

STATISTIQUES ENVIRONNEMENTALES

Notes de cours

Intervenant : Franck Cachia

Période : 7-12 décembre 2015

1. Introduction, définitions et enjeux

1.1. Historique

- Juin 1972 (Stockholm) – Conférence des Nations Unies sur l'environnement humain
=> Prise de conscience
- 1984 – Premier Cadre de Développement des Statistiques de l'Environnement
=> Les thématiques et méthodes en matière de statistiques environnementales sont définies
- 1973 – Réunions de la Commission Économique pour l'Europe
=> Premières initiatives pour développer les statistiques environnementales au niveau international
- Juin 1992 (Rio de Janeiro) – Conférence des Nations Unies sur l'environnement et le développement
=> Consensus sur l'intégration de la dimension environnementale dans les politiques de développement et émergence du concept de développement durable
- Août 2002 (Johannesburg) – Sommet mondial sur le développement durable
=> Souligne l'importance de définir des objectifs quantifiables ainsi que des périodes et des outils pour suivre la réalisation de ces objectifs
- 2010 – La Commission des Statistiques des Nations Unies demande l'actualisation du Cadre de Développement des Statistiques Environnementales
=> Définition d'un ensemble de base de statistiques environnementales
- Juin 2012 (Rio de Janeiro) – Sommet des Nations Unies sur le développement durable (« Rio+20 »)

=> Souligne l'importance de l'économie verte pour le développement durable et l'éradication de la pauvreté

1.2. Quelques définitions

Définition 1 : Environnement

L'environnement au sens large est défini comme :

- L'ensemble des éléments (biotiques ou abiotiques) qui entourent un individu ou une espèce et dont certains contribuent directement à subvenir à ses besoins »
- L'ensemble des conditions naturelles (physiques, chimiques, biologiques) et culturelles (sociologiques) susceptibles d'agir sur les organismes vivants et les activités humaines »

Dans ce cours, nous nous restreindrons à l'**environnement naturel**, c'est-à-dire à l'ensemble des éléments physiques, chimiques ou biologiques interagissant avec les organismes vivants.

Définition 2 : Statistiques environnementales

- Les statistiques environnementales fournissent une description quantitative de l'état de l'environnement, de ses changements au cours du temps et de ses interactions avec les activités humaines et les phénomènes naturels.
- Les statistiques environnementales mesurent les activités humaines et les phénomènes naturels qui affectent l'environnement, en quantifient les impacts et permettent de suivre les réponses apportées par la société.
- Les statistiques de l'environnement décrivent l'état et les tendances de l'environnement, couvrant les milieux de l'environnement naturel (air/climat, eau, terres/sols), l'ensemble des organismes vivants au sein de ces milieux et les établissements humains (Nations Unies, 1997).

Définition 3 : Bien public global

En économie, un bien public est qui possède les propriétés suivantes :

- Non-rivalité : la consommation de ce bien par une personne ne réduit pas la quantité disponible pour d'autres personnes ; et
- Non-exclusion : une fois produit, ce bien est accessible et peut être consommé par tous

On parle de bien public global (ou mondial) lorsque le bien en question ne fait pas référence à une délimitation géographique restreinte mais est au contraire disponible au niveau global. C'est le cas de beaucoup de ressources environnementales comme l'air (et sa qualité), la biodiversité, le climat, etc. Cette notion recouvre donc une importance particulière dans le domaine de l'environnement et a des implications sur les méthodes de collecte des données, de construction des statistiques et indicateurs et de partage des responsabilités aux différents niveaux.

1.3. Le champ : qu'est-ce qu'il faut mesurer ?

Il s'agit de mesurer :

- L'état de l'environnement ;
- Les changements de cet état entre deux périodes ; et
- Les interactions avec les organismes vivants, notamment avec les être humains. Cela inclut :
 - L'impact des activités humaines sur l'environnement ; et
 - Les réponses, d'ordre économique, social, technique, etc. apportées afin de répondre à un problème environnemental

En termes de thématiques :

- Se restreindre à l'environnement naturel ;
- Se restreindre à ce qui est mesurable ;
- Se référer aux composantes et sous-composantes du CDSE

1.4. Pourquoi et pour qui produire des statistiques sur l'environnement ?

Pourquoi ?

- Les enjeux nationaux

La dégradation de l'environnement et ses impacts sur les populations humaines doivent être mesurés : pour disposer d'informations pertinentes dans une perspective décisionnelle (à tous les niveaux – individuels et collectifs) et pour suivre l'effet de ces décisions sur les phénomènes environnementaux.

