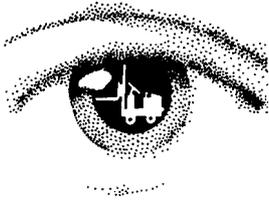


LA LOCALISATION DES ENTREPÔTS

Christian CALZADA, Hakim HAMMADOU,
Hubert JAYET, Sabine KAZMIERCZAK*



La répartition spatiale de la population est le déterminant fondamental de la géographie des implantations de locaux de stockage. C'est ce qui ressort d'une analyse dynamique de la localisation des entrepôts de stockage créés dans le Nord - Pas-de-Calais entre 1995 et 2001. Les créations de locaux de stockage sont, à densité égale, proportionnelles à la superficie de la zone d'installation. Néanmoins, pour une zone donnée, une croissance de la densité de population entraîne une croissance moins que proportionnelle du nombre de locaux créés. Alors que l'analyse ponctuelle ne permet pas d'appréhender précisément le rôle de l'accessibilité aux pôles urbains et aux infrastructures de transport, la méthode économétrique va plus loin, en confirmant le moindre rôle de l'accessibilité aux infrastructures par rapport au facteur population.

Le choix en matière d'implantation des entrepôts de stockage résulte de plusieurs types d'arbitrages : un arbitrage entre la proximité du bassin de population à desservir (clients) et la proximité aux industriels (fournisseurs), entre l'accessibilité aux infrastructures et le prix du foncier, etc. Du point de vue de l'aménageur, deux questions se posent. Quel est le rôle des infrastructures sur le choix de localisation ? Quels « retours sur investissements » la collectivité est-elle en droit d'attendre ? Ces questions sont examinées à travers le cas de la région Nord - Pas-de-Calais, dont on analyse les constructions récentes de locaux de stockage.

Forte densité de grands locaux de stockage en région Nord - Pas-de-Calais

Contrairement aux locaux de stockage de surface modeste, répartis de manière homogène sur le territoire français, les locaux dépassant 1 500 m² sont implantés de façon plus dense dans certaines régions. Ainsi, les grands locaux de stockage (plus de 10 000 m²), qui ont connu une forte croissance au cours des années 90 (+ 9,4 % entre 1995 et 2000), se concentrent sur trois régions : Île-de-France, Rhône-Alpes et Nord - Pas-de-Calais. Une analyse factorielle¹ régionale des liens entre spécialisations sectorielles et concentration des locaux de stockage montre que :

- les locaux de petite taille (moins de 1 000 m²) ne semblent pas liés à des secteurs d'activités particuliers ;
- les locaux de taille moyenne sont associés à la présence d'industries de biens intermédiaires ;
- les grands locaux correspondent aux industries agroalimentaires (IAA), aux industries de biens d'équipement et de biens de consommation, aux secteurs de l'énergie et du bâtiment.

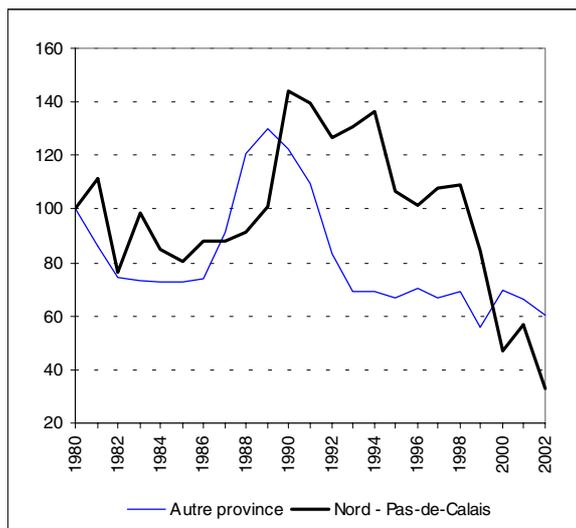
Après avoir connu un fort développement, les constructions à destination de stockage sont en net recul en région Nord - Pas-de-Calais, tout comme dans le reste de la province, notamment au cours de la période 1994-2002 (*figure 1*). En 2002, la part des constructions de stockage dans l'ensemble des constructions est passée sous la barre des 5 % en Nord - Pas-de-Calais. Entre 1995 et 2001, 1 641 locaux de stockage non agricole ont vu le jour dans cette région : jusqu'en 1998, 170 locaux étaient créés en moyenne par an, mais depuis 1999, le rythme de création annuelle enregistre une baisse de plus de la moitié du nombre de locaux (*figure 2*).

