

LE MODÈLE GRAVITAIRE EST-IL UN OUTIL ADAPTÉ À LA MODÉLISATION DES FLUX TÉLÉPHONIQUES ?

Pascal GILLON

UPRESA 6049 du CNRS - Théma
Université de Franche-Comté

Résumé

Le modèle gravitaire est un des principaux outils utilisés dans l'interaction spatiale pour appréhender les déplacements des personnes ou des biens. A partir de la masse respective des lieux et de la distance les séparant, il permet de reproduire assez fidèlement la réalité, grâce à l'optimisation de l'exposant de la fonction distance et d'un coefficient pondérant le produit des deux masses (modèle gravitaire simple).

Nous nous sommes donc servis de cet outil pour modéliser les flux téléphoniques entre les villes françaises. Il apparaît toutefois qu'il ne tient pas compte de façon explicite des occasions interposées, ni de l'effet de seuil suscité par la tarification (fonction distance continue).

Un nouveau modèle a alors été créé prenant en compte l'aspect tarifaire (fonction discrète) ainsi que les occasions interposées et intégrant une notion de fréquence des appels.

Mots-Clés

Distance - Flux téléphonique - Modèle gravitaire - Occasions interposées

Notre but est de mesurer le degré de pertinence du modèle gravitaire pour modéliser les flux téléphoniques entre les villes françaises. En effet, cet outil est le plus utilisé dans les problématiques apparentées à l'interaction spatiale dès lors qu'on traite de liens horizontaux tissés entre entités spatiales. Le modèle gravitaire, fruit d'une transposition d'un modèle physique, a surtout été employé pour traiter des problèmes de migrations alternantes ou définitives de population ainsi que pour les flux de marchandises, mais rarement pour les flux d'information. Toutefois, nous avons opté pour ce modèle et notre but est de voir s'il est apte à représenter le phénomène ou s'il faut recourir à un autre modèle.

Dans une première partie, nous vérifions que les données relatives au trafic téléphonique cadrent bien avec les deux principes du modèle gravitaire : l'interaction est proportionnelle au produit des deux masses et inversement proportionnelle au carré de la distance. Dans une deuxième partie, nous testons les performances du modèle gravitaire sans contrainte et nous introduisons des modifications afin de mesurer les améliorations possibles. Dans une dernière partie, nous proposons un modèle qui s'écarte sensiblement du modèle gravitaire, mais qui est construit à partir des enseignements tirés de l'étude de celui-ci.

1. Les flux téléphoniques vérifient-ils les hypothèses de base du modèle gravitaire ?

Nous avons tout d'abord vérifié que le total des flux émis et reçus par chaque zone à autonomie d'acheminement (ZAA : découpage technique de France Telecom s'apparentant aux bassins d'emploi) était proportionnel à leur masse, exprimée en nombre de lignes principales (LP). C'est bien le cas, comme l'indiquent les valeurs des coefficients de détermination qui expliquent toujours plus de 88 % de la variance (tab. 1).

Tableau 1 : Ajustements linéaires entre la masse et les flux émis et reçus pour le trafic national

Trafic national	Nombre de ZAA	Coefficient de corrélation R	Valeur de F	Valeur de t	Equation de la régression
Emis ZAA ₁	464	0,898	4139	***	$Y = 23 + 7,09 \cdot 10^{-3} X$
Reçu ZAA ₁	464	0,883	3482	***	$Y = -20 + 8,13 \cdot 10^{-3} X$

ZAA₁ = France moins Paris et sa petite couronne
X = nombre total de lignes
*** : non significatif < 1 %
F : Test de Fischer-Snedecor. Il indique la significativité ou l'absence de significativité de l'ensemble de la régression.
t : Test de Student. Il indique la significativité de chacune des variables explicatives.

Il s'agit ensuite de confirmer que les flux d'information sont également sensibles à la friction de la distance, ce qui peut apparaître paradoxal puisque ces techniques sont supposées abolir la distance. Toutefois, on n'utilise pas le téléphone au hasard, mais pour entrer en contact avec des connaissances, des partenaires économiques ou pour obtenir des biens ou services qu'on ne trouve pas ailleurs et c'est par ce biais que la dimension de la distance réapparaît.

Pour mesurer les flux émis en fonction de la distance, nous avons utilisé la technique de l'analyse radiale. La ville étudiée devient le centre de comptage sur lequel on superpose une grille formée de cercles concentriques dont l'écartement est défini par un pas kilométrique. On dénombre ensuite les villes positionnées sur le périmètre de chaque cercle (fig. 1). Pour chaque périmètre, le trafic (exprimé en erlangs) envoyé aux villes est sommé. On obtient ainsi une courbe que l'on qualifie de signal. Chaque ville présente un signal particulier, une « empreinte digitale », qui traduit sa position dans l'espace et l'intensité de son trafic avec les autres villes françaises (fig. 2).

