

Evaluation du profil de développement économique territorial de la Roumanie

JABA Elisabeta
Département de Statistiques
Université "A.I.Cuza" de IASI
ROUMANIE
ejaba@uaic.ro

IATU Corneliu
Département de Géographie
Université "A.I.Cuza" de IASI
ROUMANIE
corneliu_iatu@yahoo.fr

IONESCU Alina Mariuca
Département de Statistiques
Université "A.I.Cuza" de IASI
ROUMANIE
alina.ionescu@yahoo.com

Résumé

L'article veut offrir des informations fiables pour le fondement des politiques de développement régional par la prise en considération des ressources locales. L'identification du profil régional en fonction des ressources existantes et le niveau de développement constitue le but essentiel de l'analyse. La recherche s'est réalisée à l'aide de l'analyse statistique multivariée sur les données statistiques macroéconomiques disponibles après 1990. Les résultats de la recherche ont mis en évidence les conclusions suivantes :

- *en profil régional ne sont pas valorisées les ressources locales que dans une moindre mesure ;*
- *la corrélation entre ressources et le niveau de développement économique a montré la nécessité de l'adoption d'une politique de développement par laquelle se valorise mieux les ressources existantes.*

Mots clés : développement économique, régional, profil, Roumanie, analyse multivariée

Abstract

The present article is meant to provide reliable information on regional development policies by taking into consideration local resources.

The main purpose of the paper is to identify the regional profile induced both by the existing resources and the level of development. The research was made with the help of the multivariate statistical analysis of the macroeconomic statistical data after 1990. The results of the research pointed out the following conclusions:

- *local resources are used to a smaller extent at the regional level;*
- *the correlation between resources and the level of economic development underlines the necessity of adopting a development policy that would better use the present resources.*

Key words: economic development, regional, profile, Romania, multivariate analysis

1. Introduction

Le développement économique territorial constitue un enjeu essentiel à l'heure actuelle au niveau européen et mondial. Les problèmes majeurs qui se posent sont d'attractivité pour des activités économiques et dans d'autres cas de consolidation économique (il y a des régions qui se confrontent avec les problèmes de reconversion, transition etc.). Construire un profil territorial économique implique la prise en compte d'une série d'indicateurs pas seulement de nature économique mais aussi de nature socio-économique (Chih-Kai, 2008). Un rôle important revient à la demande des services, le point clé étant l'accessibilité. L'importance de la construction du profil territorial est vitale pour la diagnose d'intervention des autorités locales.

La caractérisation du développement économique territoriale a comme but la mise en évidence de la spécificité des départements et leurs perspectives de développement, la poursuite des décalages entre les territoires concernés et l'assurance des éléments de référence pour allouer les ressources et prendre les bonnes décisions pour atténuer ces décalages. La fiabilité statistique et la pertinence des indicateurs sont les éléments clés pour arriver à une caractérisation en concordance avec les réalités économiques d'un territoire.

Cette étude veut aboutir l'identification d'un profil territorial du développement socio-économique de la Roumanie par l'intermédiaire des méthodes statistiques multivariées, pour pouvoir synthétiser les multiples facettes de ce concept.

2. Méthodes

2.1. Les variables utilisées

Initialement, on a utilisé 25 variables de la base de données de l'INS Bucarest. Ces variables sont présentées dans le tableau 1.

Tableau 1. Variable utilisées

	Variabile
Démographie	Poids de la population rurale (p_rurale)
Résultats	PIB par habitant (PIB_hab); Dépenses avec l'activité de recherche-développement par habitant (dép_rech_hab)
Marché du travail	Poids de la population occupée dans les principales activités économiques : agriculture, chasse et sylviculture (pop_agr), industrie (pop_ind), constructions (pop_constr), commerce (pop_com), enseignement (pop_ensg), santé et assistance sociale (pop_san), / transactions immobilières et autres services (pop_imm), intermédiations financières (pop_fin), hôtels et restaurants (pop_hotr) du total de la population occupée; Taux de chômage (r_schom)
Education	Taux d'abandon dans l'enseignement primaire et secondaire (r_aband_éc); Poids de la population écolière de la population totale (r_pop_éc); Nombre de facultés à 100000 habitants (nr_fac)
Santé	Taux de mortalité infantile (r_minf); Durée de vie moyenne (dur_moy_vie); Lits dans les hôpitaux par 100000 habitants (lits_hop); Médecins à 100000 habitants (médecins)
Infrastructure	Poids des localités du département où se distribue l'énergie thermique (en_therm); Poids des localités du département où se distribue gaz naturel (gaz_nat); Poids des localités du département avec installations de canalisation publique (canal); Poids des localités du département avec adduction d'eau potable (eau_pot)
Tourisme	Indice d'utilisation nette de la capacité touristique en fonction (tourisme)

2.2. Méthodes utilisées

Les méthodes statistiques utilisées dans l'étude sont: l'analyse en composantes principales (ACP) – pour l'analyse préliminaire des données; l'analyse cluster - pour l'identification des groupes homogènes des départements du point de vue du développement économique; l'analyse discriminant – pour valider la solution obtenue par l'analyse cluster.

Les données sont enregistrées au niveau du département, l'année de référence étant 2005. La source des données est l'Annuaire Statistique de la Roumanie (2006) et ces données ont été analysées avec le logiciel SPSS.

3. Résultats

3.1. Analyse préliminaire des données par ACP

L'analyse en composantes principales se justifie par la dimension des données (25 caractéristiques pour les 42 départements), toutes les 25 variables étant quantitatives continues. Par l'ACP, la dimensionnalité des données est réduite (Schott, 2006).

L'ACP est utilisé pour explorer les données initiales et sélectionner les variables utilisées pour la détermination d'un profil régional du développement économique en Roumanie.

Pour vérifier l'adéquation des données pour une analyse factorielle, on a utilisé le test de *sphéricité* Barlett (pour tester l'hypothèse nulle que les variables dans la matrice de corrélation de la population ne sont pas corrélées) et l'indicateur MSA (*Mesure of Sampling Adequacy*) de Kaiser-Meyer-Olkin (pour évaluer le degré où chaque variable peut être prédite de toutes les autres variables).

