

ANALÝZA REGIONÁLNÍHO ROZVOJE S VYUŽITÍM VÍCEROZMĚRNÝCH STATISTICKÝCH METOD

Svatošová, L., Prášilová, M.

Došlo: 29. června 2006

Abstract

SVATOŠOVÁ, L., PRAŠILOVÁ, M.: *Analysis of regional development with use of multivariate statistical methods*. Acta univ. agric. et silvic. Mendel. Brun., 2006, LIV, No. 6, pp. 171–176

The paper deals with differentiation of regional entities within the Czech Republic based on study of human potential. The human factor has been defined by 22 variables from three domains: population density, demographic indicators and economic activities of inhabitants. The variables have been recorded by regions and selected districts of the C.R. in 1994–2004, for computation purposes they have been represented by their averages and standardized. Principal component method has been employed for solution, facilitating to reduce number of the variables without any considerable loss of information, to select the most significant factors for a given area and to aggregate the variables into larger groups (principal components). Two extensive methods have been constructed. The first one has been based on the C.R. regions, the second one on the data from the Vysočina Region districts. Solution results demonstrate different roles of the separate aggregate variables in regional development. While in the C.R. as a whole, the most difficult problem is population ageing, growth of urban population and unemployment, in the Vysočina Region it is the development of small villages and of countryside as a whole, and unemployment. The method used is suitable generally for study and assessment of regional development and it brings many objective information for decision-making process.

regional differentiation, regional development, human potential, community, countryside, Vysočina Region, demography, unemployment, Principal Component Method

Základním úkolem vedení každé demokratické společnosti je její rovnoměrný rozvoj. Česká republika je rozdělena do čtrnácti krajů, které mají svůj rozdílný ekonomický a historický základ. Uvnitř krajů je však situace obdobná. Rozvoj území je možno popsat řadou ukazatelů, z nichž některé jsou měřitelné, jiné neměřitelné a řada působí náhodně. Hodnocení krajů a dílčích území zajišťuje centrální statistika pomocí sběru statistických dat. Tato data se prolínají, spolu souvisejí a mohou být ve vzájemném vztahu. Úroveň souvislostí a vazeb mezi ukazateli lze odhalit pomocí statistických metod a analýz. V práci byla využita vícerozměrná metoda hlavních komponent, která pracovala s údaji charakterizujícími lidský potenciál na úrovni krajů a dále s údaji okresů v kraji Vysočina. Lidský faktor hraje v rozvoji regionů významnou roli

a přes všechny ekonomické a politické změny zůstává nejdůležitějším faktorem stability oblasti. Řešení má ukázat na možnosti statistických metod při hodnocení a diferenciaci regionů z hlediska souboru ukazatelů.

MATERIÁL A METODY

Rovnoměrný rozvoj regionů v současné době představuje jeden ze základních úkolů regionální politiky státu a Evropské unie. Jedná se o komplex procesů přispívajících k pozitivním změnám v ekonomické, sociální a environmentální situaci regionů s cílem dosažení ekonomické a sociální soudržnosti.

Sledování a hodnocení indikátorů regionálního rozvoje je věnována značná pozornost. V rámci komparace rozvoje regionů v Evropské unii bylo doporučeno sledovat původně 42 ukazatelů, v současné

době se tento počet zúžil na 14 syntetizujících indikátorů, pomocí nichž je sledována ekonomická úroveň, zaměstnanost, vzdělanost, podnikatelské aktivity, příjmová situace obyvatelstva a životní prostředí regionu.

Pro komplexní hodnocení rozvojového potenciálu regionů v rámci České republiky je vhodné vycházet z detailnějšího hodnocení a využít širší škálu ukazatelů, pomocí nichž lze charakterizovat stav a vývoj jednotlivých oblastí. Jedná se zejména o ukazatele charakterizující:

- demografickou situaci
- ekonomickou aktivitu obyvatelstva
- rozvoj výroby a služeb
- vývoj investic
- vývoj HDP
- životní úroveň obyvatelstva
- stav a ochranu životního prostředí
- infrastrukturu v regionu
- sociální ukazatele
- rozvoj venkova a zemědělství.