* Etude réalisée par l'Association ADRINORD et l'Université des Sciences et Technologies de Lille sous la direction du pôle études déconcentrées du département des études économiques du SES.

¹ Analyse en Composantes Principales (ACP) sur la période 1980-1996.

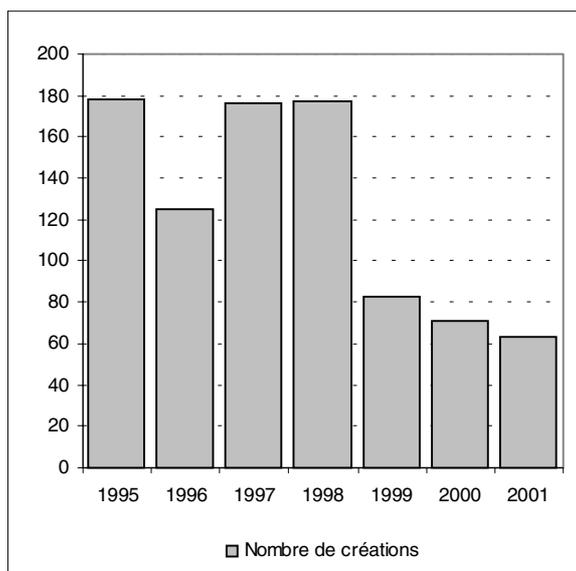
ENTREPÔTS

Figure 1 : Evolution comparée des surfaces de locaux de stockage commencés
Indice en base 100 (1980)



Source : Ministère de l'Équipement/DAEI/SES : SITADEL.

Figure 2 : Les créations d'entrepôts en région Nord - Pas-de-Calais entre 1995 et 2001



Champ : locaux de stockage non agricole construits terminés.
Source : SITADEL, SES.

On y construit le plus souvent de petits locaux : un peu plus de la moitié (55 %) a une superficie inférieure à 500 m² et ceux-ci n'occupent ensemble que 8 % de la superficie de locaux construits de 1995 à 2001. A l'opposé, 18 % des grands locaux (plus de 1 500 m²) représentent 73 % de la superficie totale (figure 3) ; la moitié d'entre eux relève d'activités de services.

Les locaux de stockage se concentrent dans les zones fortement urbanisées, en particulier les zones d'emploi. Près de 60 % des locaux, occupant un peu plus de 71 % de la surface totale, sont installés sur le territoire d'un pôle urbain. Si les pôles urbains dans leur ensemble atteignent à peine le quart de la superficie de la région, ils rassemblent un peu plus des trois quarts de sa population (figure 4). Trois zones d'emploi se distinguent : la zone de Lille avec plus de 16 % des locaux créés (près de 30 % de la surface totale), la zone de Roubaix-Tourcoing et celle de l'Artois-Ternois.

ENTREPÔTS

Figure 3 : Répartition des locaux par classe de superficie

En %

Superficie en m ²	Nombre de locaux	Superficie occupée
]0, 250]	37,4	3,4
]250, 500]	16,8	4,7
]500, 1000]	18,5	10,4
]1000, 1500]	9,1	8,7
]1500, 2000]	4,8	6,4
]2000, 5000]	8,3	20,7
]5000, 10000]	3,2	16,4
Sup. à 10000	2,0	29,2
Ensemble	100,0	100,0

Champ : 1641 locaux de stockage non agricole construits terminés entre 1995 et 2001.

Source : SITADEL, SES.