Les résultats obtenus par le traitement des données avec le logiciel SPSS sont présentés dans le tableau 2. Le niveau de signification associé au test Barlett, Sig=0,000, est sous 0,05 (seuil conventionnel), d'où on peut conclure que l'hypothèse nulle de non corrélation entre les variables est rejetée. Par conséquent, les données sont adéquates pour une analyse en composantes principales. Le même résultat est obtenu par la valeur de l'indicateur MSA de KMO (0,767), qui est supérieur à 0,5.

Tableau 2. *Mesure de l'Adéquation Statistique de Kaiser-Meyer-Olkin et le test de sphéricité Barlett, pour les variables initiales*

Mesure Kaiser-Meyer-Olkin d'adéquation de l'échantillonnage		.767
Test de sphéricité de Bartlett	aprox. Hi-carré	1235.58
	degré de liberté	300
	sig.	.000

Source: Résultats obtenus en SPSS par ACP

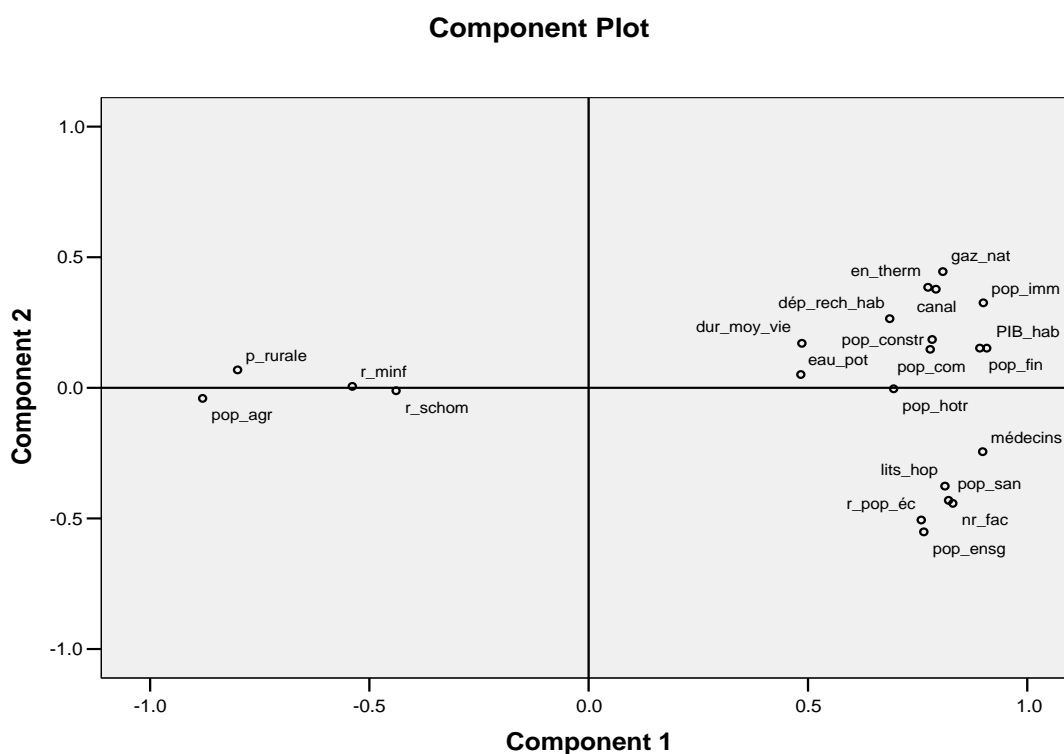
Un autre indicateur de l'adéquation des variables pour l'analyse est constitué par la matrice de corrélation anti-image. Chaque valeur de la diagonale principale de cette matrice montre la mesure de l'adéquation de l'échantillonnage (MSA) pour l'item respectif. Dans notre exemple, les variables suivantes: le poids de la population occupée en industrie, le taux d'abandon dans l'enseignement primaire et secondaire et l'indice de l'utilisation nette de la capacité touristique en fonction ont eu les valeurs de MSA sous 0,5. Dans ces conditions, ces variables seront exclues des analyses

ultérieures parce que les résultats indiquent des variables qui paraissent de ne pas se corréler avec la structure des autres variables.

Les estimations de la variance dans chaque variable, expliquées par les composantes de la solution factorielle peuvent constituer, aussi, un signe concernant les variables qui ne sont pas corrélées. Dans ce contexte, les variables taux d'abandon dans l'enseignement primaire et secondaire (*r_abandon_éc*) et l'indice d'utilisation nette de la capacité touristique en fonction (*tourisme*) ont des estimations sous 0,5 et elles ne seront pas gardées dans les analyses ultérieures parce qu'elles ne concordent pas avec la solution factorielle.

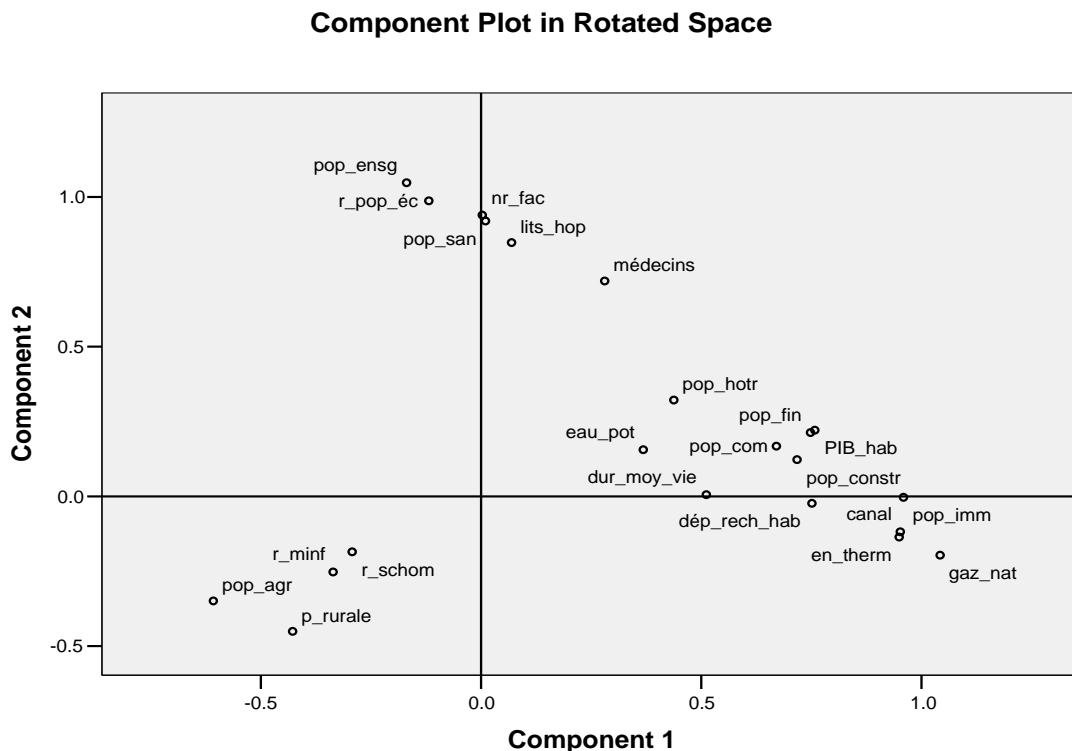
Après l'élimination de 3 variables de l'analyse, qui ne se corréllent pas avec la structure des autres, a résulté une augmentation de la mesure d'adéquation de l'échantillonnage Kaiser-Meyer-Olkin de 0,767 à 0,846, les variables retenues étant plus appropriées pour une analyse factorielle. Aussi, le pouvoir explicatif des composantes principales a augmenté à 84,139 %, la variance expliquée par les premiers deux axes factoriels sautant de 84,139 % à 67,208 % de la variance totale.

