Gaussian model which adapts and fits to each measurement. After determining the parameters, a clustering method gives opportunity to find and merge similar BTSs.

Then, an analysis based the estimated of the correlations in the space scale between each BTS is performed. These estimations were done in Paris, Colombes and Issy areas and the traffic similarities were illustrated on maps.

All these results should be used in the simulator as an input to describe user traffic in large time scale which is particularly important as the simulator must include realistic models describing Ile de France region.

7 References

[1] URC Delivrable D2.2.2 “Measurements campaign in Ile-de-France report”.

[2] URC Deliverable D2.2.3 “Field measurements processing”

8 Acronyms

BTS	Base Transceiver Station
CI	Cell Identity
DTX	Disrupted Transmission
GPS	Global Positioning System
GSM	Global System for Mobile communications
LAC	Location Area Code
MCC	Mobile Country Code
MNC	Mobile Network Code
RF	Radio Frequency
UMTS	Universal Mobile Telecommunication System
URC	Urban planning for Radio Communications
VAD	Voici Activity Detection