Figure 4 : Répartition des locaux par aire urbaine

En %

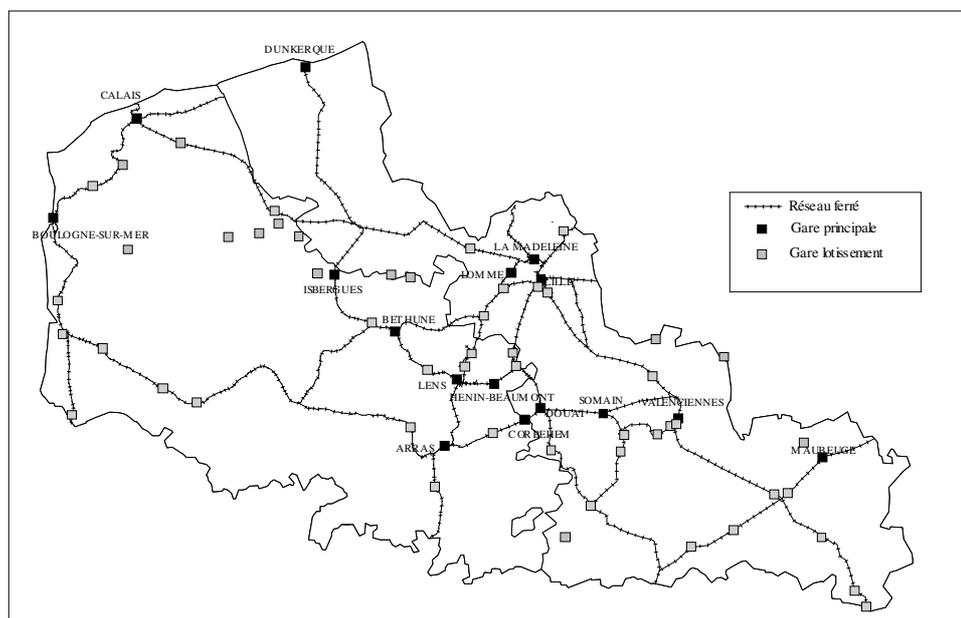
	Nombre de locaux	Surface occupée	Part de l'aire urbaine	
			Superficie	Population
Pôles urbains	59,9	71,3	24,5	75,8
Communes monopolarisées	15,5	14,6	23,5	10
Communes multipolarisées	9,0	7,5	16,2	5,8
Rural sous faible influence urbaine	7,2	1,8	17,6	4,3
Pôles ruraux	2,9	2,7	2,3	1,8
Communes sous influence pôle rural	1,6	0,6	4,6	0,7
Rural isolé	4,0	1,5	11,2	1,6
Ensemble	100,0	100,0	100,0	100,0

Champ : 1641 locaux de stockage non agricole construits terminés entre 1995 et 2001.

Source : SITADEL, SES.

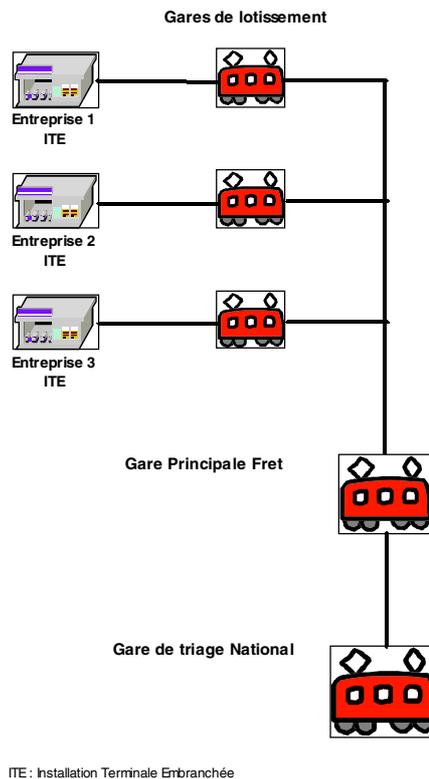
Les activités logistiques dans la région disposent d'infrastructures appuyées sur un réseau ferroviaire riche en gares et installations terminales (figure 5). Les gares principales (au nombre de 16) possèdent des espaces pour le stockage et l'équipement adéquat pour le chargement et le déchargement de marchandises et pondéreux. Elles sont situées à proximité des plates-formes multimodales puisqu'elles assurent la liaison avec la route. Plus petites, les gares de lotissement permettent d'effectuer la liaison entre les gares principales et les entrepôts. Elles ne possèdent pas d'équipement pour le chargement et le déchargement des marchandises et conteneurs. Elles font simplement du « porte à porte » avec les entreprises équipées d'une installation terminale embranchée (ITE). Une locomotive récupère les wagons chargés à chaque entrepôt équipé d'une ITE et les achemine à la gare principale la plus proche (figure 6). Le Nord - Pas-de-Calais compte 49 gares de lotissement et 320 ITE sur lesquelles se réalise 85 % du trafic de fret ferroviaire de la région.

Figure 5 : Réseau ferré et gares du Nord - Pas-de-Calais



ENTREPÔTS

Figure 6 : Schéma fonctionnel



Une répartition spatiale des locaux qui s'explique d'abord par celle de la population desservie

Les modèles de localisation-affectation permettent de déterminer le nombre optimal d'implantations des activités humaines, leur localisation, ainsi que leur aire de marché et leur taille. Cette étude n'utilise pas directement la théorie des localisations, ne serait-ce que parce que, au niveau individuel, les informations sur la structure spatiale du « parc » de locaux de stockage font défaut (*encadré 1*). Par contre, cette théorie liste des forces globales qui influencent ces localisations et, par ce biais, propose des variables explicatives. Elle conduit en particulier à élaborer des variables de distance ou d'accessibilité aux réseaux de transport et à un certain nombre de points remarquables de ces réseaux : les grands pôles d'habitat ou d'activité, les grands nœuds de communication et les points d'accès aux réseaux.