Figure 1. Position des variables étudiées sur les premiers deux axes factoriels



Source: Output obtenu en SPSS par ACP

Figure 2. Position des variables étudiées sur les premiers deux axes factoriels après la rotation des axes



Source: Output obtenu en SPSS par ACP

Dans la représentation graphique des points des variables sur les premiers deux axes factoriels (figure 1), s'observe que le premier axe oppose, d'une part, les variables qui décrivent les poids de la population rurale et de la population occupée en agriculture et, d'autre part, les variables qui décrivent le développement de l'infrastructure et les résultats économiques.

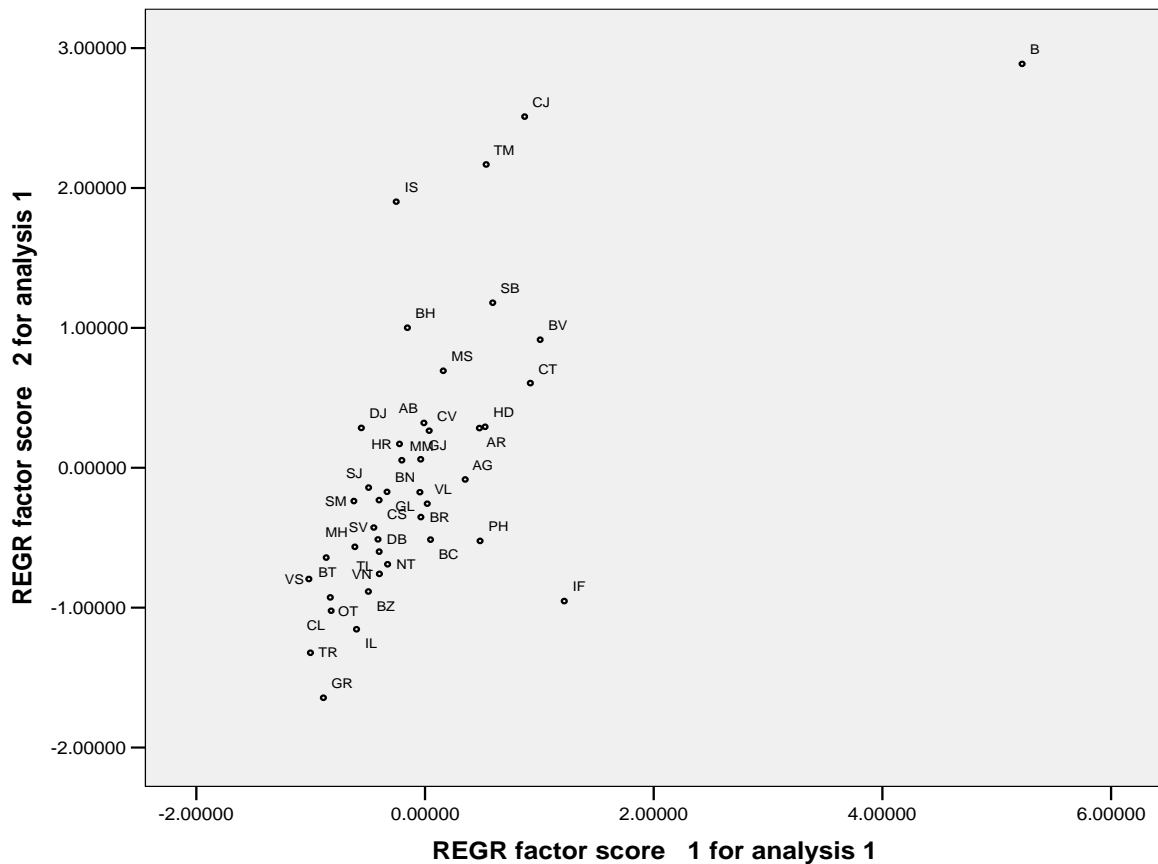
Parce que certaines variables ont des coefficients de corrélation proches comme valeur des deux axes, pour une meilleure interprétation des résultats de l'analyse en composantes principales, on a fait une rotation des axes par la méthode de rotation oblique *Direct Oblim*, disponible dans le programme SPSS.

Après la rotation des axes, la corrélation des variables avec les deux axes est meilleure et s'observe plus clairement les composantes principales: le potentiel de développement (le premier axe) et la qualité des ressources (le deuxième axe).

L'analyse des cartes factorielles obtenues avant et après la rotation des axes (figure 2) montrent qu'il y a des différences en profil territorial du point de vue du développement économique et des ressources disponibles des départements.