Figure 1 : Les principales étapes de la constitution du signal

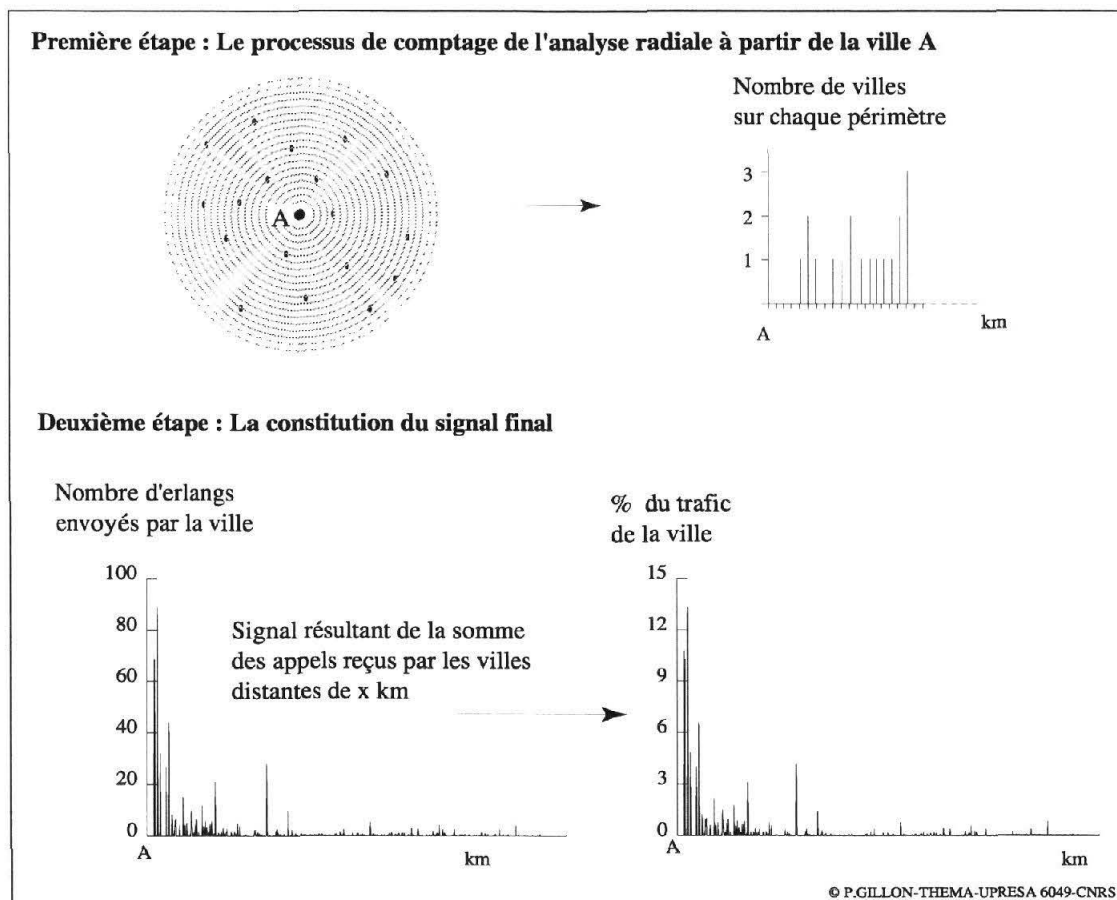
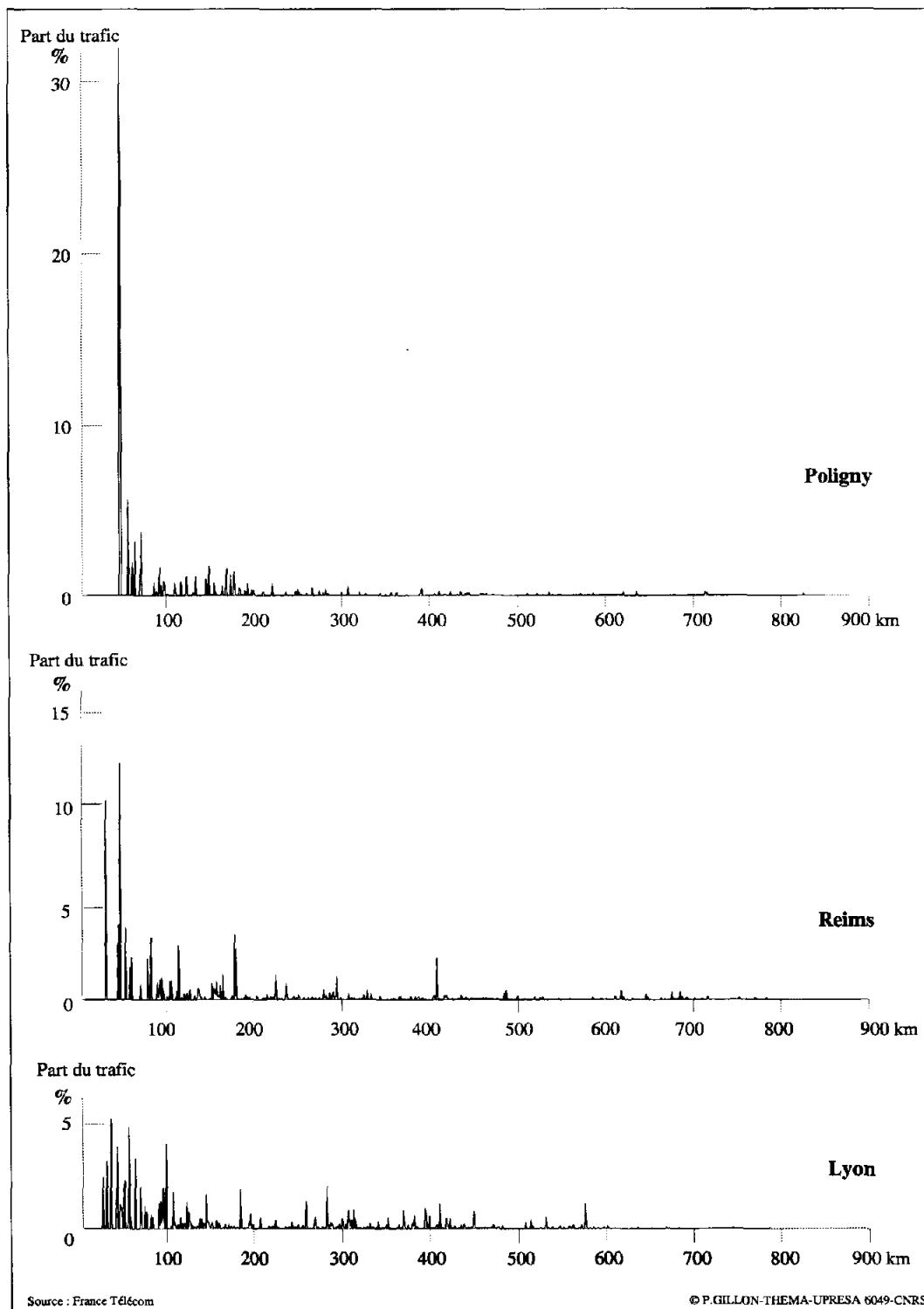


Figure 2 : Le signal, une empreinte spatiale distincte pour chaque ville



Cependant, les courbes présentent toutes la même forme générale. Le signal observé correspond au trafic émis divisé par le nombre de lignes de la ville recevant le flux. La forme de la courbe correspond à celle de la fonction $y = ax^{-b}$. Elle traduit la diminution de l'interaction entre deux lieux en fonction de la distance. La décroissance est fortement accentuée dans une zone proche (moins de 100 kilomètres), puis elle devient beau-

coup plus douce. Cette analyse rapide de la répartition du trafic en fonction de la distance met en évidence le frein porté à la communication par la distance.

Ce survol permet de montrer que les flux d'information répondent bien aux deux hypothèses de base du modèle gravitaire. Il s'agit maintenant de tester celui-ci et de mesurer ses performances.

2. Le modèle gravitaire sans contrainte

Le modèle gravitaire sans contrainte se présente sous la forme :

$$I_{ij} = \alpha M_i M_j d_{ij}^{-\beta}$$

avec I_{ij} : volume de trafic traduisant une interaction entre la ville i et la ville j

M_i : masse de la ville i

M_j : masse de la ville j

d_{ij} : distance entre la ville i et la ville j

α et β étant les deux paramètres à estimer (α la propension à communiquer et β l'intensité de la friction).

Les autres variables étant connues, les deux paramètres α et β sont à estimer. Pour ce faire, on peut opter pour plusieurs hypothèses.

La première serait qu'il existe un α et un β communs à toutes les villes, ce qui est peu réaliste et implique que les flux entre les entités i et j soient symétriques. La seconde serait qu'il existe pour chaque ville des paramètres différents. Ceci suppose de lever l'hypothèse d'un comportement homogène global (même comportement pour tous les abonnés) pour admettre un comportement homogène partiel (même comportement des abonnés de toute la ville) spatialement différencié. Pour cela, on admet que α n'est pas une constante générale, mais un paramètre de position. Cette hypothèse permet, de plus, d'éviter le phénomène de symétrie imposé par la première formule. En effet, il s'agit de calculer :

$$I_{ij} = \alpha_i M_i M_j d_{ij}^{-\beta_i} \quad \text{et} \quad I_{ji} = \alpha_j M_j M_i d_{ji}^{-\beta_j}$$

α_i est différent de α_j , ainsi que β_i de β_j , car la combinaison du trafic d'une zone à autonomie d'acheminement vers les autres est unique.

