



---

**Commission économique pour l'Europe****Conférence des statisticiens européens****Soixante-huitième réunion plénière**

Genève, 22-24 juin 2020

Point 5 b) de l'ordre du jour provisoire

**Session 2 : Enseignements et résultats des mesures concrètes déjà prises  
par les organismes nationaux de statistique et les spécialistes  
de l'information géospatiale pour moderniser leur rôle****Exploiter les multiples possibilités offertes par l'intégration  
des données géospatiales et statistiques<sup>1</sup>****Note établie par Statistics Poland***Résumé*

Le présent document est soumis à la Conférence des statisticiens européens, pour examen lors du séminaire sur les nouveaux rôles des organismes nationaux de statistique et d'information géospatiale dans les récents écosystèmes nationaux de données, organisé dans le cadre de la session 2 « Enseignements et résultats des mesures concrètes déjà prises par les organismes nationaux de statistique et les spécialistes de l'information géospatiale pour moderniser leur rôle ».

---

<sup>1</sup> Il a été convenu que le présent document serait publié après la date normale de publication en raison de circonstances indépendantes de la volonté du soumetteur.



## I. Introduction

1. L'organisme polonais de statistique produit et utilise des données géospatiales depuis plusieurs années. Prenant en compte les normes modernes en matière de technologies de l'information et les outils technologiques nécessaires aux recensements et aux enquêtes, Statistics Poland a entrepris de cartographier les données statistiques en se fondant sur des références spatiales et des systèmes d'information géographique. Les informations géospatiales prennent une place de plus en plus importante dans le processus de production de statistiques et consolident le rôle des organismes de statistique dans l'intégration des données géospatiales et statistiques.

## II. Modèle de production de statistiques

2. Statistics Poland élabore actuellement un modèle de production de statistiques en se fondant sur le Modèle générique du processus de production statistique (GSBPM)<sup>2</sup> et en l'enrichissant avec des composantes géospatiales. Afin de pouvoir analyser avec précision si les processus appliqués par Statistics Poland pour produire des données géospatiales sont conformes à ceux que prévoit le GSBPM, les processus de production ont été mappés par rapport à ce Modèle, qui est une norme de référence en la matière.

3. L'analyse a reposé sur une comparaison des processus de chaque phase du GSBPM avec ceux actuellement suivis, ou prévus, pour produire des statistiques, l'accent étant mis sur les données géospatiales.

4. Pour réaliser l'analyse de conformité, les diverses activités menées par Statistics Poland ont été mappées par rapport aux phases, processus et sous-processus recommandés dans le GSBPM. L'application des processus relatifs aux données géospatiales et leur mise en correspondance avec le modèle GSBPM a révélé que celui-ci ne prenait pas en compte certains domaines importants et qu'il souffrait donc de potentielles lacunes. Ces lacunes concernaient principalement les aspects liés à la « spatialisation » des données statistiques, de l'étape de la conception de la collecte de données à celles du géocodage, de l'analyse et de la diffusion de produits statistiques présentant des caractéristiques géospatiales. Les sous-processus géospatiaux qui devraient être appliqués dans l'architecture de la production statistique sont les suivants :

- Phase 2, sous-processus 2.5a « conception du cadre de géocodage, de l'échantillon et de la collecte » ;
- Phase 4, sous-processus 4.1a « Géocodage du cadre et de l'échantillon » et 4.3a « géocodage de la collecte » ;
- Phase 6, sous-processus 6.2a « Préparation des analyses géospatiales et des cartes » ;
- Phase 7, sous-processus 7.2a « Gestion des analyses géospatiales et de cartes au moyen de systèmes d'information géographique ».

5. L'ajout des sous-processus énumérés ci-dessus permettra d'intégrer une composante géospatiale dans le modèle de production de statistiques, ce qui contribuera à mieux appréhender les données géospatiales et leur rôle dans la production de statistiques, ainsi qu'à normaliser les méthodes ayant trait à la fusion des données statistiques et des données géospatiales. Ces nouveaux sous-processus géospatiaux (les blocs jaunes dans le schéma ci-dessous) se retrouvent dans le modèle polonais de production de statistiques, sur lequel se fondent la production statistique et l'intégration des éléments statistiques et géospatiaux, l'objectif étant d'améliorer et d'approfondir l'analyse géospatiale.

---

<sup>2</sup> <https://statswiki.unece.org/display/GSBPM/IV.+Levels+1+and+2+of+the+GSBPM>.