Encadré 1

Source, concepts et variables

A ce jour, aucune statistique n'est disponible sur les surfaces de stockage incluses dans les bâtiments industriels ou de vente ; on ne dispose d'informations chiffrées que sur les nouvelles constructions de locaux de stockage indépendants. La construction de bâtiments à vocation principale d'entrepôt fait l'objet d'autorisations de construire qui sont suivies statistiquement sous l'intitulé « lieux de stockage non agricole » dans la source statistique SITADEL du SES.

L'étude s'intéresse aux locaux de stockage non agricole construits et terminés, c'est-à-dire ayant fait l'objet d'une déclaration d'achèvement de travaux, sur la période 1995-2001 dans la région Nord - Pas-de-Calais. L'étude aurait pu également s'appuyer sur un autre indicateur issu de la même source, la statistique des mises en chantier. Les ouvertures de chantier, comme les achèvements de travaux font l'objet de déclarations volontaires de la part du maître d'ouvrage ; mais les premières sont mieux connues du statisticien² en raison de leur utilisation comme indicateur avancé de l'activité des entreprises de construction. Si on fait l'hypothèse que la sous-déclaration des achèvements de travaux n'est pas corrélée au phénomène ici étudié, l'utilisation des surfaces terminées ne doit pas être de nature à perturber l'analyse.

² Les maîtres d'ouvrage ayant obtenu une autorisation administrative de construire mais n'ayant pas déclaré de mise en chantier sont relancés.

A chaque local a été affecté les coordonnées du centroïde³ de sa commune d'appartenance, ainsi qu'un certain nombre de caractéristiques de la commune, à savoir :

- la superficie ;
- la population résidente et la population au lieu de travail au recensement de 1999 ;
- le type d'urbanisation, défini par le zonage en aires urbaines et rurales : pôle urbain, commune monopolarisée, multipolarisée, commune sous forte influence urbaine, sous faible influence urbaine, pôle rural, commune sous influence d'un pôle rural, commune rurale isolée ;
- des variables indicatrices précisant si l'on trouve sur la commune un accès à une infrastructure de transport : accès autoroutier, gare et autres accès ferroviaires, port maritime ou fluvial.

Enfin, pour chaque commune a été calculé, la distance à vol d'oiseau :

- à l'agglomération lilloise ;
- au pôle urbain le plus proche ;
- au pôle rural le plus proche ;
- à l'échangeur autoroutier le plus proche ;
- au nœud autoroutier le plus proche ;
- à la gare principale la plus proche ;
- à la gare de lotissement la plus proche ;
- au port le plus proche.

Les locaux ont été répartis selon le type d'activité de l'entreprise à laquelle ils étaient rattachés : industrie, commerce de gros, commerce de détail, services, non renseigné. Cette information a été obtenue par recherche manuelle à partir du nom et de l'adresse du maître d'ouvrage. Elle n'a été réalisée que pour les locaux dont la surface est au moins égale à 1 500 m².

Étudier des localisations, c'est les mettre en rapport avec d'autres composantes de l'espace : des pôles, des réseaux, des points d'accès. Pour ce faire, l'on a recours aux techniques statistiques de l'analyse spatiale des nuages de points « Spatial Point Pattern Analysis » (*encadré 2*).

Les locaux de stockage créés en région Nord – Pas-de-Calais depuis 1995 sont très fréquemment⁴ implantés à proximité des échangeurs, des nœuds autoroutiers et des gares. La distance à laquelle s'exerce cette attractivité n'est cependant pas la même selon les types d'infrastructure : jusqu'à 5 km pour les échangeurs et les gares de lotissement ferroviaire, 10 km pour les nœuds autoroutiers et les gares principales ferroviaires (*figure 7*). Les plus grands locaux apparaissent plus polarisés et plus sensibles à l'accessibilité aux infrastructures que les plus petits (*figure 8*). Lorsque l'on s'intéresse non pas au nombre de locaux, mais à la surface des locaux créés, on aboutit à des conclusions semblables : les surfaces de stockage sont sur-représentées à proximité des infrastructures.