Systém potom zahrnuje řadu ukazatelů, na jejichž základě lze provádět celou řadu hodnocení, jako například:

- místní porovnání úrovně hodnot ukazatele
- analýzu trendů v čase
- vzájemné souvislosti změn jednotlivých ukazatelů
- stanovení prognóz
- stanovení agregovaných ukazatelů postihujících jednotlivé zkoumané oblasti regionálního rozvoje.

Prvním úkolem pro dosažení relevantních závěrů je výběr ukazatelů, které mají zásadní vliv na rozvojový potenciál dané oblasti zkoumání. Při zpracovávání rozsáhlého souboru ukazatelů, kdy výchozí počet znaků je značný a pro interpretaci málo přehledný, můžeme s výhodou užít některou z metod vícerozměrných statistických analýz. Redukovaný soubor proměnných může být posléze zpracován jednodušším způsobem a zejména zaručuje výběr takových ukazatelů, které na rozvoj dané oblasti mají zásadní vliv.

K redukci nadměrného počtu proměnných, respektive zhuštění informací do menšího počtu neměřitelných hypotetických veličin s co nejmenší ztrátou informace, která je obsažena ve sledovaných znacích, lze použít analýzu hlavních komponent.

Principem **analýzy hlavních komponent** je transformace souboru původních proměnných do nových hypotetických proměnných. Tyto proměnné nazýváme komponenty. Komponenty jsou seřazeny podle velikosti svého příspěvku k vysvětlení celkového rozptylu původních proměnných. Jde tedy v zásadě o to, že vytvoříme z původního velkého množství dat proměnné obecnější, ale ztráta informace, kterou obsahují původní proměnné, je minimální.

Základní model je vyjádřen prostřednictvím normovaných proměnných a neobsahuje reziduum:

$$K_i = a_{i1}u_1 + a_{i2}u_2 + \dots + a_{ip}u_p,$$

kde K_i – hodnota i -té komponenty ($i = 1 \dots p$)

u_1, u_2, \dots, u_p – hodnoty normovaných proměnných

$a_{i1}, a_{i2}, \dots, a_{ip}$ – váhy jednotlivých proměnných i -té komponenty.

V modelech, kde proměnné mají stejnou či podobnou úroveň a stejnou měrnou jednotku, není nutné provádět normování. Potom i interpretace komponent může být konkrétnější.

Metodický postup lze charakterizovat následovně:

- vytvoření matice dat
- výpočet kovarianční matice
- vyjádření charakteristických čísel jím odpovídajících charakteristických vektorů kovarianční matice
- vytvoření hlavních komponent
- vyjádření podílu celkového rozptylu vyčerpaného jednotlivými komponentami.

Podrobný postup analýzy hlavních komponent obsahuje např. HEBÁK (Vícerozměrné statistické metody (3)). Výhodou použití této vícerozměrné metody je to, že získáme nejenom soubor redukovaných proměnných, ale z výsledků analýzy hlavních komponent lze stanovit i jejich významnost (váhu), což lze použít například při konstrukci agregovaných ukazatelů.

VÝSLEDKY

Výše uvedená metodika byla použita pro zkoumání stavu lidského potenciálu v regionech. Lidský potenciál a podmínky jeho rozvoje je nutné posuzovat z několika hledisek (oblastí), z nichž každé zahrnuje řadu ukazatelů, které mají vliv na úroveň dané oblasti.

Oblasti zkoumání lidského potenciálu:

1. Demografický vývoj – střední stav obyvatelstva, hustota obyvatelstva, rozložení obyvatelstva dle místa bydliště, věková struktura obyvatelstva, sňatečnost, rozvodovost, přirozený přírůstek obyvatelstva, migrace, průměrný věk obyvatelstva, index stárí.
2. Ekonomická struktura obyvatelstva – míra ekonomické aktivity, index hospodářského zatížení, index závislosti mladých, index závislosti starých.
3. Úroveň zaměstnanosti – míra nezaměstnanosti, struktura nezaměstnaných dle věku, pohlaví, vzdělání, délky nezaměstnanosti.
4. Životní úroveň obyvatelstva – průměrná mzda celkem a v členění dle sektorů, struktura výdajů obyvatelstva.