Une fois les calculs réalisés pour cette hypothèse, nous ferons varier la forme de la fonction traduisant la friction de la distance afin de voir si la fonction puissance inverse est la plus adéquate.

2.1. La méthode d'estimation des paramètres α et β

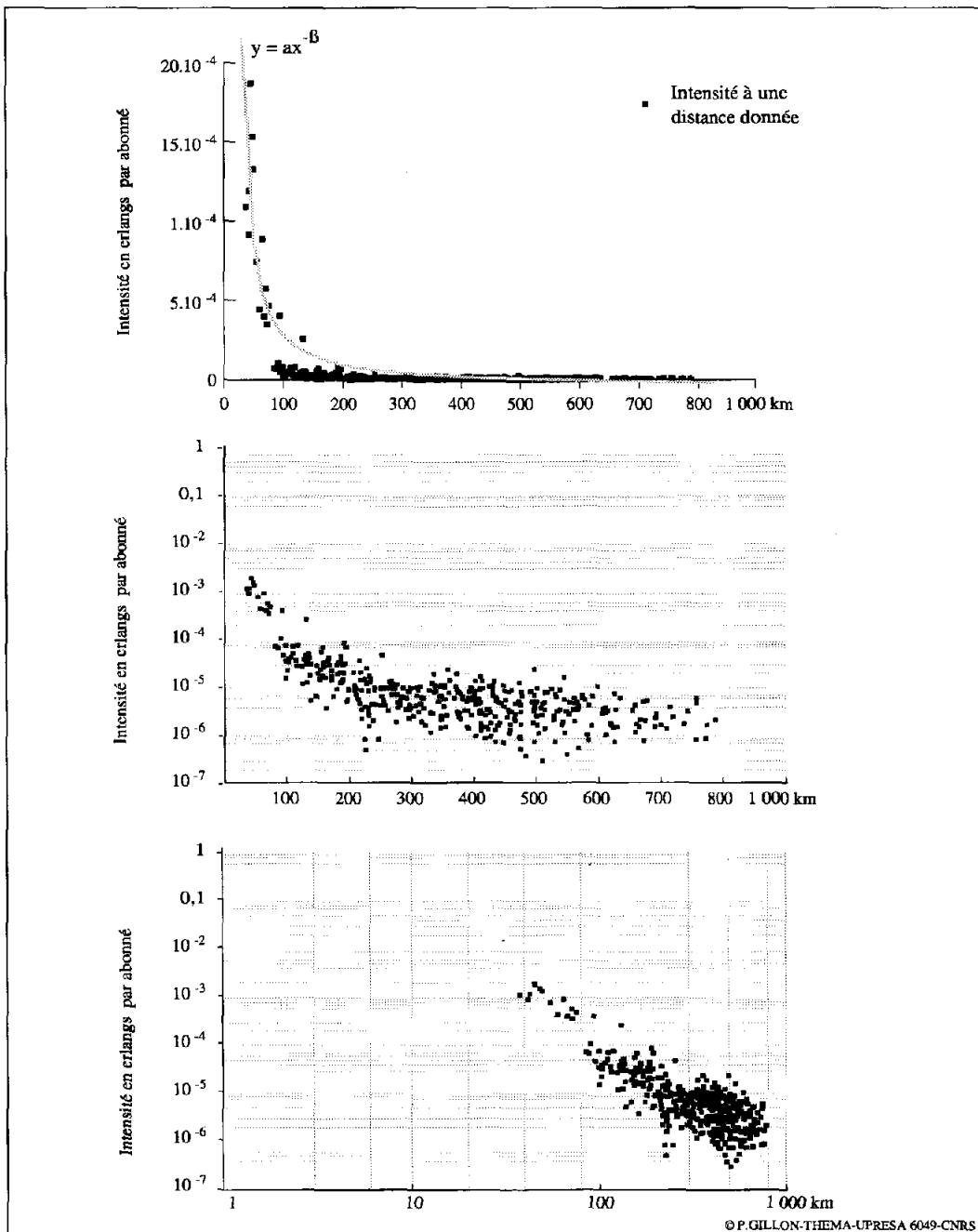
C'est celle de la régression linéaire qui a été retenue. Elle nécessite une transformation initiale des données. Il s'agit de redresser le nuage de points en opérant une transformation bi-logarithmique (fig. 3). Cette opération effectuée, il est possible de calculer les paramètres α et β .

$$I_{ij} = \alpha_i M_i M_j d_{ij}^{-\beta} \quad \text{soit} \quad \frac{I_{ij}}{M_i \cdot M_j} = \alpha_i \cdot d_{ij}^{-\beta} \quad \text{soit} \quad \log \frac{I_{ij}}{M_i \cdot M_j} = -\beta \cdot \log d_{ij} + \log \alpha_i$$

$$\text{soit} \quad \log y = -\beta \log x + \log \alpha_i$$

On pose $Y = \log y$, $X = \log x$ et $\alpha' = \log \alpha$, et on opère alors l'ajustement linéaire sur l'équation $Y = \beta X + \alpha'$.

Figure 3 : Transformation semi- puis bi-logarithmique du nuage de points original



2.2. La mesure de la performance

Une fois les paramètres α et β estimés, on calcule les valeurs théoriques de l'interaction entre les villes. Il reste ensuite à évaluer la qualité des valeurs théoriques calculées par le modèle. La somme des carrés des écarts donne une idée générale de sa valeur. Le principe de cette méthode est relativement simple. Il s'agit de calculer l'écart entre la valeur observée et la valeur théorique, puis de le porter au carré et d'en faire la somme sur tous les couples et sur différents intervalles de distance. Plus cette somme est faible, plus le modèle se rapproche de la réalité.

$$\sum_{c=1}^N (y_i - y_i^{\wedge})^2 \quad \text{avec } N \text{ nombre de couples observés}$$

2.3. Des performances d'estimation médiocres

Les paramètres α et β sont calculés pour chaque zone à autonomie d'acheminement. Nous avons ainsi calculé 444 régressions pour obtenir 444 α et 444 β . Les valeurs des tests de Fischer et de Student sont bonnes et le R^2 est significatif.

Toutefois, les performances sont médiocres puisque la valeur totale du trafic théorique émis est de 61 214 erlangs soit à peine 47,9 % du total réel ! Plus de la moitié du trafic des zones à autonomie d'acheminement est sous-estimé à plus de 50 %. Pour tous les intervalles de distance, le total du trafic est toujours inférieur à la réalité (tab. 2 et 3).