Si l'on procède à des rapprochements avec la population des zones considérées (et non plus leur superficie), le nombre de locaux est a contrario légèrement sous-représenté à proximité des infrastructures. Quant à la répartition spatiale des superficies de locaux, elle n'apparaît plus comme marquée par une sur-représentation évidente. Une implantation un peu plus dense persiste, sauf à faible distance pour les nœuds autoroutiers. Elle disparaît pour les gares dans leur ensemble ; la contrainte de proximité reste néanmoins plus forte pour les gares de lotissement que pour les gares principales.

En résumé, il existe en général une forte concentration spatiale des locaux créés autour de lieux caractéristiques (pôles urbains, infrastructures de transport), mais cette concentration suit aussi d'assez près la répartition spatiale de la population, elle-même fortement polarisée. La méthode utilisée ne permet pas de dire si cette attraction est le fait des infrastructures elles-mêmes ou des agglomérations où elles sont localisées.

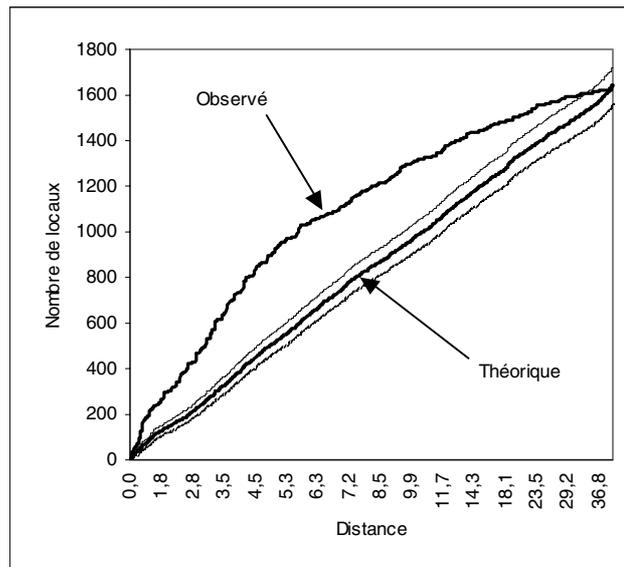
³ Centroïde : point situé au centre d'un polygone.

⁴ La distribution observée se situe significativement au-dessus de la distribution théorique (loi binomiale négative).

ENTREPÔTS

Figure 7 : Localisation de l'ensemble des locaux par rapport aux échangeurs autoroutiers

Référence : superficie



Légende :

- nombre observé de locaux de stockage à proximité du ou des échangeurs autoroutiers,
- nombre moyen théorique de locaux de stockage, proportionnel à la superficie (référence) des communes concernées.

Les deux traits entourant le nombre moyen théorique déterminent l'intervalle de confiance (+/- 2 écarts-types).

Lecture : le nombre de locaux créés à proximité des échangeurs autoroutiers est nettement supérieur au nombre théorique de locaux, quand celui-ci est proportionnel à la surface des communes concernées. L'écart entre ces deux nombres augmente fortement jusqu'à une distance de 5 km, pour diminuer ensuite lentement.

Source : SITADEL, SES. Traitements ADRINORD.

Figure 8 : Implantation des locaux et accessibilité

Distance à ...	Le nombre moyen de locaux est proportionnel à ...			La surface des locaux est proportionnelle à ...	
	superficie	population	emploi	superficie	population
Commune de Lille					
Ens. des locaux	(35 km)				
Locaux sup >= 1500 m ²		(30 km)	(30 km)		
Pôles urbains de la région					
Ens. des locaux		(légèr.)			(8 km)
Locaux >= 1500 m ²	(12 km)	(légèr.)			
Pôles urbains (hors Lille) de la région					
Ens. des locaux		(légèr.)			(10 km)
Locaux >= 1500 m ²		(légèr.)			
Echangeurs autoroutiers					
Ens. des locaux	(5 km)	(légèr.)		(5 km)	
Locaux >= 1500 m ²	(5 km)	(légèr.)			
Nœuds autoroutiers					
Ens. des locaux	(10 km)			(10 km)	
Locaux >= 1500 m ²	(10 km)	(légèr.)			
Gares					
Ens. des locaux	(5 km)	(légèr.)		(5 km)	ns
Locaux >= 1500 m ²	(5 km)	(légèr.)			
Gares principales					
Ens. des locaux	(10 km)			(10 km)	ns
Locaux >= 1500 m ²	(10 km)	ns			
Gares de lotissement					
Ens. des locaux	(5 km)	ns		(5 km)	(5 km)
Locaux >= 1500 m ²	(5 km)				

Légende :

Très forte sur-représentation des locaux de stockage
Sur-représentation des locaux de stockage
Sous-représentation des locaux de stockage

Entre parenthèses : portées moyennes en km.