5. Vzdělanostní úroveň obyvatelstva – struktura obyvatelstva dle stupně vzdělání.
6. Sociální úroveň – průměrný měsíční objem vyplácených dávek sociální podpory, počty bytů, celková plocha bytu na 1 obyvatele, počet obyvatel v zařízeních sociální péče, počet učitelů na 1 žáka či studenta, počet obyvatel na 1 lékaře, počet nemocničních lůžek, kriminalita.
7. Zdravotní stav obyvatelstva – střední délka života, kojenecká úmrtnost, struktura zemřelých dle příčiny nemoci, dle věkové struktury.

Pro demonstraci metodiky byly vybrány ukazatele charakterizující první tři oblasti zkoumání lidského potenciálu:

- Hustota obyvatelstva
- Počet obcí na 100 km²
- Podíl měst na celkovém počtu obcí
- Podíl obcí do 200 obyvatel na celkovém počtu obcí
- Podíl obcí s 200–500 obyvateli na celkovém počtu obcí
- Podíl obcí s 500–2000 obyvateli na celkovém počtu obcí
- Podíl obyvatel ve městech z celkového počtu obyvatel regionu
- Podíl obyvatel v obcích do 200 obyvatel z celkového počtu obyvatel regionu
- Podíl obyvatel v obcích s 200–500 obyvateli z celkového počtu obyvatel regionu
- Podíl obyvatel v obcích s 500–2000 obyvateli z celkového počtu obyvatel regionu
- Počet narozených na 1000 obyvatel regionu
- Počet zemřelých na 1000 obyvatel regionu
- Přirozený přírůstek na 1000 obyvatel regionu
- Přírůstek stěhováním na 1000 obyvatel regionu
- Index stárí (podíl obyvatel ve věku do 15 let k počtu obyvatel ve věku nad 65 let)
- Míra nezaměstnanosti
- Podíl nezaměstnaných ve věku do 25 let z celkového počtu nezaměstnaných
- Podíl nezaměstnaných ve věku 45–60 let z celkového počtu nezaměstnaných
- Podíl nezaměstnaných s délkou nezaměstnanosti delší než 1 rok z celkového počtu nezaměstnaných
- Průměrný věk žen
- Průměrný věk mužů
- Míra ekonomické aktivity obyvatelstva (počet ekonomicky aktivních ke střednímu stavu obyvatelstva).

Proměnné byly sledovány za kraje a okresy ČR v letech 1999–2004. Do modelu analýzy hlavních komponent, který posloužil k redukci proměnných

a zejména k posouzení jejich vypovídací schopnosti, byly použity průměry za toto období. Analýza hlavních komponent sledovala dva cíle:

1. redukci proměnných – výběr ukazatelů, které měly v daném období největší vliv na rozvoj regionu – kraje z daného hlediska,
2. stanovení ukazatelů, jejichž vliv lze charakterizovat jako obecný, z hlediska regionálního rozvoje, tedy platných ve všech krajích a stanovení specifických ukazatelů, jejichž rozvoj má význam pro jednotlivé kraje.

Ukazatele obecně platné pro rozvoj regionů získáme z modelu analýzy hlavních komponent, do kterého byly zařazeny průměry ukazatelů za kraje ČR, ukazatele specifické pak z modelů analýzy hlavních komponent pro jednotlivé kraje, kde vycházíme z průměrných ukazatelů za okresy.

Výsledky analýzy hlavních komponent – model pro kraje ČR

Model analýzy hlavních komponent charakterizující úroveň lidského potenciálu v krajích byl vysvětlen z 90,5 % pomocí pěti komponent. Ukazatele, jejichž vliv je v jednotlivých komponentách zásadní, jsou uvedeny v tabulce I.

Vybrané ukazatele postihují problémové oblasti vlastní všem regionům. Nejsilněji se zde odráží problémy stárnutí populace, růst městské populace a nezaměstnanost obyvatel. Vytvoření podmínek pro růst přirozeného přírůstku obyvatelstva, programy na rozvoj venkovského osídlení a řešení nezaměstnanosti v regionu jsou součástí rozvojových strategií všech regionů ČR.