- Les enjeux internationaux

De nombreux phénomènes environnementaux ont une portée globale et ne s'arrêtent pas aux frontières administratives : une pollution à un endroit donné peut générer un impact à un autre

endroit, les effets de la dégradation d'un écosystème spécifique peuvent aller au-delà de ce même écosystème, etc.

Le caractère de bien public global de nombre de phénomènes environnementaux justifie l'action au niveau international, en particulier en matière d'harmonisation des méthodes de collecte des données afin de permettre agrégations et comparaisons dans le temps et l'espace.

L'articulation des différents niveaux d'actions - locaux, nationaux et internationaux – est essentielle afin de garantir une division adéquate du travail.

- Les engagements internationaux

Plusieurs conventions internationales existent (sur les changements climatiques, la biodiversité, la désertification, etc.). Dans le cadre de ces conventions, les pays peuvent s'engager à réaliser certaines actions (comme par exemple un engagement de réduction des émissions de gaz à effet de serre), à communiquer des informations sur l'état de l'environnement, le suivi des engagements, etc. Les pays membres de ces conventions sont donc contraints, directement ou indirectement, à produire un ensemble d'informations statistiques.

Par ailleurs, en tant que membres d'organisations internationales (ONU, FAO, etc.), les pays s'engagent à répondre aux requêtes de ces dernières en information statistiques. Ils doivent notamment répondre aux différents questionnaires (sur l'utilisation des sols, les polluants atmosphériques, etc.) qui leur sont transmis.

Pour qui ?

- La société civile et les ONG ;
- Les preneurs de décisions à tous niveaux ;
- Les chercheurs et scientifiques ;
- Les analystes de politiques, experts et conseillers ;
- Les organisations internationales.

Il s'agit-là des mêmes utilisateurs que pour d'autres thématiques statistiques. En matière de statistiques environnementales, les utilisateurs au niveau local, qu'ils proviennent de la société civile ou d'autorités locales, ont une importance majeure compte-tenu de la portée locale de beaucoup de phénomènes environnementaux (pollution des eaux, de l'air et des sols, érosion, etc.).

Conclusions :

- La demande en statistiques environnementales est croissante à tous les niveaux (local, national, régional et international) ;
- Les utilisateurs et usages sont variés et les niveaux d'analyse multiples : cela appelle à une stratégie de collecte et de diffusion des données adaptée ;

- Le caractère transversal des statistiques environnementales constitue un des défis qui doit être surmonté, sur le plan technique, opérationnel et de la gouvernance ;
- La transversalité et l'ampleur des questions environnementales fait de l'identification des principales thématiques et sous-thématiques un préalable à la mise en place de tout système structuré de production de statistiques environnementales.

=> C'est ce à quoi s'est attaché à répondre le CDSE.

2. Thématiques environnementales

2.1. Le Cadre de Développement des Statistiques de l'Environnement (CDSE)

- Le CDSE, présenté pour la première fois en 1984, est le résultat d'un travail coordonné par la Division des statistiques des Nations-Unies impliquant plusieurs organisations internationales et un groupe d'experts. Une nouvelle version du CDSE a été présentée en 2013, comportant notamment la définition de l'ensemble des statistiques environnementales de base. Le CDSE a été validé par la Commission des statistiques des Nations-Unies, la plus haute instance du système statistique mondial.
- Le CDSE propose une organisation des statistiques de l'environnement claire et flexible en composantes, sous-composantes, sujets statistiques et indicateurs ;
- Son point de départ sont les 6 thématiques ou composantes de base qui englobent les caractéristiques de l'environnement naturel et ses interactions avec le système humain ;
- Le CDSE fournit une hiérarchie des indicateurs en fonction de leur niveau d'importance, identifiant ainsi un ensemble de base des statistiques de l'environnement sur lequel les pays peuvent concentrer leur efforts ;
- Le CDSE est un outil qui vise à aider les organisations nationales en charge de la production de statistiques de l'environnement à identifier les thématiques à couvrir, à planifier et prioriser leur programme de collecte et à présenter les données de manière structurée, synthétique, claire et cohérente.