La représentation des départements sur les premiers deux axes factoriels (figure 3) met en évidence l'existence d'un *outlier* (Bucarest). La capitale du pays a des caractéristiques de développement économique très différentes par rapport aux autres unités administratives – territoriales, elle nécessitant une analyse individuelle des caractéristiques, sans l'inclure dans les analyses ultérieures.

Figure 3. Représentation des départements sur les premiers deux axes factoriels après la rotation des axes



3.2. Résultats de l'analyse des clusters

L'analyse des clusters a été utilisée pour l'identification des groupes homogènes des départements du point de vue du développement économique.

A l'aide de cette analyse, on a fait la représentation visuelle du développement économique en profil territorial (Del Campo, Monteiro, Soares, 2008), par l'identification des clusters homogènes par départements du point de vue des ressources existantes et du niveau du développement, le but étant celui de contribuer à l'optimisation des décisions de politique économique.

Parce que le volume de la population étudiée est relativement petit ($n=41$ départements après l'enlèvement des *outliers*), les méthodes appliquées ont été celles de classification hiérarchique et comme type de mesure des distances celle de la distance euclidienne carrée, fréquemment utilisée pour des données de type intervalle.

Après l'application des méthodes de classification hiérarchique disponibles en SPSS, on a observé que la méthode de la liaison moyenne intergroupes, la méthode de la liaison complète (le voisin le plus éloigné) et la méthode de Ward ont été celles qui ont groupé le plus clairement les départements en fonction des variables considérées et ont trouvé des clusters plus compacts et plus équilibrés (Jaba & Grama, 2004).

Pour déterminer le nombre optimum des clusters il n'y a pas des critères préétablis mais les informations utiles dans ce sens peuvent être offertes par le dendrogramme et la

série des coefficients du schéma d'agglomération, qui montrent le mode dans lequel les départements sont groupés en clusters dans chaque étape de l'analyse.

Par l'analyse de le dendrogramme et de la série des coefficients du schéma d'agglomération pour les trois méthodes (séparées), ont été identifiées 3 solutions possibles, chaque solution groupant les départements en 5 clusters (la solution optimale sera présentée plus détaillée dans la section 6).

3.3. Résultats de l'analyse discriminant

L'analyse discriminant aide à l'identification et la description des différences significatives entre les groupes des départements.

L'utilisation de l'analyse discriminant (Vaughn, Wang, 2008) se fait pour déterminer la solution pour laquelle existe une combinaison des variables qui sépare le plus objectivement possible les clusters par départements obtenus.

Dans cette étude, les variables discriminateurs (variables prédicteurs) sont représentées par les 22 variables indépendantes sélectionnées par ACP, et la variable de groupement, respectivement la variable qui fait l'objet de la classification, est le cluster d'appartenance, obtenue par l'analyse cluster.

Les différences significatives entre les groupes sont mises en évidence à l'aide des fonctions discriminant, représentant des combinaisons linéaires non corrélées des variables explicatives initiales et ayant la forme suivante : $D = b_1X_1 + b_2X_2 + \dots + b_pX_p + c$, où D = la fonction discriminant; X_j = le vecteur des variables discriminateurs; $j = 1, p$; b_j = les coefficients discriminants; c = une constante (Jaba s.a., 2006, pp.7-15).

L'application de l'analyse discriminant suppose la vérification des hypothèses concernant : la normalité des distributions multivariées (les variables prédicteurs doivent avoir des distributions multivariées normales), l'homogénéité des variances (homoscedasticité) et l'absence de la multicollinéarité parfaite entre les variables.

Le test de la normalité des variables a été réalisé en SPSS à l'aide du test Kolmogorov – Smirnov, les exemples d'application étant dans la littérature de spécialité assez présents (D'Alimonte, Cornford, 2008; Solomonoff, 2008) et pour le test de l'homogénéité des variances a été utilisé le test Levene (Jaba & Grama, 2004).

Les résultats des tests ont montré, généralement, la confirmation des hypothèses initiales, avec de très petites déviations en ce qui concerne les hypothèses de normalité et homogénéité. Lachenbruch¹ (1975) affirme que l'analyse discriminant est relativement robuste même s'il y a des petits changements dans les hypothèses de normalité et d'homogénéité des variances. Suite à ces constatations, l'analyse discriminant peut être appliquée sans influencer les conclusions qui se basent sur celle-ci.

Le tableau 3 présent le poids des départements, classifiés correctement par l'analyse discriminant pour chaque de 3 solutions cluster trouvées. Ainsi, dans l'étude réalisée, pour toutes les 3 solutions, la fonction discriminant classifie correctement 100 % du total des cas, c'est-à-dire tous les 41 départements. Un cas est considéré à être classifié correctement s'il est encadré, par le score de la fonction discriminant, dans le groupe dont appartient en réalité.

¹ Lachenbruch, P. A. (1975), *Discriminant analysis*. NY: Hafner, in <http://www2.chass.ncsu.edu/garson/pa765/discrim.htm>

Les résultats de la classification initiale peuvent offrir des estimations excessivement optimistes. La validation croisée essaye de résoudre ce problème par le fait que chaque cas de l'analyse est classifié sur la base des fonctions dérivées de tous les cas, autres que le cas considéré.

Tableau 3. *Les cas classifiés correctement, issues de l'analyse discriminant*

Nombre des clusters de la solution	Cas groupés initialement classifiés correctement	Cas groupés par la validation croisée, classifiés correctement
5 (Liaison moyenne)	100.0 %	73.2 %
5 (Liaison complète)	100.0 %	78.0 %
5 (Ward)	100.0 %	70.7 %

Source: Résultats obtenus en SPSS par l'analyse discriminant

La validation croisée est une méthode utilisée dans l'évaluation des règles de classification par l'estimation des taux d'erreur (Lachenbruch & Mickey, 1968).