## Schéma

## Modèle de production de statistiques

1 Specify Needs	2 Design	3 Build	4 Collect	5 Process	6 Analyse	7 Disseminate	8 Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instrument	4.1 Create frame & select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update outputs systems	8.1 Gather evaluation inputs
1.1a Prepare Statistical Programme (SSPPS)	2.2 Design variable descriptions	3.2 Reuse or build processing & analysis components	4.1a Geocode frame & sample	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.2 Consult & confirm needs	2.3 Design collection	3.3 Reuse or build dissemination components	4.2 Set up collection	5.3 Review & validate	6.2a Prepare spatial analyses & maps	7.2a Manage spatial analyses & maps using GIS	8.3 Agree an action plan
1.3 Establish output objectives	2.4 Design frame & sample	3.4 Configure workflows	4.3 Run collection	5.4 Edit & input	6.3 Interpret & explain outputs	7.3 Manage release of dissemination products	
1.4 Identify concepts	2.5 Design processing & analysis	3.5 Test production system	4.3a Geocode collection	5.5 Derive new variables & units	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5a Design geocoding frame, sample & data collection	3.6 Test statistical business process	4.4 Finalise collection	5.6 Calculate weights	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production system & workflow	3.7 Finalise production system	4.5 Verify plans	5.7 Calculate aggregates		7.6 Satisfaction research	
1.7 Prepare a general plan for the implementation of the production process	2.7 Design quality indicators	3.8 Verify projects and construction plans		5.8 Finalise data files		7.7 Prepare an improvement plan for the next research cycle	
1.8 Prepare a detailed plans for subsequent phases	2.8 Verify the general plan and detailed plans						
1.9 Verify the concept to improve satisfaction							

### III. Les systèmes d'information géographique à l'appui du processus de production de statistiques

6. Les composantes géospatiales jouent un rôle de plus en plus important dans le processus polonais de production de statistiques. Il faut en effet disposer de données statistiques géolocalisées de manière très précise pour pouvoir réaliser des analyses qui mettent en lumière la dimension territoriale de l'évolution de différents phénomènes. Seul l'organisme national de statistique collecte toutes les unités de données localisées par point (coordonnées x et y) qui sont indispensables au traitement et à la production d'agrégats et permettent de réaliser des analyses géospatiales fiables et de qualité. Il est aussi crucial de disposer de données géoréférencées au moyen de coordonnées x et y pour les recensements.

7. Statistics Poland a eu recours à un système d'information géographique pour présenter les résultats du recensement agricole de 2010 et du recensement de la population et du logement de 2011. Afin de faciliter la diffusion et la visualisation de données statistiques géoréférencées, le Portail géostatistique<sup>3</sup> a été créé à l'intention des utilisateurs internes et externes.

8. Le Portail géostatistique est un outil interactif permettant de présenter sur des cartes les statistiques agrégées produites à partir de divers types de données et d'analyses géospatiales. Conçu pour recueillir des informations et les mettre à disposition d'un large public, y compris des autorités publiques, des entrepreneurs, des utilisateurs individuels et des instituts de recherche, le Portail s'avère très utile pour éclairer la prise de décisions stratégiques à tous les niveaux d'administration. Cet outil global est conforme aux normes

<sup>3</sup> <https://geo.stat.gov.pl/imap/?locale=en>.

européennes et toutes les données sont traitées de manière à les dépersonnaliser, afin de préserver leur caractère confidentiel.

9. À l'heure actuelle, Statistics Poland tire profit des nombreuses possibilités qui se font jour pour développer le Portail géostatistique dans le cadre d'un nouveau projet, à savoir la conception d'un outil innovant et ultramoderne qui lui permettra d'assurer son rôle d'organisme de production de statistiques géospatiales et de mener des analyses plus complexes et plus approfondies.

10. Intitulé « Statistiques géospatiales et système national d'information », le projet de développement vise à élargir la portée et améliorer la disponibilité des statistiques et méthodes d'analyse géostatistique en faisant appel aux ressources dont dispose l'organisme officiel de statistique. Lancé pour répondre aux besoins exprimés après la mise en ligne de l'actuel Portail géostatistique, ce projet prend en compte de nouvelles sources de données et de nouveaux outils informatiques et logiciels libres. Dans ce contexte, Statistics Poland s'emploie à concevoir une plateforme informatique moderne qui viendra favoriser le recours aux systèmes d'information géographique dans les études statistiques et permettra de mener des analyses plus complexes et plus approfondies.