Tableau 2 : Répartition du trafic réel et calculé par le modèle gravitaire sans contrainte en fonction de la distance

En km	Trafic réel	% réel	Trafic théorique	% théorique
<15	4 230	3,31	1 715	2,80
15-25	13 466	10,54	5 230	8,54
25-50	35 675	27,92	9 827	16,05
50-100	29 320	22,95	11 884	19,41
100-250	23 770	18,60	16 489	26,94
250-500	13 560	10,61	11 309	18,47
500-750	6 447	5,05	4 089	6,68
>750	1 297	1,02	672	1,10
Total	127 768	100	61 214	100

Tableau 3 : Répartition des couples suivant l'écart (en points) entre valeurs théoriques du modèle gravitaire sans contrainte et valeurs estimées, selon les intervalles de distance

Intervalles en km	Valeur théorique < valeur réelle Sous-estimation (en points)				Théorique = réelle 0	Valeur théorique > valeur réelle Surestimation (en points)				
	<5	4,9 / 1	0,9 / 0,5	0,4 / 0,1		0,1 / 0,4	0,5 / 0,9	1 / 4,9	>5	
<15	23	18	3	2	0	0	0	6	5	
15-25	133	118	17	13	0	17	8	34	13	
25-50	474	539	158	202	3	277	149	185	7	
50-100	230	490	315	998	55	3 485	961	535	4	
100-250	45	256	228	2 576	496	32 661	2 101	742	1	
250-500	11	73	153	6 522	4 305	76 258	702	141	0	
500-750	5	48	67	6 332	6 328	44 278	80	6	0	
>750	0	4	25	1 303	6 328	7 268	2	0	0	
Total	921	1 546	966	17 945	17 515	164 244	4 003	1 649	30	

Les trames grisées indiquent la part des couples sur ou sous-estimés par rapport au total des couples existants dans l'intervalle de distance donné

Pas de couple pour l'intervalle de distance	De 5 à 9,9 %
Moins de 1 %	De 10 à 25 %
De 1 à 4,9 %	Plus de 25 %

2.4. Un apport important du point de vue de la connaissance du modèle

Au-delà d'une performance médiocre, ce paramétrage du modèle apporte certains enseignements à propos de la fonction distance et du rôle de l'exposant β .

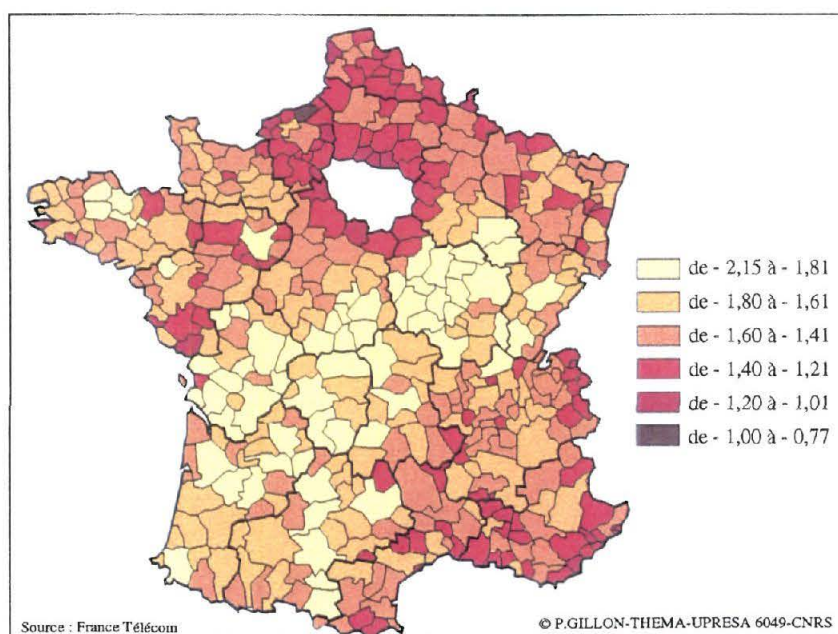
La «dureté» de la friction de la distance varie assez fortement entre les différentes zones à autonomie d'acheminement puisque β oscille entre 2,152 et 0,766. Les valeurs du coefficient peuvent s'interpréter de la façon suivante :

- les zones à autonomie d'acheminement ayant les coefficients les plus élevés (2 et plus) sont caractérisées par une forte friction de la distance et donc par une intensité de communication qui décroît rapidement. Elles auraient ainsi tendance à lier des contacts étroits avec des partenaires proches ;
- les zones à autonomie d'acheminement ayant les coefficients les plus bas sont définies par une faible friction de la distance et donc par une intensité de communication qui décroît lentement. Elles auraient donc des contacts avec des partenaires plus éloignés que les précédentes.

La répartition des valeurs de β présente une configuration spatiale nette (fig. 4), divisant la France en quelques grands ensembles. Cette distribution des valeurs de β semble prendre en compte, au moins en partie, celle de la densité de la population en France.

Ainsi, les villes situées dans un espace de faible densité ont tendance à communiquer plus intensément avec leurs proches voisins. Cet isolement relatif limiterait les échanges, la distance à franchir étant trop importante. Les villes positionnées dans un espace de forte densité entrent plus facilement en contact avec des partenaires un peu plus éloignés, mais peut-être plus accessibles.

Figure 4 : Valeur du coefficient β pour le modèle gravitaire sans contrainte



Cette influence de l'environnement autour de la ville a déjà été remarquée par de nombreux auteurs. Ils ont observé que la géométrie du système spatial avait une influence sur le coefficient β , comme le montre l'exemple donné par R.J. JOHNSTON [4]. La seule densité autour du point émetteur déforme l'espace et modifie donc la fonction distance.

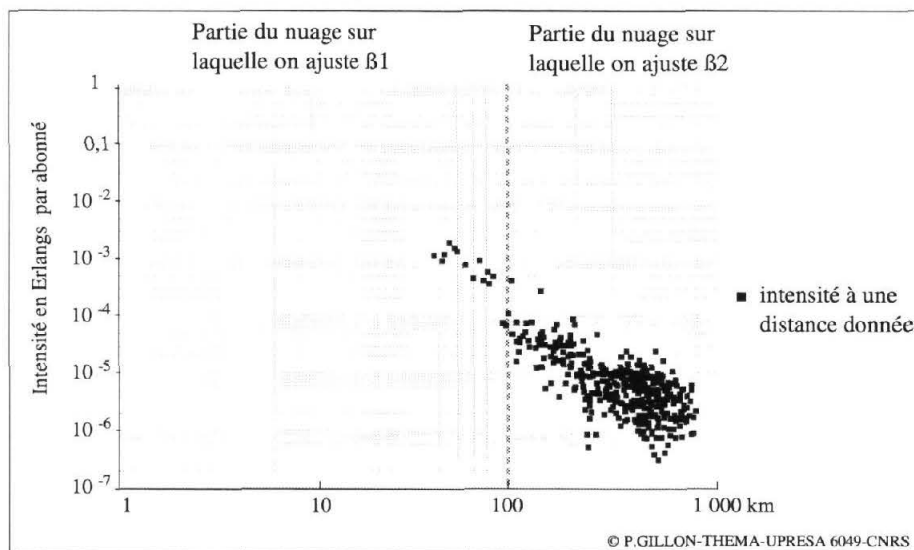
3. Un modèle gravitaire avec deux fonctions "distances"

La traduction de la friction de la distance n'est pas bonne dans le premier modèle puisque la répartition du trafic en fonction de la distance ne correspond visiblement pas à la réalité. Les performances du modèle sont peut-être mauvaises simplement parce que l'ajustement des paramètres α et β n'est pas optimal. Nous allons, par conséquent, ajuster les paramètres par une nouvelle approche.