2.2. Thématique 1 : Condition et qualité de l'environnement

2.3. Thématique 2 : Les ressources environnementales et leur utilisation

2.4. Thématique 3 : Résidus

2.5. Thématique 4 : Phénomènes extrêmes et catastrophes

2.6. Thématique 5 : Les établissements humains et santé environnementale

2.7. Thématique 6 : Protection, gestion et engagement en matière d'environnement

Conclusion :

- Malgré la complexité des phénomènes environnementaux, le CDSE fournit un cadre clair et flexible permettant d'identifier les différentes thématiques et sous-thématiques ;
- La variété et la nature des différentes thématiques fait des statistiques environnementales un domaine par essence transversal. Cela a des implications sur :
 - Les méthodes de collecte des données
 - La présentation des indicateurs
 - Le dispositif institutionnel (le partage des responsabilités) permettant de produire les statistiques environnementales requises et de le diffuser (si possible de manière centralisée).

3. Collecte de données environnementales

3.1. Spécificité des données environnementales

Définitions

Définition 4 : Donnée environnementale

- Il s'agit d'une information généralement de nature quantitative mesurée sur ou collectée auprès d'une unité statistique (individu, ménage, entreprise, unité spatiale, milieu écologique bien défini, etc.) directement ou indirectement associée à une thématique environnementale.
- Par exemple, une donnée environnementale peut s'agir du nombre d'individus d'une certaine espèce compté sur une période et pour une parcelle donnée. Dans cet exemple, cette donnée est directement associée à un objectif environnemental (de mesure de la diversité biologique). Mais il peut également s'agir de données dont la portée environnementale n'est pas évidente et dépend de l'usage qui en est fait ou de la réalisation de transformations appropriées. Par exemple, la consommation de charbon de bois d'un ménage est une donnée qui peut être utilisée à des fins d'analyse environnementale (mesure des émissions de gaz à effet de serre, évaluation de l'impact de la consommation de charbon de bois sur les surfaces forestières, etc.) mais également à d'autres fins (aspects économiques, sanitaires, etc.).
- Les données environnementales ne sont généralement pas publiées en tant que telles, mais sous la forme de statistiques ou d'indicateurs, c'est-à-dire après transformation et/ou agrégation. Cependant, l'accès aux données brutes ou micro-données (c'est-à-dire la donnée collectée, sans aucune transformation ou agrégation) se généralise de plus en plus avec la

mise en place de protocoles d'accès standardisés au niveau international sous l'impulsion d'initiatives comme celle de Paris²¹.

- Dans ce cours, la notation y_i fera référence à une donnée environnementale, i pouvant représenter un individu (personne, ménage, entreprise), une unité temporelle, une unité spatiale (coordonnées GPS, zone, etc.) ou toute combinaison de ces 3 dimensions. Par exemple y_i peut faire référence à la teneur en particules fines mesurée quotidiennement par une station de surveillance géo-référencée.

Définition 5 : Statistique environnementale

- Par définition, une statistique est une transformation fonctionnelle des données: une moyenne, le maximum ou le minimum d'un ensemble de données, un ratio, etc. Elle implique donc un certain degré d'agrégation temporelle et/ou individuelle des données. Par exemple :
 - La transformation $y_S: (y_1, \dots, y_i, \dots, y_I) \rightarrow \sum_{i=1}^I y_i$ est une statistique. Il peut s'agir par exemple du nombre total d'espèces comptabilisées dans les différentes zones de recensement au cours d'une période donnée
 - La transformation $\bar{y}: (y_1, \dots, y_i, \dots, y_I) \rightarrow \frac{1}{I} \sum_{i=1}^I y_i$ est également une statistique. Il peut s'agir de la consommation moyenne d'électricité des ménages enquêtés.
- Il est important de ne pas confondre statistique environnementale et donnée environnementale

Spécificité

Les phénomènes environnementaux se distinguent par exemple des phénomènes économiques ou sociaux par :

- Le caractère diffus des sources de pollution et de leurs impacts, ne se limitant généralement pas à une unité économique ou institutionnelle, et souvent difficiles à circonscrire :
 - Même si dans certains cas les sources de pollution peuvent être uniques et bien définies (une usine) elles sont généralement nombreuses et mouvantes (les émissions de gaz à effet de serre provenant des véhicules, par exemple) ;
 - Les impacts sont souvent diffus (pollution de l'air, des eaux, etc.) et à échelle variable, de la plus petite (habitat spécifique mis en danger par la déforestation) à la plus grande (changements climatiques ayant des impacts globaux) ;
- La temporalité de sources de pollution et des impacts : d'un flux continu (pollution de l'air par les émissions de particules fines) à des impacts ponctuels (une marée noire générée par un accident maritime) ;
- La multiplicité et transversalité des thématiques environnementales : phénomènes biologiques, chimiques ou physiques d'une part et aspects économiques, sociaux,

urbanistiques, géographiques, etc. d'autre part, y compris les relations de ces composantes entre elles. Par exemple, l'impact de l'occupation des terres sur la déforestation, la réduction de la biodiversité, les impacts sur les écosystèmes forestiers, les émissions de gaz à effet de serre, etc.