11. Dans le cadre de ce nouveau projet, les fonctionnalités de l'actuel Portail géostatistique seront étendues et incluront notamment la possibilité de réaliser des analyses statistiques pour toutes les divisions spatiales, qu'elles soient définies par l'utilisateur, téléchargées à partir de services externes de données géospatiales ou fondées sur des cellules « dynamiques », ainsi que la possibilité de combiner des données avec celles de l'utilisateur ou de géocoder les caractéristiques de l'utilisateur pour s'en servir dans l'analyse géostatistique. En outre, de nouvelles fonctionnalités seront mises au point afin de faciliter l'analyse exploratoire à partir des données géostatistiques et la modélisation géostatistique et de compléter les données des utilisateurs avec des données et analyses géostatistiques.

12. Ce projet permettra donc d'intégrer trois fonctionnalités externes entièrement nouvelles :

- L'analyse exploratoire à partir des données géostatistiques disponibles sur le Portail ;
- La modélisation géostatistique ;
- L'enrichissement des contenus générés par les utilisateurs avec des données et analyses géostatistiques disponibles sur le Portail.

13. Les fonctionnalités d'analyse exploratoire à partir des données géostatistiques disponibles sur le Portail et de modélisation géostatistique permettront aux utilisateurs :

- D'importer des données statistiques disponibles sur le Portail afin de compléter celles dont ils disposent déjà ;
- De compléter les données avec des variables supplémentaires (aussi appelées variables dérivées) calculées sur la base des données qu'ils auront déjà importées au moyen des outils mathématiques et géostatistiques disponibles ;
- D'étudier les données collectées (et complétées par les variables dérivées) selon une perspective géostatistique, d'évaluer leur qualité, de les compléter et de les modifier ;
- De visualiser et de publier les résultats des analyses.

14. Un vaste éventail de méthodes d'analyse statistico-spatiale sera mis à disposition, notamment : la mesure de la tendance centrale, la mesure de la dispersion, l'analyse par grappe et l'autocorrélation spatiale. Grâce à la fonctionnalité de modélisation géostatistique, il sera également possible de concevoir et d'appliquer un modèle probabiliste qui permettra d'estimer la valeur de la variable dépendante sur la base des résultats de l'échantillonnage probabiliste (c'est-à-dire sur la base des données collectées).

15. L'enrichissement des contenus générés par les utilisateurs avec les données et analyses géostatistiques disponibles sur le Portail permettra aux utilisateurs :

- D'importer le contenu qu'ils ont généré pour l'analyser ;
- D'analyser le contenu généré en utilisant des outils de « fouille de texte » pour repérer les mots clefs renvoyant à des catégories d'objets disponibles sur le Portail et de catégoriser ce contenu de manière appropriée ;
- De sélectionner les objets disponibles sur le Portail qui correspondent le mieux aux mots et expressions clefs repérés et d'établir des corrélations entre les objets recherchés sur le Portail, afin de pouvoir choisir parmi les différents types d'analyse disponibles sur le Portail ;
- De configurer les analyses géospatiales qu'ils auront sélectionnées et de les présenter sous forme de graphique ;
- D'avoir accès aux résultats des analyses géospatiales sélectionnées sous forme de graphique ou de tableau.

16. Qui plus est, le projet de développement du Portail permettra d'intégrer une nouvelle fonctionnalité interne qui facilitera la maintenance des bases de données d'adresses.

#### **IV. Conclusions et recommandations**

17. Il est essentiel de disposer de données statistiques géolocalisées avec précision pour pouvoir réaliser des comparaisons et améliorer la production de statistiques géospatiales. Les organismes de statistique peuvent tirer profit des possibilités multiples qui s'offrent à eux pour accroître leur rôle en créant des outils d'analyse statistique ultramodernes.

18. Le Portail géostatistique, qui est la source de statistiques géospatiales la plus importante de Pologne, permet de publier des données statistiques agrégées selon diverses modalités d'analyse spatiale et de les présenter sur des cartes tout en respectant la confidentialité. Le nouveau Portail permettra de mener des analyses très approfondies, notamment des analyses exploratoires à partir de données statistiques, et d'établir des modèles géostatistiques. Le point fort de l'actuel Portail et de sa nouvelle version, et ce qui leur confère ce caractère exceptionnel, c'est leur capacité de réaliser des analyses géospatiales fiables à partir des seules unités de données dont dispose l'organisme national de statistique.

19. Le nouveau Portail sera conçu selon le modèle polonais de production de statistiques et fera partie de l'architecture plus vaste de la production de statistiques officielles en Pologne. Il permettra de publier les données statistiques et les résultats des analyses géostatistiques, qui sont essentiels au bon fonctionnement des pouvoirs publics, à l'échelle nationale et locale, et des communautés locales. Les données seront commodément présentées dans des graphiques, ce qui accélèrera le processus de prise de décisions. L'autre objectif sera de présenter sur ce Portail les résultats des recensements du cycle 2020-2021.