On constate une sur-représentation du nombre moyen de locaux de stockage dans les communes à proximité des échangeurs autoroutiers.

La distance à laquelle s'exerce cette attractivité s'étend jusqu'à 5 km.

Source : SITADEL, SES. Traitements ADRINORD.

Le rôle marquant de l'accessibilité à la métropole lilloise

On peut également caractériser le processus de localisation par le nombre d'établissements d'un type donné créés dans chacune des 1 500 communes de la région. Cette statistique s'analyse avec des modèles économétriques de comptage, utilisant des variables représentatives de la structure de l'espace, en particulier les distances à d'autres communes (*encadré 2*).

Les facteurs d'accessibilité ont un impact significatif : accessibilité à la métropole et dans une moindre mesure aux autres pôles urbains, au réseau autoroutier via les échangeurs, au réseau ferroviaire via les gares de lotissement⁵. Les quatre effets estimés sont de même importance (*figure 9*). Un accroissement de la distance de 10 % conduit, toutes choses égales par ailleurs, à une réduction du nombre attendu de créations de locaux de stockage : de 2,0 % quand on s'éloigne de Lille et de 1,7% quand on s'éloigne d'une gare de lotissement. L'influence des pôles ruraux apparaît significative, quoique faible. L'effet estimé indique que, toutes choses égales par ailleurs, un pôle rural accueille un tiers de locaux de stockage de plus que les autres communes. Le coefficient de dispersion prend une valeur élevée, proche de 1,2, qui est très significativement différente de zéro : on est donc loin d'un processus spatial homogène.

Figure 9 : Modélisation de la création de locaux de stockage

Variables	Ensemble des locaux			Locaux d'au moins 1500 m ²		
	Coefficient estimé	Ecart-Type	Statistique de Wald	Coefficient estimé	Ecart-Type	Statistique de Wald
Constante	-8,142	0,469	//	-5,762	0,611	//
Années :			121,2	//	//	//
1995	0,965	0,123	//	//	//	//
1996	0,808	0,126	//	//	//	//
1997	0,848	0,124	//	//	//	//
1998	0,798	0,126	//	//	//	//
1999	0,259	0,137	//	//	//	//
2000	0,203	0,139	//	//	//	//
2001 (Référence)	0	0	//	//	//	//
Log(Superficie)	0,521	0,054	94,6	0,498	0,073	45,6
Log(Population)	0,496	0,032	226,0	0,513	0,045	126,6
Pôle rural	0,334	0,192	2,9	0,323	0,26	1,5
Log(1+distance à Lille)	-0,197	0,053	14,0	-0,16	0,074	4,6
Log(1+distance pôle urbain)	-0,113	0,049	5,2	-0,053	0,072	0,6
Log(1+distance échangeur)	-0,125	0,037	11,2	-0,152	0,051	8,8
Log(1+dist. gare lotissement)	-0,173	0,041	17,7	-0,181	0,058	9,8
Dispersion	1,184	0,126	//	0,856	0,091	//
Log-vraisemblance		-3 336			-182	

Champ : locaux de stockage non agricole construits terminés entre 1995 et 2001.

Source : SITADEL, SES. Traitements ADRINORD.

L'année de construction est une variable explicative très significative, avec des valeurs élevées des coefficients estimés en début de période (1995-1998), plus faibles ensuite (1999-2001). Ceci traduit la baisse brutale depuis 1999, de plus de la moitié du nombre des locaux créés.

La prise en compte simultanée de la superficie et de la population des communes améliore la qualité de l'ajustement du modèle. La somme des coefficients du logarithme de la superficie et de la population (1,017) n'est pas significativement différente de l'unité. Cela s'interprète comme l'existence de rendements d'échelle constants, au sens où une superficie deux fois plus importante, abritant une population deux fois plus nombreuse, conduit à un doublement du nombre moyen attendu de créations de locaux de stockage. Autrement dit, à densité inchangée, le nombre moyen attendu de créations est proportionnel à la superficie de la commune. Par contre, à superficie inchangée, une augmentation de la population et donc de la densité s'accompagne d'une augmentation moins que proportionnelle du nombre moyen de créations de locaux de stockage. De ce point de vue, les zones denses ne sont pas des localisations privilégiées.