Lorsqu'on examine les nuages de points traduisant la relation entre l'intensité du trafic et la distance, on s'aperçoit que la transformation bi-logarithmique ne remplit pas totalement son office. En fait, le nuage est souvent bien redressé, mais on remarque que, dans une première partie, il est assez allongé et que, dans la seconde (une majorité de points), la forme est moins étirée.

Il est possible de segmenter un nuage de points pour mieux l'ajuster. Nous avons donc procédé de la sorte en le partageant en deux parties (fig. 5). Il est bien sûr possible de le découper en plusieurs parties, mais chaque segment doit être ajusté. La détermination du seuil entre les segments pose le problème du nombre de points les constituant. En effet, il faut au moins une trentaine de points pour pouvoir réaliser une régression. Le seuil ne doit donc pas être trop proche de l'émetteur afin de satisfaire à cette condition. C'est pourquoi, nous avons retenu la limite des 100 km. Ce choix est également motivé par le fait qu'à 100 km, 50 % des flux sont déjà écoulés pour toutes les zones à autonomie d'acheminement.

Figure 5 : Partition du nuage bi-logarithmique



3.1. Une amélioration des performances...

Cette nouvelle approche améliore sensiblement les performances du modèle. Le volume global théorique de flux émis est de 106 426 erlangs soit 83,3 % du total réel. De plus, la répartition du pourcentage qui mesure l'écart entre la valeur réelle et la valeur théorique a bien changé. Il n'y a plus que 23 zones à autonomie d'acheminement contre 256 qui ont un total inférieur à 50 % de la valeur réelle.

Tableau 4 : Répartition du trafic réel et calculé par le modèle gravitaire sans contrainte (avec 2 fonctions distance)

En km	Trafic réel	% réel	Trafic théorique	% théorique
<15	4 230	3,31	7 285	6,84
15-25	13 466	10,54	23 349	21,94
25-50	35 675	27,92	27 572	25,91
50-100	29 320	22,95	19 864	18,66
100-250	23 770	18,60	12 288	11,55
250-500	13 560	10,61	10 785	10,13
500-750	6 447	5,05	4 494	4,22
>750	1 297	1,02	790	0,74
Total	127 768	100	106 426	100

Pour la répartition du trafic en fonction de la distance, les flux de proximité calculés (moins de 25 km) sont deux fois plus importants que les flux réels ! Les flux émis à moins de 50 km représentent 54,69 % du total des flux théoriques contre 41,77 % dans la réalité. On retrouve ensuite une situation plus classique avec une sous-estimation du reste du trafic. Toutefois, les deux classes pivots autour des 100 km sont les plus fortement sous-estimées (tab. 4 et 5).

Tableau 5 : Répartition des couples suivant l'écart (en points) entre valeurs théoriques du modèle gravitaire sans contrainte (deux fonctions distance) et valeurs estimées, selon les intervalles de distance

Intervalles en km	Valeur théorique < valeur réelle Sous-estimation (en points)				Théorique = réelle 0	Valeur théorique > valeur réelle Surestimation (en points)			
	<5	4,9 / 1	0,9 / 0,5	0,4 / 0,1		0,1 / 0,4	0,5 / 0,9	1 / 4,9	>5
<15	4	8	3	3	0	1	1	8	29
15-25	32	38	10	11	0	10	15	74	163
25-50	275	337	95	162	4	206	184	527	204
50-100	205	488	320	1 418	110	3 201	729	577	22
100-250	96	533	590	10 305	2 305	25 047	188	42	0
250-500	11	121	243	14 093	7 909	65 615	153	20	0
500-750	5	49	98	9 683	7 473	39 802	30	4	0
>750	0	6	27	1 850	7 473	6 408	0	1	0
Total	628	1 580	1 386	37 525	25 274	140 290	1 300	1 253	418

Les trames grisées indiquent la part des couples sur ou sous-estimés par rapport au total des couples existants dans l'intervalle de distance donné

Pas de couple pour l'intervalle de distance	De 5 à 9,9 %
Moins de 1 %	De 10 à 25 %
De 1 à 4,9 %	Plus de 25 %

3.2. ...mais une approche quelque peu limitée

L'approche des échanges des flux de communication par le modèle gravitaire sans contrainte apparaît cependant limitée.

Tout d'abord, la fonction puissance, censée traduire la dureté de la friction de la distance, est en fait plus complexe. L'exposant prend en compte, de façon implicite, la position de la zone à autonomie d'acheminement émettrice dans l'espace français. Ce manque de transparence est gênant. Ensuite, la technique de la régression ne permet pas d'introduire plusieurs paliers pour les distances proches, puisque le nombre de points est trop faible pour obtenir un paramètre β significatif. Or, nous avons constaté que l'introduction d'un seuil à 100 km apporte un gain global considérable dans l'ajustement. Nous avons pu voir également que la distance proche n'était pas encore bien calibrée, avec une forte surestimation de nombreux couples.

Le seul véritable apport du modèle gravitaire a été de montrer toute l'importance qu'il faut accorder à la position de l'émetteur et à son environnement. L'examen des paramètres β a souligné le rôle des occasions interposées mises en lumière par E. ULLMAN [9]. Dans la modélisation il faut prendre en compte la densité de villes entre l'émetteur et le récepteur et donc la répartition spatiale.

4. Vers un modèle plus explicatif

Pour ce faire, nous proposons un modèle qui pourrait en partie corriger les problèmes inhérents au modèle gravitaire. Nous avons pris en compte la tarification et les occasions interposées, ce qui transforme complètement la philosophie du modèle gravitaire avec notamment le passage d'une fonction distance continue à une

fonction discrète pour représenter la distance. Les variables étant interdépendantes, le remplacement de l'une a un effet direct sur les autres. Ce changement de représentation de la distance entraîne d'autres modifications, notamment celle des masses.

4.1. Une fonction distance tenant compte de la tarification et de la perception

Le choix de la fonction distance, c'est-à-dire du terme $d_{ij}^{-\beta}$ dans le modèle gravitaire, doit tenir compte plus explicitement du coût entraîné par la taxation. En effet, l'écart de prix pour les appels entre le proche et le lointain est tel qu'il existe une différence de comportement des abonnés. Le coût d'appel entre 25 et 50 km est cinq fois moins élevé que celui d'un appel à plus de 100 km.

Mais la fonction doit aussi intégrer la perception de la distance par les usagers, notamment l'aspect logarithmique. Il s'agit donc de déterminer la forme de la fonction qui traduit en même temps l'effet tarifaire (paliers) et la perception de l'individu. La distance kilométrique à vol d'oiseau est conservée comme unité, tout comme la géométrie euclidienne, puisque la tarification est explicitement basée sur celle-ci.

4.1.1. L'intégration de la tarification

La tarification fonctionne par paliers, c'est-à-dire que le prix d'une communication ne varie pas à l'intérieur des intervalles de distance définis par l'opérateur, mais qu'il est différent entre ceux-ci. Le coût d'un appel augmente en fonction de la distance et renforce en quelque sorte le frein naturel lié à celle-ci. L'opérateur a défini cinq tarifs délimités par cinq distances. Les 100 premiers kilomètres sont ainsi coupés en quatre segments, ce qui était impossible à réaliser avec la technique d'ajustement choisie dans le modèle gravitaire.