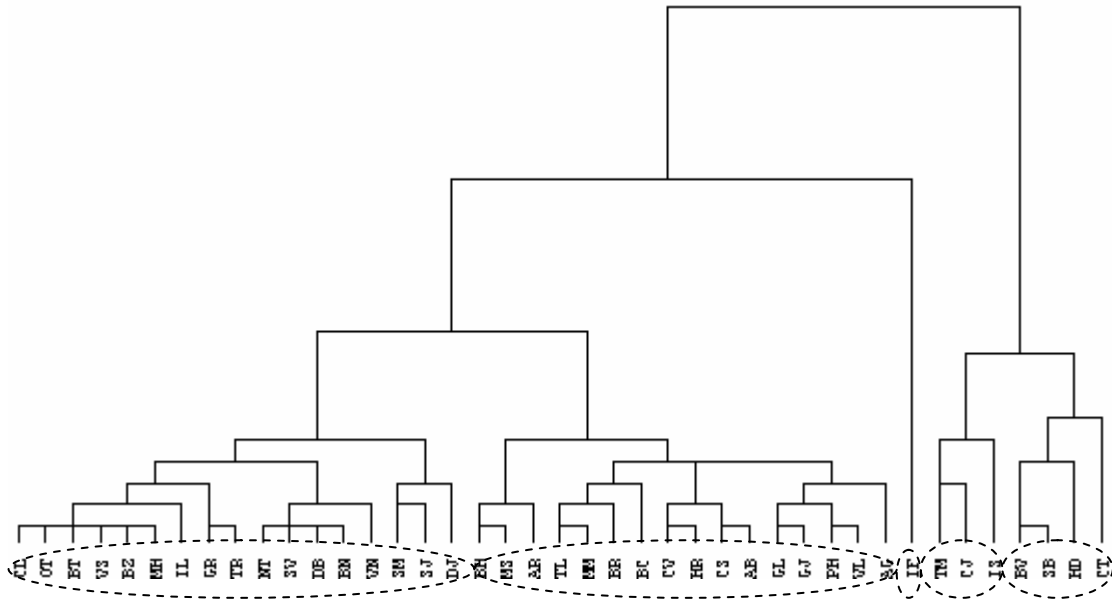
Les résultats de la validation croisée montrent que la méthode de la liaison complète a classifié correctement les plus nombreux des cas (78 %, c'est à dire 32 d'un total de 41 départements), générant la plus petite erreur (22 %). Par suite, cette méthode est la solution optimale pour le groupement des départements en fonction des variables analysées.

4. Clusters des départements

4.1. Classifications des départements de Roumanie par groupes

Le dendrogramme présenté dans la figure 4 montre clairement le groupement des départements en 5 clusters principaux. Ceux-ci peuvent être interprétés comme suit: l'un des clusters (3) a les départements les plus développés et avec une activité de recherche-développement très importante (Timis - TM, Cluj - CJ, Iasi - IS); le cluster 4 est celui des départements développés en matière d'infrastructure et utilités publiques, reconnus par leurs poids industriel pendant le régime communiste (Brasov - BV, Sibiu - SB, Hunedoara HD, Constanta - CT); le cluster 5 est formé d'un seul département Ilfov – IF, qui est profondément influencé par sa position autour de la capitale Bucarest; le cluster 1 est formé par les départements avec un développement économique modéré et assez contrasté (Bihor - BH, Mures - MS, Arad - AR, Tulcea - TL, Maramures - MM, Braila - BR, Bacau - BC, Covasna - CV, Harghita - HR, Caras Severin - CS, Alba - AB, Galati - GL, Gorj - GJ, Prahova - PH, Valcea - VL, Arges - AG); le cluster 2 groupe les départements les moins développés économiquement, avec un poids important de l'agriculture (Calarasi - CL, Olt - OT, Botosani - BT, Vaslui - VS, Buzau - BZ, Mehedinti - MH, Ialomita - IL, Giurgiu - GR, Teleorman - TR, Neamt - NT, Suceava - SV, Dambovita - DB, Bistrita Nasaud - BN, Vrancea - VN, Satu Mare - SM, Salaj - SJ, Dolj - DJ).

Figura 4. Groupement des départements dans les 5 clusters



Les clusters des départements obtenus sont mis en évidence dans la figure 5. La superposition de la représentation graphique des clusters des départements (figure 5) sur la carte des variables obtenue par ACP (figure 3) donne quelques caractéristiques des clusters obtenus.

Les départements du cluster 3 se différencient nettement des autres par la forte association positive avec les deux axes factoriels, ayant le plus grand niveau de développement. Le fait que ce groupe de départements a les plus grands coordonnées positives sur l'axe de l'accès aux services de santé et enseignement peut être expliqué par la composante du cluster, les trois départements ayant une tradition forte dans l'enseignement (universitaire surtout), culture, santé (services sanitaires spécialisées) et, en même temps, un poids démographique important.

Un autre cluster où les départements se remarquent par un niveau élevé de développement est le cluster 4, les départements de ce cluster ont les plus grands poids positifs sur le premier axe factoriel, ce poids étant déterminé par une infrastructure de bonne qualité et des bons indicateurs économiques.

Les départements du cluster 1 sont disposés autour de l'origine du système des axes factoriels, ce qui signifie un niveau modéré de développement. Ces départements se remarquent par des disparités intra-départementales importantes, ayant des noyaux d'industrialisation en forte opposition avec d'autres zones, qui sont moins développées.

Dans le deuxième cluster, les départements sont situés dans le troisième cadran du plan factoriel, ce qui imprime le plus bas niveau de développement avec les plus grands coordonnées négatifs sur les deux axes factoriels. Les activités du secteur primaire sont dominantes.