⁵ L'accessibilité aux ports a été laissée de côté, du fait que ces derniers sont localisés dans de grands pôles urbains de la région et engendre donc une confusion dans l'évaluation des impacts réciproques.

ENTREPÔTS

Pour les locaux d'au moins 1 500 m², les effets de la superficie communale et de la population sont également du même ordre. La somme des coefficients de ces deux variables est là encore égale à l'unité, avec un partage proche de 0,5/0,5 : il y a invariance d'échelle et l'élasticité par rapport à la densité de population est de 0,5, pour les locaux de stockage d'au moins 1500 m², comme pour l'ensemble des locaux.

Globalement, les autres variables introduites dans le modèle sont moins explicatives de l'implantation des locaux dépassant 1 500 m² que de celle de l'ensemble des locaux. L'effet des pôles ruraux qui était positif et faiblement significatif pour l'ensemble des locaux ne l'est plus dès lors que l'on exclut de l'analyse les locaux de petite taille. L'élasticité à la distance à Lille est un peu plus faible que pour l'ensemble des locaux de stockage. L'effet de la distance aux autres pôles urbains n'est plus significatif. Les élasticités de la distance à l'échangeur et à la gare de lotissement la plus proche sont significativement négatives dans les deux cas, avec des ordres de grandeur similaires.

Les résultats des deux méthodes d'analyse utilisées dans cette étude sont convergents. Il ressort que la répartition spatiale de la population est, dans la région Nord - Pas-de-Calais, le déterminant fondamental de la géographie des créations de locaux de stockage et ce, quelle que soit leur taille. En particulier, la méthode économétrique indique que les créations de locaux de stockage sont, à densité égale, proportionnelles à la superficie de la zone. Néanmoins, pour une zone donnée, une croissance de la densité, entraîne une croissance moins que proportionnelle du nombre de locaux créés. Par exemple, un quadruplement de la densité conduit seulement à un doublement du nombre de locaux créés.

Encadré 2

Méthodes

L'analyse du nuage de points

L'analyse statistique d'un nuage de points, caractérisés par leurs coordonnées géographiques et diverses informations, permet de répondre à deux types de questions :

- Peut-on considérer que les points de ce nuage sont issus d'une répartition spatiale purement aléatoire ?
- Si ce n'est pas le cas, quels éléments de l'espace sont susceptibles d'attirer ou au contraire de repousser les points, contribuant ainsi à expliquer les localisations observées ? Quels rapports entretiennent entre eux des nuages de points différents ?

On dispose pour chaque local de stockage de ses coordonnées géographiques et d'informations telles que le type d'ouvrage, le mode d'utilisation principale, la SHON, etc. L'objectif est de juger de la répartition spatiale de ces locaux par rapport à un point ou un ensemble de points de référence. Ces points de référence correspondent à des déterminants du processus de localisation et peuvent être considérés comme indépendants du processus de création des locaux de stockage : métropole lilloise, ensemble des pôles urbains de la région, ensemble des échangeurs et nœuds autoroutiers, ensemble des gares de marchandises (gares principales / gares de lotissement), ensemble des ports, maritimes et fluviaux.

La distance entre un local et un point de référence est la distance géométrique entre le centroïde de la commune d'implantation du local et le centroïde de la commune du point de référence. On considère ensuite l'ensemble des communes dont la distance à la commune du point de référence le plus proche est inférieure ou égale à un niveau fixé. On compare alors le nombre de locaux de stockage observés dans cet ensemble, au nombre moyen obtenu avec un processus purement aléatoire. Ce dernier dépend de la fonction de densité du processus :

- processus homogène dans l'espace géographique : le nombre moyen de locaux est alors proportionnel à la superficie de l'ensemble des communes considérées,

- processus ayant une densité proportionnelle à la population locale : le nombre moyen de locaux est alors proportionnel à la population cumulée de l'ensemble des communes considérées,
- processus ayant une densité proportionnelle à l'emploi local : le nombre moyen de locaux est alors proportionnel à l'emploi cumulé de l'ensemble des communes considérées.

L'économétrie des données de comptage

La variable à expliquer est le nombre de locaux de stockage Y_i construits entre 1995 et 2001 dans chacune des 1 547 communes i de la région Nord - Pas-de-Calais, soit 10 829 observations au total. D'un point de vue économétrique, la famille de modèles adaptée à l'étude de ce type de variables est celle des modèles économétriques sur données de comptage⁶.