Il en résulte une extrême diversité des données environnementales, de leurs sources et des méthodes appropriées afin de les collecter, de les compiler en statistiques et indicateurs et de les diffuser.

3.2. Les méthodes dites traditionnelles : enquêtes, recensements et registres administratifs

Les méthodes de collecte et de recueil des données dépendent du phénomène étudié.

Enquêtes et recensements

- Ces méthodes sont particulièrement adaptées à l'identification et à la mesure de certaines pressions environnementales, c'est-à-dire aux sources de pollution ou de dégradation de l'environnement, lorsque ces sources sont bien identifiées et peuvent être traduites en unités statistiques à même d'être enquêtées.

C'est notamment le cas des pressions environnementales exercées par des unités économiques ou institutionnelles bien identifiées, comme les ménages, les entreprises, les sites industriels. Par exemple, l'utilisation d'enquêtes est pertinente pour mesurer :

- La consommation d'énergie des ménages et des entreprises, à partir desquelles il est possible d'estimer les émissions de gaz à effet de serre associées ;
 - Les quantités d'engrais chimiques utilisés par les agriculteurs, qui permettront notamment d'évaluer leur impact sur l'eutrophisation des sols et des cours d'eau.
- Des enquêtes statistiques sont également réalisées pour mesurer certaines caractéristiques environnementales d'un milieu, d'une ressource ou d'une espèce. Un objectif fréquent de ces enquêtes est de caractériser l'état d'une ressource (d'un milieu, d'une espèce, etc.) et le changement de cet état.

La réalisation d'enquêtes environnementales basées sur une approche probabiliste se heurte à des défis spécifiques, notamment :

- La nécessité de définir une population d'intérêt suffisamment ample ;
- La nécessité de disposer et d'utiliser les informations auxiliaires ;
- La dépendance spatiale des données ;
- La difficulté de disposer de bases de sondage adaptées ;

- Les difficultés d'accès aux unités de la population ;
- L'existence d'objectifs multiples et la présence d'une variété de parties prenantes.

La construction de bases de sondage adaptées est un pré-requis à la réalisation d'enquêtes environnementales. Le type de base de sondage est fonction d'un ensemble de variables : la nature de l'unité statistique (discrète ou continue, par exemple), le budget, etc. :

- Les bases de sondage de listes : elles peuvent être construites/utilisées lorsque l'unité statistique est discrète et bien défini (un lac, un milieu spécifique, un cours d'eau, etc.). Ce type de base est coûteux à construire et doit être actualisé régulièrement afin de garantir sa complétude et son exactitude : certains milieux peuvent changer, se déplacer, disparaître...
- Compte-tenu de la spatialité des phénomènes environnementaux, les bases de sondage les plus souvent utilisées sont les bases aréolaires. Elles peuvent par ailleurs être complétées par des données sur la couverture et l'utilisation de sols afin de mieux structurer la population d'intérêt. Cependant, il est nécessaire de disposer d'images/cartes avec un niveau de résolution suffisant afin de pouvoir identifier les différents éléments de la population, ce qui peut s'avérer coûteux. Par exemple, certains éléments peuvent être identifiés à tort comme des éléments de la population d'intérêt : un lac peut être confondu avec une mare ou une zone inondée. Les unités identifiées peuvent également contenir des éléments autres que ceux qui devraient être inclus dans la population d'intérêt.
- Il est également possible de procéder à un maillage d'un territoire de manière homogène en utilisant des formes géométriques. Celles-ci sont généralement régulières (par exemple des carrés), de surface identique et prédéfinie (par exemple des carrés de 10km de côté). Ce maillage permet de définir une base de sondage à partir de laquelle un échantillon des différentes unités (mailles) sera sélectionné. Il est important de pouvoir combiner ce maillage avec des informations auxiliaires sur la topographie, la zone écologique, l'utilisation des sols etc. afin de stratifier la population d'intérêt et de procéder à un tirage de l'échantillon qui réponde aux objectifs de l'enquête.
- Compte-tenu du caractère spatial des données environnementales, il est au minimum nécessaire de pouvoir localiser les éléments de la population dans un plan (latitude, longitude). Idéalement, il faudrait également pouvoir l'associer à d'autres caractéristiques géographiques, agro-environnementales, topographiques, etc. afin de pouvoir utiliser ces informations auxiliaires pour le tirage de l'échantillon.