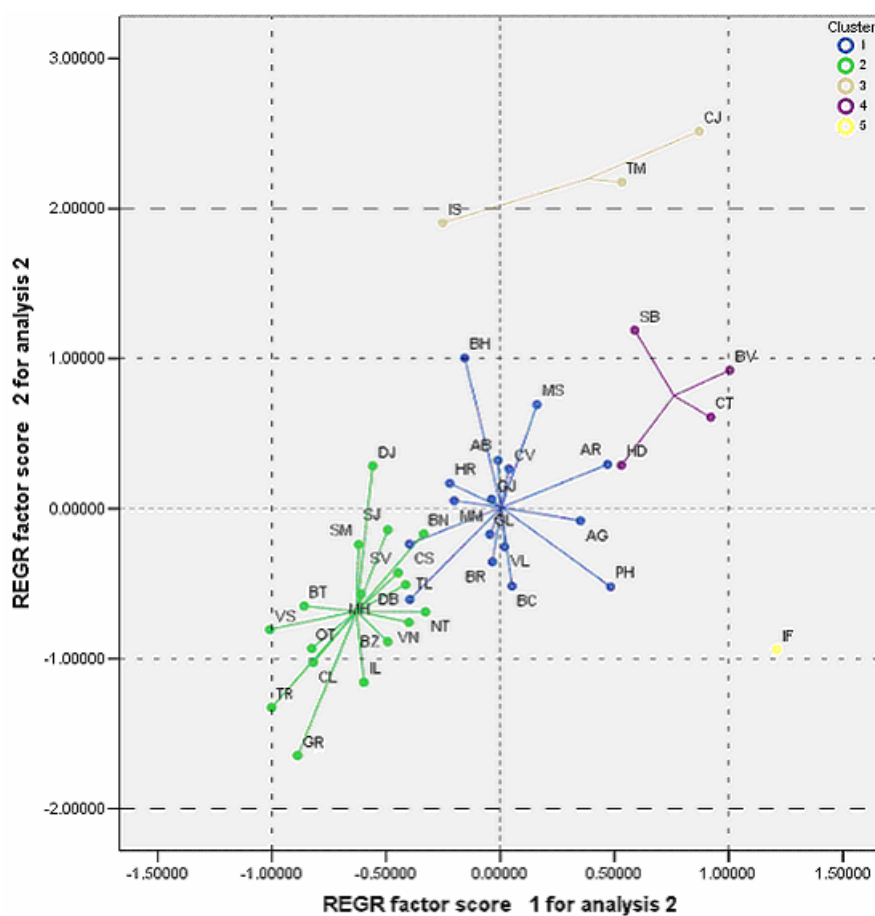
La représentation graphique de la figure 5 sort en évidence un seul département, celui de Ilfov comme atypique: la plus grande coordonnée positive sur le premier axe factoriel, comparable avec ce des plus développés départements et sur l'autre axe une coordonnée négative, similaire aux moins développés départements. Un rôle important revient au voisinage, le département de Ilfov étant proche de la capitale du pays,

Bucarest. L'influence de Bucarest se ressent fortement et explique ce paradoxe de positionnement sur les axes factoriels.

4.2. L'analyse de la position des départements et des clusters sur la carte factorielle

Cette analyse permet l'identification par la règle de 3σ , des plus développés et des moins développés départements par rapport à chaque axe factoriel. Pour les distinguer, il faut chercher les départements qui se trouvent au dehors des intervalles $\bar{x} \pm \sigma$, $\bar{x} \pm 2\sigma$ et, respectivement, $\bar{x} \pm 2\sigma$, correspondants aux deux axes et marqués sur le graphique à l'aide des lignes pointillées ($\bar{x} = 0, \sigma = 1$) (Jaba, 2007).

Figure 5. Représentation des clusters des départements sur la carte factorielle



Les axes des résultats économiques et de l'infrastructure d'utilités publiques identifient les départements de Brasov (BV) et Ilfov (IF) comme les plus développés. Ceux-ci sont suivis, proche de la limite de l'intervalle $\bar{x} \pm \sigma$, par Constanta (CT) et Cluj (CJ). Les plus faibles résultats économiques et accès aux réseaux des utilités publiques sont dans les départements de Vaslui (VS) et Teleorman (TR), départements avec une composante agricole majoritaire.

Les départements les plus développés en matière d'infrastructure sanitaire et d'enseignement sont Bihor (BH), Sibiu (SB), Iasi (IS), Timis (TM) et Cluj (CJ), qui se trouvent à droite de l'intervalle $\bar{x} \pm \sigma$. Les moins développés, de ce point de vue, sont Calarasi (CL), Ialomita (IL), Tulcea (TL) et Giurgiu (GR), positionnés à gauche de l'intervalle $\bar{x} \pm \sigma$, sur la deuxième axe factoriel. Le retard des derniers s'explique soit

par le manque de diversité des activités économiques (CL, IL, GR) soit par l'isolement et l'accès difficile (TL).

La carte de la figure 5 met en évidence aussi les départements situés à la limite des clusters, n'importe quelle modification des caractéristiques de développement pouvant déterminer le glissement vers le cluster voisin. C'est le cas pour des départements de Caras Severin (CS) et Tulcea (TL), qui se trouvent dans le cluster 1 mais ayant un grand degré de ressemblance avec les départements les moins développés du cluster 2. Par contraste, le département de Bistrita-Nasaud (BN), du deuxième cluster, présente des traits similaires avec ceux du premier cluster.

Le département de Hunedoara, même s'il est assimilé dans le groupe des départements plus développés (activités importantes dans l'industrie sidérurgique et minière en recul après 1990), est positionné très proche du groupe des départements avec un développement modéré. Sibiu a les plus grandes chances d'accéder dans le cluster des plus développés départements, étant le plus proche de ceux-ci.

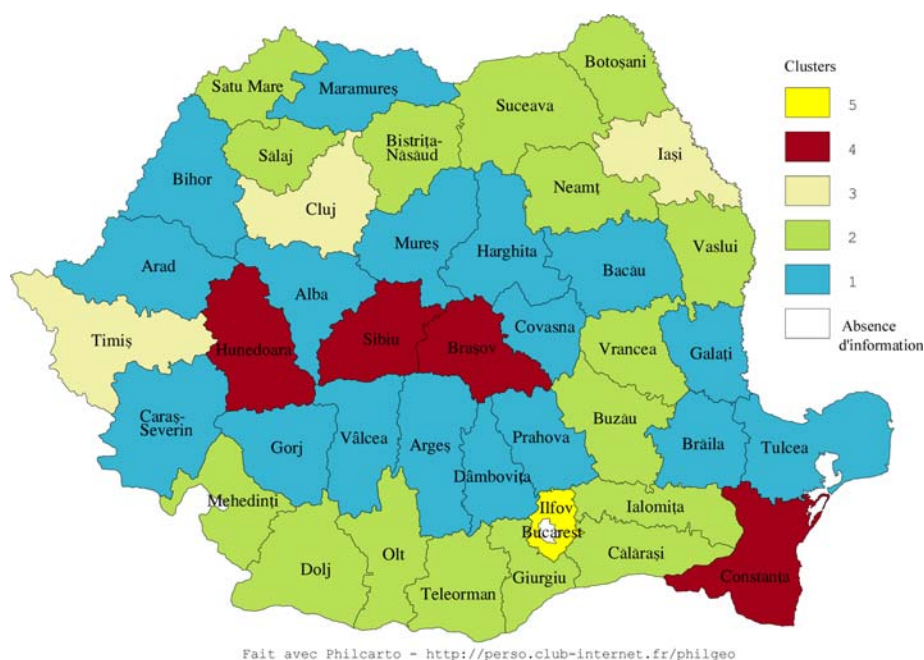
En analysant la distribution territoriale des clusters par départements pour la solution considérée optimale, on observe le fait qu'elle reproduit dans une grande mesure la répartition géographique, groupant des départements voisins.