Výsledky analýzy hlavních komponent – model pro kraj Vysočina

Model analýzy hlavních komponent charakterizující úroveň lidského potenciálu v kraji Vysočina byl vysvětlen z 99,9 % pomocí čtyř komponent. Tabulka II shrnuje výsledky za kraj Vysočina.

V modelu analýzy hlavních komponent kraje Vysočina se kromě ukazatelů obecného charakteru objevují i proměnné pro kraj specifické. Jde zejména o podíl obcí do 200 obyvatel na celkovém počtu obcí, neboť z hlediska rozmístění obyvatelstva patří kraj Vysočina ke krajům s vysokým podílem obyvatelstva žijícího v malých obcích. Počet obcí činí 10,53 na 100 km² zatímco průměr ČR činí 7,92 obce na km². Podíl obyvatel žijících ve městech činí 57,8 %. 33,2 % obyvatel kraje Vysočina žije v obcích do 2000 obyvatel, 8,8 % obyvatel kraje pak žije v malých obcích do 200 obyvatel. Tento podíl vysoce překračuje průměr ČR, který činí 2 %.

I: Výsledky analýzy hlavních komponent (kraje ČR), průměr let 1999–2004

Komponenta	% z celkové variability	Ukazatel	Koeficient korelace s komponentou
I.	42,2	Index stáří	0,289
		Podíl měst na celkovém počtu obcí	0,290
		Podíl obyvatel v obcích s 200–500 obyvateli z celkového počtu obyvatel regionu	0,311
		Průměrný věk mužů	0,280
		Podíl městského obyvatelstva z celkového počtu obyvatel regionu	0,290
II.	20,8	Podíl obcí s 500–2000 obyvateli z celkového počtu obcí	0,460
		Míra ekonomické aktivity obyvatelstva	0,360
		Hustota obyvatelstva	0,310
III.	13,2	Podíl nezaměstnaných ve věku 45–60 let z celkového počtu nezaměstnaných	0,576
		Podíl nezaměstnaných ve věku do 25 let z celkového počtu nezaměstnaných	0,447
IV.	9,0	Počet narozených na 1000 obyvatel	0,444
		Počet zemřelých na 1000 obyvatel	0,572
V.	5,3	Průměrný věk žen	0,413
		Podíl nezaměstnaných s dobou nezaměstnanosti delší než 1 rok z celkového počtu nezaměstnaných	0,410

II: Výsledky analýzy hlavních komponent (kraj Vysočina), průměr let 1999–2004

Komponenta	% z celkové variability	Ukazatel	Koeficient korelace s komponentou
I.	58,2	Index stáří	0,287
		Průměrný věk žen	0,288
		Průměrný věk mužů	0,284
		Počet obcí na 100 km ²	0,274
		Podíl obcí do 200 obyvatel na celkovém počtu obcí	0,261
II.	22,1	Podíl obyvatel v obcích s 500–2000 obyvateli z celkového počtu obyvatel regionu	0,447
		Hustota obyvatelstva	0,404
		Podíl nezaměstnaných ve věku do 25 let z celkového počtu nezaměstnaných	0,415
III.	12,2	Podíl obcí s počtem obyvatel 500–2000	0,437
		Podíl obyvatel v obcích s 200–500 obyvateli z celkového počtu obyvatel regionu	0,426
IV.	7,4	Podíl nezaměstnaných s dobou nezaměstnanosti delší než 1 rok z celkového počtu nezaměstnaných	0,468
		Míra nezaměstnanosti	0,331

Dalším ukazatelem, který je z hlediska rozvoje regionu důležitým, je podíl nezaměstnaných s dobou nezaměstnanosti delší než 1 rok. Ukazatel odráží situaci zejména v okresech Jihlava, Třebíč a Žďár nad

Sázavou, kde je vysoká nezaměstnanost a podíl nezaměstnaných delší dobu než 1 rok vysoce překračuje celorepublikový průměr.

Prioritou strategických opatření rozvoje regionu Vysočina by tedy měla být opatření na rozvoj venkova, zejména malých obcí, a řešení problému nezaměstnanosti v kraji.