Pour maintenir le diviseur constant entre les seuils tarifaires, la solution la plus simple est de prendre le centre de la classe et de lui appliquer un exposant β . Ainsi, pour toutes les liaisons qui se tissent entre 50 et 100 km, on applique un β à la valeur 75. Il s'agit ensuite de définir la valeur de β pour chaque classe. En fait, celle-ci est fixée dès que le β de la première classe est choisi. En effet, pour tenir compte de l'effet tarifaire, il faut que le dénominateur d_{ij}^{β} respecte les rapports de prix entre les différentes distances. Une fois le premier exposant déterminé, les autres sont calculés en fonction de lui.

La première classe comprend les relations entre deux villes à moins de 15 km. Le centre de la classe a été choisi à 10 km. L'exposant β a été fixé à 4. Cette valeur est directement tirée de l'observation des ajustements menés pour le modèle gravitaire. Le dénominateur est donc égal à :

$$10^4 = 10000$$

Ainsi toutes les liaisons de moins de 15 km sont divisées par 10 000.

Les autres exposants se calculent alors de la façon suivante :

– le «saut tarifaire» entre le tarif à moins de 15 km et celui appliqué entre 15 et 25 km est égal à :

$$\frac{0,61}{0,12} = 5,08$$

– le rapport entre le dénominateur appliqué pour les liaisons de moins de 15 km et celui utilisé pour celles entre 15 et 25 km doit ainsi être de 5,08. Il faut donc déterminer l'exposant qui remplit la condition :

$$10000 * 5,08 = 50800$$

– il suffit de calculer l'exposant qui donne 50 000 pour la distance de 20 km, centre de la classe :

$$20^{3,6} \approx 50000$$

– l'opération est ensuite reconduite pour les autres paliers tarifaires (tab. 6).

Tableau 6 : Calcul des exposants pour traduire les différents paliers tarifaires

	< 15 km	15 à 25 km	25 à 50 km	25 à 50 km	> 100 km
Centre de la classe	10	20	35	75	400
Tarif pour une minute de communication (en francs)	0,12	0,61	0,98	1,82	3,13
Rapport entre les seuils tarifaires	—	5,08	1,6	1,85	1,69
Valeur du dénominateur	10 000	50 000	80 000	150 000	255 000
Valeur de l'exposant β	4	3,6	3,18	2,76	2,08

4.1.2. L'intégration de la perception

La forme de la fonction de la tarification traduit, dans l'ensemble, les phénomènes de perception. En effet, l'espace proche est fortement discriminé, avec une hausse assez importante entre les paliers, et l'espace lointain n'est plus du tout différencié. C'est à ce niveau qu'une critique peut être émise. Au-delà de 100 km, la tarification ne fait plus aucune différence, alors que l'espace peut être encore discrétisé au-delà, notamment en ce qui concerne les déplacements ; la propension à se déplacer ne sera pas la même pour une distance de 150 ou de 700 km. Nous avons donc introduit des seuils à 250 km, 500 km et 750 km selon une progression géométrique pour les deux premiers seuils. La dernière valeur (750 km) a été retenue aussi pour pouvoir comparer les résultats avec les modèles gravitaires précédents. La rupture entre chaque seuil est faible, puisqu'une fois la décision de se déplacer prise, le supplément de distance à parcourir n'intervient pas de façon cruciale. Les valeurs du dénominateur ne sont donc pas trop différentes (tab. 7). On obtient ainsi une fonction qui prend en compte le coût des appels ainsi que la perception des personnes. La forme de la fonction inverse traduit l'effet combiné de la tarification et de la perception sur la communication (fig. 6).

La fonction distance ainsi définie n'a pas à être ajustée puisque les valeurs de l'exposant β sont fixées une fois pour toute. L'intérêt de cette démarche est d'obtenir une fonction unique pour toutes les villes sans recourir à la phase d'ajustement. De plus, cette fonction est construite selon des hypothèses clairement identifiées. Le rôle de la tarification et celui de la perception sont intégrés de façon réaliste.

Tableau 7 : Calcul des exposants pour traduire la fonction tarifaire et la perception

	<15 km	15 à 25 km	25 à 50 km	50 à 100 km	100 à 250 km	250 à 500 km	500 à 750 km	>750 km
Centre de la classe	10	20	35	75	175	375	675	850
Tarif pour une minute de communication (en FHT)	0,12	0,61	0,98	1,82	3,13	3,13	3,13	3,13
Rapport entre les seuils tarifaires	—	5,08	1,6	1,85	1,69	1,15	1,1	1,06
Valeur du dénominateur	10 000	50 000	80 000	150 000	255 000	300 000	330 000	350 000
Valeur de l'exposant β	4	3,6	3,18	2,76	2,08	2,15	1,95	1,87


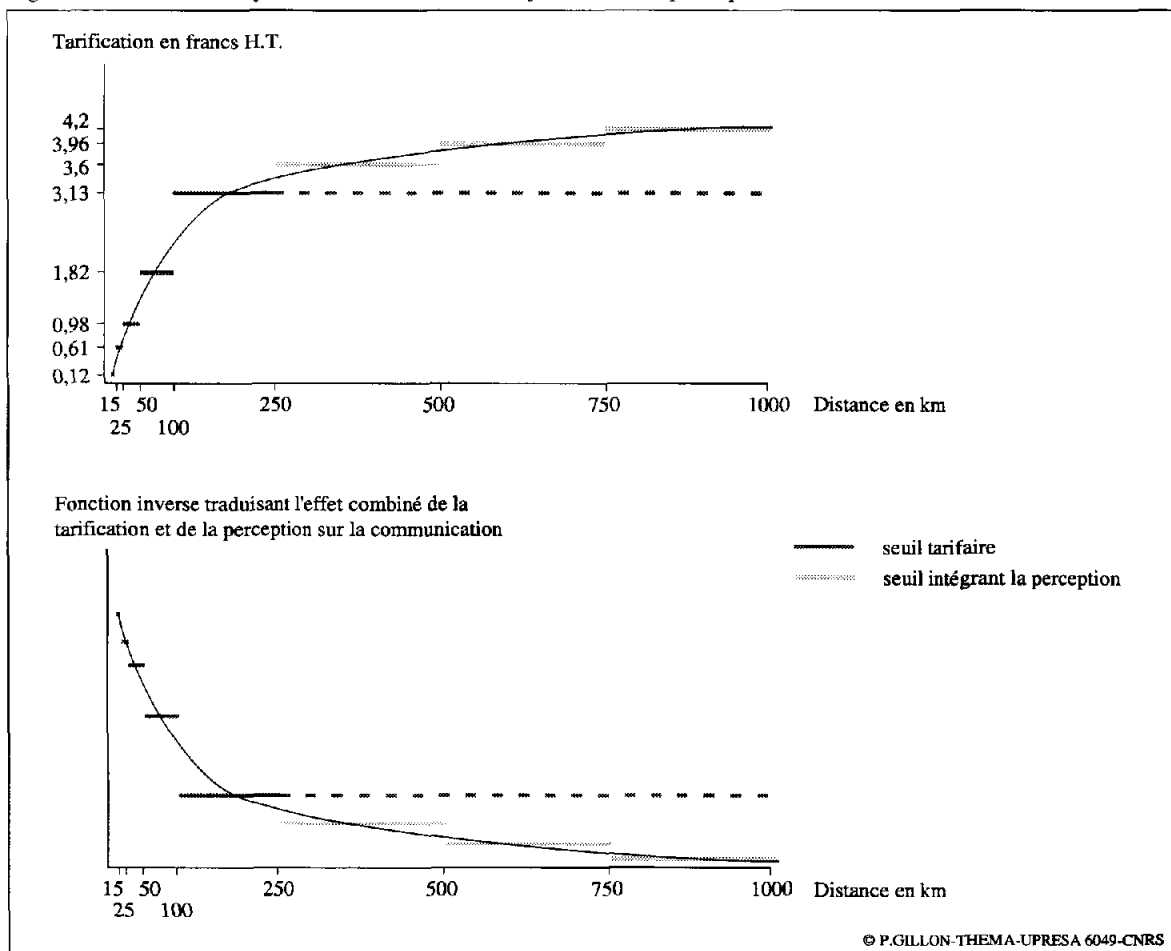
 Introduction de la perception dans la fonction