L'utilisation de méthodes d'échantillonnage est nécessaire pour sélectionner les éléments de la population: par exemple, il peut s'agir de la sélection d'un échantillon de zones géographiques (auquel cas, une base de sondage aréolaire sera nécessaire) à l'intérieur desquelles il sera procédé à différents comptages (d'individus d'une espèce donnée par exemple) ou de mesures (mesures biochimiques sur les sols, les cours d'eau, etc.). Les principales méthodes d'échantillonnage utilisées en matière d'enquêtes environnementales sont :

- L'échantillonnage en plusieurs étapes, ou par grappes : il consiste à partitionner un ensemble géographique en différentes zones (unités primaires) et sous-zones (unités secondaires) et de procéder à des tirages aléatoires successifs afin de sélectionner un échantillon d'unités primaires, puis d'unités secondaires, et enfin d'unités finales (par exemple des segments) qui feront l'objet de la visite d'enquêteurs.
- Que l'échantillon soit tiré en une seule ou plusieurs étapes, il est essentiel que la population soit stratifiée de manière spatiale, en plus d'une stratification analytique (usage des sols, zone écologique, etc.). Les strates peuvent s'agir de formes géométriques régulières comme des carrés ou de polygones délimitant des zones écologiques ou administratives distinctes. La nature des strates dépend de l'objectif de l'étude.
- Un moyen de tenir compte de la dépendance spatiale des données est de faire en sorte que l'échantillon soit spatialement équilibré, c'est-à-dire que le centre de gravité (densité spatiale) de l'échantillon corresponde à celui de la population.
- La méthode dite du sondage par partitionnement aléatoire permet d'obtenir un échantillon satisfaisant la propriété d'équilibre spatial. Cette méthode peut être utilisée pour des unités statistiques correspondant à des points (lacs), des flux linéaires (cours d'eau) ou des surfaces (zones forestières). Elle consiste d'abord à réaliser un partitionnement successif de la zone d'étude, puis à associer un identifiant à chacune des zones ainsi délimitées et les ordonner selon cet identifiant, mais selon un ordre hiérarchique inverse. Les zones sont ensuite sélectionnées par sondage systématique, stratifié ou non.
- La méthode de sondage adaptative est appropriée lorsque la population d'intérêt est de faible taille et lorsque les unités qui la composent ont tendance à se regrouper (forte corrélation spatiale), comme cela peut être le cas par exemple pour des espèces végétales dans une zone donnée. La méthode consiste à sélectionner et à enquêter un échantillon initial d'unités. Les unités adjacentes seront ajoutées à l'échantillon si les résultats de l'enquête sur l'échantillon initial vérifient une règle préétablie (nombre d'individus recensés supérieur à un certain seuil, par exemple). Si cette méthode permet de tenir compte de la dépendance spatiale des données, elle rend difficile le contrôle de l'échantillon (taille, dispersion géographique des unités, etc.), ce qui pose des problèmes empiriques (calcul de la variance, précision) et opérationnels (coût d'enquête d'unités dispersées).

Registres administratifs

Souvent, des législations spécifiques encadrent les activités de secteurs particulièrement polluants. Exemples (en Europe) :

- La législation sur les quotas d'émission de gaz à effet de serre : les établissements soumis à cette législation sont identifiées dans un registre, qui fournit également certaines

informations permettant de les caractériser (niveau de pollution, chiffre d'affaire, niveau de production, etc.) ;

- La législation REACH, qui régit l'usage par les industries de substances chimiques néfastes pour la santé. Cette législation a permis de générer des données et informations sur l'utilisation de substances toxiques, notamment d'identifier dans quelle mesure et par qui elles sont utilisées ;
- La directive Seveso impose aux États membres de l'Union européenne d'identifier les sites industriels présentant des risques d'accidents majeurs, dits « sites SEVESO », et d'y maintenir un haut niveau de prévention.