DISKUSE

Uvedená metodika je obecně použitelná pro všechny oblasti sledování a hodnocení regionálního rozvoje. Umožňuje jak redukci velkého počtu ukazatelů na menší počet relevantních ukazatelů, aniž dojde k podstatné ztrátě informace, tak i definování ukazatelů z hlediska regionálního rozvoje obecných a ukazatelů specifických pro rozvoj jednotlivých regionů.

Vybrané ukazatele pak mohou být podrobeny detailním analýzám. Analyzovat lze každý ukazatel odděleně – porovnat jeho úroveň, variabilitu, vývoj

v čase, nebo ve vztahu k ostatním vybraným ukazatelům pomocí metod regresní a korelační analýzy.

Při porovnávání stavu a vývoje jednotlivých regionů (krajů) či okresů v rámci kraje je možné na základě vybraných ukazatelů vytvořit i ukazatel agregovaný. Jednoduchá forma tohoto ukazatele shrnuje vybrané proměnné se stejnou vahou, pro vytvoření vážené formy ukazatele poslouží váhy jednotlivých komponent a koeficienty korelace vybraných ukazatelů s komponentou. Pomocí takto konstruovaných agregovaných ukazatelů lze provádět meziregionální porovnání, stanovení pořadí regionů či skupin regionů z hlediska vývoje dané oblasti nadprůměrných, průměrných, podprůměrných či definování kritických oblastí rozvoje jednotlivých regionů.

SOUHRN

Příspěvek se zabývá diferenciací územních celků České republiky na základě zkoumání lidského potenciálu. Lidský faktor byl definován 22 proměnnými ze tří oblastí. Jedná se o oblast zalidněnosti, o skupinu demografických ukazatelů a dále o oblast ekonomické aktivity obyvatel. Proměnné byly sledovány za kraje a vybrané okresy České republiky v letech 1994–2004 a pro účely výpočtů byly zastoupeny svými průměry a byly standardizovány. K řešení bylo použito metody hlavních komponent, která umožnila zredukovat počet proměnných bez větší ztráty informace, vybrat pro dané území faktory nejvýznamnější a proměnné agregovat do větších skupin (hlavních komponent). Byly sestaveny dva rozsáhlé modely. První vycházel z údajů krajů České republiky a druhý z údajů okresů kraje Vysočina. Výsledky řešení ukazují na rozdílné role jednotlivých agregovaných proměnných v rozvoji oblastí. Zatímco v České republice jako celku je nejpalčivějším problémem stárnutí populace, růst městské populace a nezaměstnanost, v kraji Vysočina je to rozvoj malých obcí a venkova jako celku a nezaměstnanost. Použitá metodika je obecně vhodná pro sledování a hodnocení regionálního rozvoje a přináší řadu objektivních podkladů pro proces rozhodování.

diferenciace oblastí, rozvoj regionů, lidský potenciál, obec, venkov, kraj Vysočina, demografie, nezaměstnanost, agregace proměnných, metoda hlavních komponent

Příspěvek vznikl v rámci řešení projektu QG60030 „Strategie využití sociálního a lidského kapitálu při revitalizaci venkova v kraji Vysočina“.

LITERATURA

HEBÁK, P., HUSTOPECKÝ, J., PECÁKOVÁ, I., PRŮŠA, M., ŘEZANKOVÁ, H., SVOBODOVÁ, A., VLACH, P.: Vícerozměrné statistické metody (3), 1. vyd., Informatorium, 2005, 255 s., ISBN 80-7333-039-3,
MELOUN, M., MILITKÝ, J., HILL, M.: Počítačová

analýza vícerozměrných dat v příkladech, 1. vyd., Academia, 2005, 449 s., ISBN 80-200-1335-0
SVATOŠOVÁ, L., BOHÁČKOVÁ, I., HRABÁNKOVÁ, M.: Regionální rozvoj z pozice strukturální politiky, 1. vyd., JU České Budějovice, 2005, 173 s., ISBN 80-7040-749-2

Adresa

Doc. Ing. Libuše Svatošová, CSc., Ing. Marie Prášilová, CSc., Katedra statistiky, Česká zemědělská univerzita v Praze, Kamýcká 129, 165 21 Praha 6-Suchbát, Česká republika, e-mail: svatosova@pef.czu.cz, prasilova@pef.czu.cz