4.3. Distribution géographique des clusters des départements roumains

Dans une certaine mesure, la solution reproduit la carte géographique des départements (figure 6), pouvant être efficace aussi pour différents cas comme: l'application des programmes sur des territoires zonaux réduits ou pour mettre les fondements des programmes pour des territoires plus étendus comme les régions, qui peut prendre en compte la distribution territoriale des départements visés.

Prendre en compte seulement les départements dans le cas des décisions de développement sera une faute. Il faut tenir compte aussi des contrastes importants de développement intra-départementaux. La politique des zones défavorisées, appliquée en Roumanie après 1990, a essayé de résoudre ces problèmes intra-départementaux mais sans grand succès.

Figure 6. *Distribution territoriale des clusters des départements*



Pour une description plus en détail des clusters par l'intermédiaire de leur niveau de développement et des ressources disponibles, dans le tableau 4 ont été calculés les valeurs moyennes standardisées obtenues pour chaque cluster. Les profils caractéristiques des clusters sont:

Cluster 1

- les valeurs proches de la moyenne pour la majorité des indicateurs (tableau 4)
- les départements de ce cluster se caractérisent par un accès modéré aux services de santé et d'enseignement, les sommes allouées pour les activités de recherche-développement par rapport au nombre d'habitants étant réduites
- l'accès aux utilités est proche de la moyenne, mais au-dessus de la moyenne à toutes les utilités, par rapport aux autres clusters qui enregistrent des poids grands des localités avec accès à une part des utilités, mais qui se confrontent avec des privations importantes à une autre part
- la même appropriation de la moyenne caractérise aussi les indicateurs socio-économiques (PIB/habitant, taux de chômage, poids de la population rurale)
- la population occupée en agriculture a un poids plus grand que dans le cas des autres clusters, mais aucune activité ne se distingue pas en comparaison avec les autres groupes des départements.

Cluster 2

- le taux le plus élevé de la mortalité infantile, associé avec la plus petite espérance de vie et le plus faible accès aux services sanitaires
- l'accès à l'enseignement est très bas
- l'infrastructure est moins développée, ayant le plus faible accès aux réseaux de gaz naturel et d'installation de canalisation publique
- les départements ont un niveau bas de développement économique (le PIB par habitant, investissements minimaux dans l'activité de recherche-développement)
- le degré d'urbanisation est réduit et le taux de chômage est élevé, avec les plus grands poids de la population occupée dans l'agriculture et des poids faibles de l'occupation dans les autres activités économiques

Cluster 3

- par opposition avec le cluster 2, les départements de ce cluster se distinguent par un niveau élevé de développement, avec un PIB/habitant élevé, des importants investissements dans les activités de recherche-développement, un niveau bas du chômage, un degré élevé d'urbanisation, la meilleure infrastructure médicale et d'enseignement, la plus grande espérance de vie et l'une des plus bas taux de mortalité infantile
- l'accès aux utilités est supérieur à la moyenne mais les réseaux de gaz naturel et de canalisation publique ne sont pas très répandus
- les plus élevés poids de la population occupée dans la santé et dans l'enseignement (dans les chefs-lieux des départements se trouvent les plus importants centres universitaires du pays)

- le poids relativement élevé de la population occupée dans les domaines des transactions immobilières et constructions
- poids bas de la population occupée dans l'agriculture

Cluster 4

- les départements du cluster sont aussi développés (valeurs élevés du PIB par habitant, dépenses élevées pour les activités de recherche-développement, accès très bon à toutes les utilités et un réseau de distribution de l'énergie thermique bien développé)
- accès élevé aux services de santé et d'enseignement mais à un niveau plus bas que les départements du cluster précédent
- les poids les plus petits de la population rurale et de la population occupée en agriculture
- les plus grands poids de la population occupée dans l'hôtellerie et restaurants (activité touristique bien développée), constructions
- le commerce et les transactions immobilières sont bien représentés par le poids occupé par la population qui travaille dans ces secteurs
- le taux le plus élevé du chômage

Tableau 4. Valeurs moyennes des variables analysées standardisées obtenues pour chaque cluster

Variable standardisée	Cluster				
	1 (BH, MS, AR, TL, MM, BR, BC, CV, HR, CS, AB, GL, GJ, PH, VL, AG)	2 (CL, OT, BT, VS, BZ, MH, IL, GR, TR, NT, SV, DB, BN, VN, SM, SJ, DJ)	3 (TM, CJ, IS)	4 (BV, SB, HD, CT)	5 (IF)
ZPIB_hab	0.24	-0.82	1.18	1.10	2.17
Zdép_rech_hab	-0.17	-0.45	1.14	0.75	4.09
Zp_rurale	-0.25	0.72	-0.80	-1.95	1.93
Zpop_agr	-0.41	0.96	-0.85	-1.61	-0.79
Zpop_constr	0.17	-0.65	1.04	1.42	-0.39
Zpop_com	0.04	-0.62	0.62	1.35	2.68
Zpop_eng	0.09	-0.53	2.39	0.48	-1.59
Zpop_san	0.06	-0.65	2.59	0.72	-0.48
Zpop_imm	0.03	-0.70	1.27	1.21	2.82
Zpop_fin	0.10	-0.58	1.49	1.41	-1.88
Zpop_hotr	0.11	-0.68	0.59	1.66	1.49
Zr_schom	-0.03	0.20	-0.84	0.39	-2.01
Zr_pop_éc	-0.03	-0.45	2.31	0.77	-1.84
Znr_fac	-0.01	-0.54	2.45	0.81	-1.20
Zr_minf	-0.17	0.43	-0.91	0.02	-1.82
Zdur_moy_vie	-0.01	-0.17	0.58	0.25	0.41
Zlits_hop	0.17	-0.67	2.16	0.61	-0.22
Zmédecins	-0.09	-0.54	2.74	0.64	-0.29
Zen_therm	0.11	-0.47	0.54	1.37	-0.80
Zgaz_nat	0.21	-0.70	-0.18	1.56	2.91
Zcanal	0.46	-0.78	0.09	0.99	1.75
Zeau_pot	0.34	-0.47	0.62	0.37	-0.78