Figure 6 : Forme de la fonction traduisant la tarification et la perception



4.2. Un correctif à la fonction distance : la fréquence d'appels

La fonction distance introduit les effets dus à la tarification et à la perception. Toutefois, pour que le modèle traduise la réalité, il faut tenir compte de la fréquence des appels en liaison avec la distance, celle-ci étant influencée par les occasions interposées. Nous allons donc introduire une variable qui sert à l'ajustement du modèle et qui tient compte de ces deux aspects.

4.2.1. La fréquence d'appels : un correctif à l'effet tarifaire

La structure tarifaire a pour effet de favoriser les appels proches et de pénaliser les appels à plus de 100 km ce qui se traduit par une intensité de trafic forte dans l'espace proche et faible dans l'espace distant. Or, le simple produit des deux masses en fonction de la distance tarifaire ne traduit pas complètement le phénomène. Il faut, par conséquent, introduire un coefficient correcteur qui donne un volume émis plus que proportionnel au produit des deux masses dans l'espace proche et moins que proportionnel dans l'espace lointain. Ce coefficient varie en fonction de la distance qui sépare les deux villes en relation. Il prend, de ce fait, une valeur différente pour chaque intervalle.

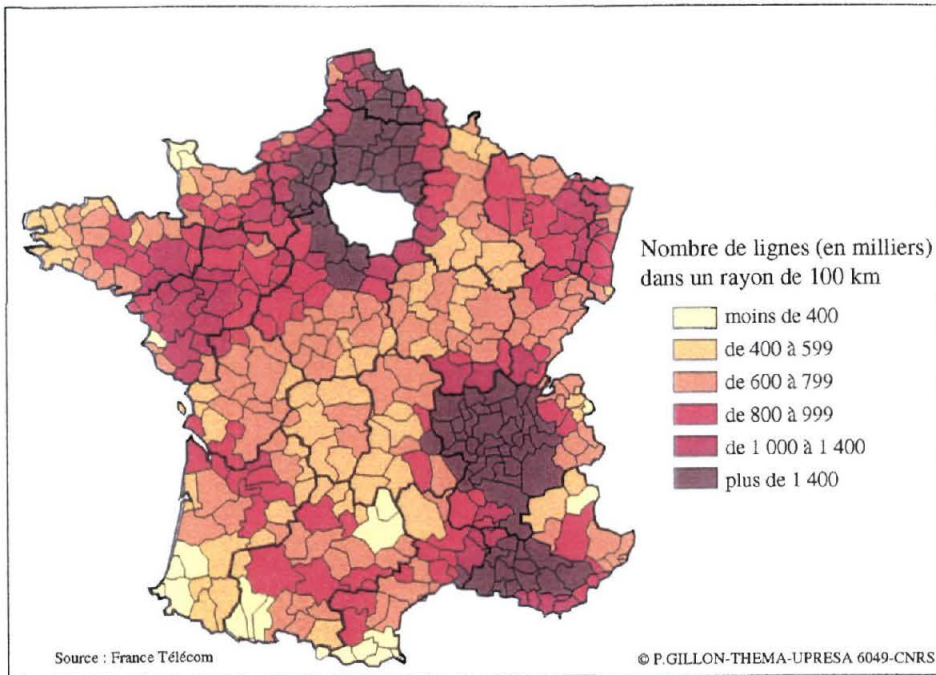
4.2.2. Les occasions interposées : une distorsion de l'espace

Ce coefficient varie en fonction de la distance par intervalle. Mais nous avons vu que la distance subit une distorsion selon le nombre et la taille des villes qui se situent dans l'espace. Les variations du paramètre β , dans le modèle gravitaire calculé par régression, ont montré que la position de la ville dans l'espace n'est pas neutre. Nous avons donc introduit explicitement l'hypothèse des occasions interposées.

Pour mesurer celles-ci, nous avons compté le nombre de lignes présentes dans un rayon de 100 km autour de la ville émettrice, et ce pour chaque zone à autonomie d'acheminement. Nous avons retenu le seuil de 100 km, puisqu'il correspond toujours à la distance à laquelle plus de 50 % du trafic est écoulé.

Pour simplifier le paramétrage, le nombre de lignes a été discrétisé en six classes. Une première regroupe les zones à autonomie d'acheminement entourées par moins de 400 000 lignes principales dans un rayon de 100 km, soit une densité inférieure à 12 lignes au km². Les seuils sont ensuite fixés à 600 000, 800 000, 1 000 000 et 1 400 000 lignes, soit des densités de 18, 25, 30 et plus de 42 lignes/km² (fig. 7). Cette distribution spatiale de la densité des lignes n'est pas sans rappeler la carte représentant les valeurs du β pour le deuxième modèle gravitaire.

Figure 7 : Densité calculée autour des zones à autonomie d'acheminement dans un rayon de 100 km



On pose ensuite l'hypothèse selon laquelle plus la densité est faible autour de la ville émettrice et plus le trafic est concentré dans l'espace proche, c'est-à-dire que la ville augmente sa fréquence d'appels. L'introduction de l'effet de la densité se fait donc par l'intermédiaire de la fréquence d'appels qu'on dénomme f . Celle-ci varie ainsi en fonction des intervalles de distance d'une part, et de la densité autour de la ville émettrice d'autre part. La fonction f doit être calibrée pour les deux variables, la densité et la distance.