On fait l'hypothèse que le nombre de locaux Y_i suit une loi de probabilité de Poisson de paramètre λ_i , dépendant de variables explicatives X_i , telle que :

$$\lambda_i = \exp(X_i \beta) \quad i = 1, \dots, n$$

où :

- X_i est un vecteur (1 x k) de variables explicatives : existence et importance des marchés locaux, accessibilité à des pôles ou à des points remarquables des réseaux d'infrastructures, etc.,
- β est un vecteur (k x 1) de paramètres à estimer.

La distribution de Poisson a un seul paramètre λ_i positif qui détermine à la fois la moyenne et la variance conditionnelle de la loi :

$$E(Y_i | X_i) = \text{Var}(Y_i | X_i) = \lambda_i = \exp(X_i \beta)$$

L'expression analytique de la loi de Poisson est donnée par :

$$\Pr(Y_i = y_i | X_i) = \frac{\exp[-\exp(X_i \beta)] * \exp[y_i X_i \beta]}{\text{Fact}(y_i)}$$

L'estimation du modèle se fait par maximum de vraisemblance.

Utiliser le modèle de Poisson, c'est faire notamment l'hypothèse qu'il n'y a pas d'hétérogénéité dans l'échantillon. Cette hypothèse « d'équi-dispersion » est très restrictive. En pratique, la variance est supérieure à la moyenne, on parle alors de « sur-dispersion ». La présence de « sur-dispersion » fait que les écarts-types des paramètres sont sous-estimés, conduisant à une surévaluation des seuils de signification statistique du modèle. Il est possible de tester l'hypothèse d'égalité entre moyenne et variance par l'intermédiaire d'un test de sur-dispersion.

Une solution à ce problème consiste à introduire [Hausman et al., 1984] un aléa supplémentaire dans le paramètre de la loi de Poisson. Ce nouveau modèle est une extension du modèle de Poisson, tel que :

$$\lambda_i = \exp(X_i \beta + \mu_i) \quad i = 1, \dots, n$$

avec $\exp(\mu_i) = \varepsilon_i$

On fait ensuite l'hypothèse que ε_i suit une loi Gamma (Γ) de moyenne 1 et de variance ∂^2 . Le choix de la loi Gamma est justifié par des facilités de calcul. En intégrant la loi conditionnelle de Y_i aux valeurs de X_i et ε_i , on obtient une loi binomiale négative :

⁶ Pour en savoir plus : Rainer Winkelmann (1994, 2000), *Econometric Analysis of Count Data*, Third edition, Berlin: Springer-Verlag et Colin Cameron and Pravin K. Trivedi, *Regression Analysis of Count Data* Econometric Society Monograph N° 30, Cambridge University Press, 1998.

$$\Pr(Y_i = y_i | X_i) = \frac{\Gamma(\partial^{-2} + y_i)}{\Gamma(\partial^{-2})\Gamma(y_i + 1)} \left[\partial^2 \exp(X_i \beta) + 1 \right]^{\partial^{-2}} \left(\frac{\partial^2}{\partial^2 \exp(X_i \beta) + 1} \right)^{y_i}$$

lorsque $\partial \rightarrow 0$, le modèle converge vers un modèle de régression poissonienne.

Cette loi binomiale négative a pour paramètres ∂^2 et β , les moyennes et variance de Y_i sont alors :

$$E(Y_i | X_i) = \exp(X_i \beta)$$

$$\text{Var}(Y_i | X_i) = \exp(X_i \beta)(1 + \exp(X_i \beta) * \partial^2)$$

La variance est une transformation croissante convexe de la moyenne.

L'estimation du modèle revient à calculer les dérivées analytiques du logarithme népérien d'une fonction Gamma. Plusieurs procédures d'estimations sont disponibles : moindres carrés non linéaires (MCNL), moindres carrés non linéaires quasi généralisés (MCNLQG), méthode du pseudo maximum de vraisemblance quasi généralisé (PMVQG), etc.

Pour aller plus loin

- Becker D., « Le développement des implantations logistiques en France et ses enjeux pour les politiques d'aménagement », Rapport n° 2001-0104-01, Conseil Général de Ponts et Chaussées.
- Beckman M.J, Thisse J.F, 1986, « The location of production activities », Chapitre 2 du *Handbook of Regional Economics*, édité par P Nijkamp, Amsterdam : North Holland.
- Hakimi, S.L., 1964, « Optimal location of switching centers and the absolute centers and medians of a graph », *Operations Research*, 12, 450-459.
- Upton, G.J.G. and Fingleton, B., 1985, « Spatial Data Analysis by Example », Volume 1 & 2, *Wiley*, Chichester.

■