Source: Résultats obtenus en SPSS

Cluster 5

- ce cluster est formé par un seul département, celui de Ilfov (IF), qui représente un cas un peu atypique, à cause du voisinage de la capitale
- la proximité de Bucarest explique en grand part les principales caractéristiques du département, majoritairement rural: le PIB par habitant le plus élevé, les plus grandes sommes allouées pour les dépenses de recherche-développement (présence des institutions importantes de recherche), le plus faible taux de chômage, les plus grands poids de la population occupée dans le commerce (les plus grands dépôts en-gros du pays se trouvent ici) et transactions immobilières (le *boom* immobilier est entretenu par les bucarestois qui aiment un espace plus agréable que celui d'une ville agglomérée)
- dans le cas de l'infrastructure, il y a un paradoxe: même si le département de Ilfov a les plus développés réseaux de distribution de gaz naturel et de canalisation publique, le même département a l'accès le plus faible à l'énergie thermique et aux installations d'alimentations en eau potable
- Bucarest a les plus nombreux et les plus importants centres médicaux et d'enseignement. La population du département de Ilfov bénéficie aussi de ces services et c'est pour ça que le département n'a pas une infrastructure très développée dans le domaine de la santé. La plus grande espérance de vie et le plus bas taux de la mortalité infantile sont des autres caractéristiques du département de Ilfov.

Conclusions

Les clusters déterminés sont partiellement homogènes dans l'intérieur mais très différents en ce qui concerne les caractéristiques et le niveau de développement des départements au niveau inter-clusters.

La solution avec 5 clusters donne la possibilité d'identification des directions principales, sur la base duquel peuvent être mises les bases des programmes, des politiques et des stratégies de développement:

- le développement du réseau de distribution du gaz naturel et de canalisation publique du cluster 3, formé par les plus développés départements (Timis – TM, Cluj – CJ, Iasi – IS)
- le principal problème du cluster 4, même si le niveau de développement est élevé, est le taux élevé du chômage. Par conséquent, les mesures qui doivent améliorer le développement visent la stimulation de l'entrepreneuriat dans le développement des affaires dans des domaines moins valorisés, pour créer des nouveaux emplois, qui assurent la rémunération de la population sur une période longue et toute l'année et pas d'une manière saisonnière. Dans le département de Hunedoara (HD), le taux élevé du chômage s'explique pas seulement par la saisonnalité touristique mais par les la récession du secteur minier (extraction du charbon) après 1990. Les nouveaux emplois sont plus que nécessaires dans ce département pour offrir une alternative professionnelle aux mineurs qui ont été obligés de quitter l'industrie minière
- l'amélioration du développement dans le département de Ilfov (cluster 5) suppose le développement du réseau de distribution de l'énergie thermique et

d'alimentation en eau potable mais aussi le développement de l'infrastructure de santé et d'enseignement

- des mesures spécifiques d'amélioration des indicateurs analysés pour le cluster des départements développés modérément (cluster 1).
- l'allocation des ressources importantes et l'implémentation des programmes radicaux pour le cluster des départements qui sont les moins développés (cluster 2). Les mesures doivent viser premièrement l'amélioration de la qualité de vie par un accès meilleur aux services sanitaires (l'attraction des cadres médicaux dans les zones défavorisées, l'allocation des ressources pour la construction des nouvelles unités médicales et l'extension des unités existantes) et le développement de l'infrastructure. Une autre solution vise les facilités pour les firmes qui veulent investir dans les zones défavorisées, le but essentiel étant la dynamisation de l'activité économique de ces départements.

Bibliographie

Chih-Kai Chen, (2008) - Construct Model of Knowledge-based Economy Indicators, Journal of American Academy of Business, Cambridge; Mar 2008; 13, 1; ABI/INFORM Global, pg. 215.

D'Alimonte Davide , Cornford Dan , (2008) - Outlier detection with partial information: application to emergency mapping, Stochastic Environmental Research and Risk Assessment. Berlin: Aug 2008. Vol. 22, N° 5; pg. 613, 8 pgs.

Del Campo Cristina , Monteiro Carlos M. F. , Soares Joao Oliveira , (2008) - The European regional policy and the socio-economic diversity of European regions: A multivariate analysis, European Journal of Operational Research. Amsterdam: Jun 1, 2008. Vol. 187, N° 2; pg. 600.

Jaba, E. (2007), Regula „3 sigma” pentru identificarea disparităților interregionale, în Evaluarea statistică a dezvoltării economico-sociale (coord. Elisabeta Jaba), Editura Junimea, Iași

Jaba, E., Grama, A. (2004), Analiza statistică cu SPSS sub Windows, Editura Polirom, Iași.

Jaba, E., Șerban, D., Viorică, E.D., Balan, C.B. (2006), Analiza discriminant, metodă statistică de bază pentru evaluarea campaniilor de marketing, articol în Revista Română de Statistică nr.4/2006

Lachenbruch, P. A., Mickey, M. R. (1968) - Estimation of error rates in discriminant analysis. Technometrics 10, 1-10.

Schott James R. , (2006) - A high-dimensional test for the equality of the smallest eigenvalues of a covariance matrix, Journal of Multivariate Analysis. New York: Apr 2006. Vol. 97, N° 4; pg. 827.

Solomonoff Ray J., (2008) - The probability of "undefined" (non-converging) output in generating the universal probability distribution, Information Processing Letters. Amsterdam: Jun 15, 2008. Vol. 106, N° 6; pg. 238.

Vaughn Brandon K., Wang Qui, (2008) - Classification Based on Tree-Structured Allocation Rules, The Journal of Experimental Education, 76(3), 315–340, Heldref Publications.