4.3. La formulation du modèle

L'équation du modèle s'écrit de la façon suivante :

$$I_{ij} = f_{(d_{ij}, O_i)} \cdot \frac{I_i \cdot I_j}{d_{ij}^\beta}$$

avec : I_{ij} : volume émis entre les villes i et j (en erlangs)

I_i : volume total émis de la ville i

I_j : volume total reçu par la ville j

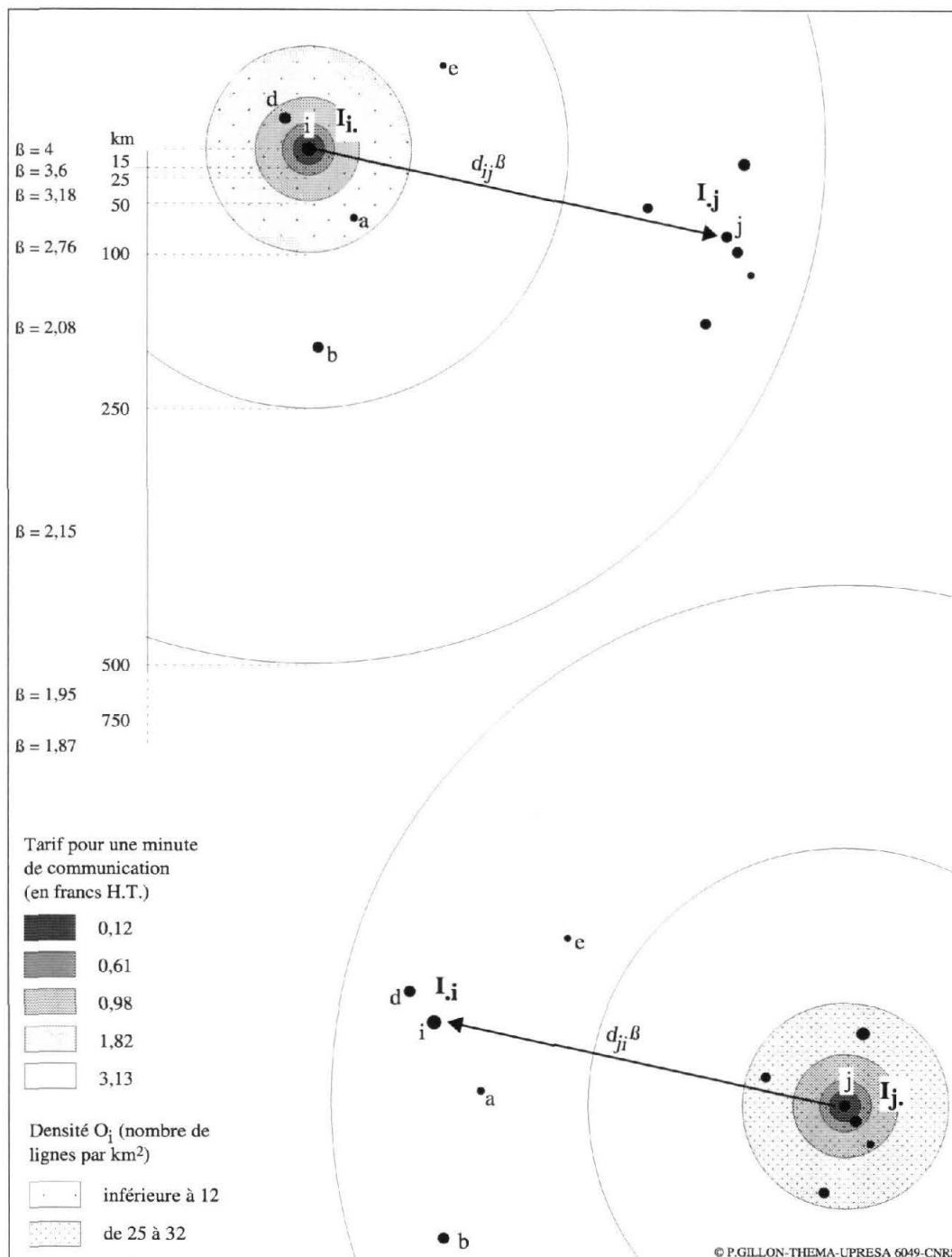
d_{ij} : distance entre les villes i et j , ramenée au centre de l'intervalle auquel elle appartient, intervalle défini par la tarification

β : exposant dont la valeur varie en fonction des tarifs

$f_{(d_{ij}, O_i)}$: fréquence de communication qui varie en fonction de la distance entre les deux villes (d_{ij}) et de la densité autour de l'émetteur (O_i).

Cette équation peut se traduire par un schéma (fig. 8).

Figure 8 : Représentation schématique du modèle



Le volume de flux émis entre deux villes est défini par le produit des deux masses, divisé par la fonction de tarification et corrigé par un paramètre qui varie suivant la distance kilométrique et la densité autour de l'émetteur. C'est ce paramètre qu'il faut déterminer puisque c'est la seule inconnue dans l'équation. Soulignons que l'exposant β , qui exprimait la rugosité de l'espace dans le modèle gravitaire, disparaît au profit de la tarification. Il est complété par l'intégration des occasions interposées.

Ce modèle ne partage plus la philosophie du modèle gravitaire classique. En effet, il tente d'être avant tout explicatif et se base sur des hypothèses explicites.

On retrouve le produit des masses des deux villes, qui est le moteur de la communication, et le frein dû à la distance. Cependant, la fonction distance est ici plus complexe. Au dénominateur, elle traduit le rôle de la tarification, qui est un outil de contrôle et de régulation du trafic, entre autres, pour l'opérateur. Sa fonction mathématique met en avant un espace fortement discriminé entre le proche et le lointain, traduisant ainsi la perception logarithmique de l'espace. De plus, elle prend en compte le coût qui est une variable incontournable et fortement incitative. Celle-ci provoque une adaptation du consommateur qui a tendance à communiquer plus dans l'espace proche sous-tarifé. La tarification influe ainsi sur le processus de communication.

Toutefois, introduire la tarification n'est pas suffisant pour expliquer complètement le rôle de l'espace. C'est pourquoi nous faisons appel à un paramètre d'ajustement, prenant en compte explicitement la tendance qui consiste à téléphoner plus dans l'espace proche (déviations dues à la tarification), ainsi que la position de la ville dans l'espace. Cette notion de densité de lignes autour du point émetteur se rattache aux occasions interposées.

Bibliographie

- [1] FUSTIER B., 1979 : *Les interactions spatiales en économie*, Paris, Sirey
- [2] GILLON P., 1997 : *Contribution à l'analyse des échanges interurbains : modélisation des flux téléphoniques entre les villes françaises*, Université de Franche-Comté, Thèse de Doctorat en Géographie, Besançon, 396 pages
- [3] HURIOT J.M., PERREUR J., 1990 : « Distances, espaces et représentations : une revue », *Revue d'économie régionale et urbaine*, n° 2, pp. 195-237
- [4] JOHNSTON R.J., 1976 : « On regression coefficients in comparative studies of the frictions of distance », *Tijdschrift voor economische en sociale geografie*, Vol. 67, pp. 15-28
- [5] LINNEMAN H.V., 1966 : *An econometric study of international trade flows Amsterdam*, North Holland Publishing Co
- [6] NIJKAMP P., RIETVELD P., SALOMON I., 1990 : « Barriers in spatial interactions and communications. A conceptual exploration », *Annals of regional science*, pp. 237-252
- [7] SCHNEIDER M., 1959 : « Gravity models and trip distribution theory », *Papers and proceedings of the regional science association*, n° 5, pp. 51-56
- [8] THOUVENOT G., 1983 : *L'espace géographique des adolescents*, Thèse IIIe cycle, Université de Lyon II
- [9] ULLMAN E., 1980 : *Geography as spatial interaction*, Seattle, University of Washington Press, 231 pages