Ces registres peuvent être utilisés directement comme base pour produire des informations à caractère environnemental : identification des principaux pollueurs, de la nature de cette pollution et de son ampleur.

Ces registres peuvent également être utilisés comme base de sondage, à partir de laquelle des enquêtes plus approfondies peuvent être réalisées : par exemple, des enquêtes sur les investissements en matière de protection environnementale.

3.3. Les méthodes spécifiques : systèmes de suivi, imagerie aérienne et satellitaire et études techniques et scientifiques

Compte-tenu de la spécificité et de la complexité des phénomènes environnementaux, la mise en œuvre de méthodes spécifiques est souvent nécessaire.

Stations de suivi, surveillance ou contrôle

Ces stations sont généralement utilisées pour :

- Collecter des données météorologiques, en particulier sur la pluviométrie, l'humidité et les températures.
- Procéder à des mesures sur la qualité de l'air. Ces stations sont souvent situées dans ou proches de zones densément peuplées (zones urbaines) et effectuent des mesures sur la concentration de l'air en particules fines, le niveau d'acidification de l'air (émissions de SO₂ et NO_x notamment), les émissions de polluants organiques persistants, pour ne citer que les principaux.

Les données provenant de ces stations sont souvent caractérisées par une forte variabilité spatiale : le niveau de pollution de l'air est par exemple fortement dépendant de la localisation de la station de surveillance : proche ou non d'une source de pollution (voie de communication, etc.), d'un espace vert, orientation au vent, etc.

Se pose donc la question de la représentativité de ces mesures pour une zone spatiale donnée : est-il possible d'agréger certaines mesures, en réalisant par exemple des moyennes pour l'ensemble des stations de surveillance de la zone ? Ou faut-il présenter les données uniquement en référence à la station de mesure de laquelle elles proviennent ? Quel est le bon niveau d'agrégation ?

Il est clair par exemple que les données climatiques sont souvent « moyennées », parfois même au niveau régional ou national, même si cela n'a pas toujours de sens d'un point de vue scientifique et/ou analytique.

Exemple : Airparif

- C'est un organisme qui réalise une surveillance de la qualité de l'air en région parisienne. Son dispositif de mesure est constitué de 70 stations de surveillance. Elles sont réparties sur un rayon de 100 km autour de Paris et elles mesurent la qualité de l'air respiré par la population (plus de 11 millions d'habitants dans toute la région).
- La localisation des stations de mesure répond aux préoccupations en matière de santé publique : certaines sont situées proche des voies de circulation et d'autres en sont éloignées.
- De la mesure à l'information :
 - La concentration de l'air en polluants est mesurée de manière automatique et en continu par les stations de surveillance ;
 - Ces données sont transmises à la base de données centrale sous la forme de moyenne par quart d'heure ;
 - A partir de ces « données quart d'heure », des moyennes horaires et journalières des différentes concentrations sont ensuite effectuées. Ce sont les données de base environnementales dont nous parlions plus haut (ou les micro-données, selon la terminologie propre aux enquêtes) ;
 - Enfin, ces données sont utilisées par les ingénieurs d'AIRPARIF pour calculer un indicateur journalier de la qualité de l'air, se référant à une zone géographique précise (par exemple, l'agglomération parisienne). Il s'agit d'un indice composite car il résume différentes dimensions de la qualité de l'air (acidification, particules fines, etc.).
- Ces données et indicateurs sont utilisés pour construire et ajuster des modèles permettant notamment de prévoir la qualité de l'air en fonction de certaines variables clés, comme les conditions atmosphériques et le niveau de trafic routier.
- Ces données illustrent également les différents temps qui peuvent être identifiés dans le processus de construction des données :
 - La fréquence de collecte de l'information : dans cet exemple, il s'agit d'une collecte d'information en temps quasi-continu
 - La fréquence de transmission des données : il s'agit ici des moyennes sur un quart d'heure

- La fréquence relative aux micro-données ou données de base : il s'agit ici des moyennes horaires et journalières
- La fréquence et la nature des indicateurs/statistiques publiés : il s'agit ici de l'indicateur composite journalier mesurant la qualité de l'air en région parisienne

Imagerie aérienne et satellitaire

- Le caractère spatial de bon nombre de phénomènes environnementaux (aussi bien au niveau des pressions, des impacts associés sur les milieux que des réponses biophysiques et humaines) fait que les données/statistiques/indicateurs environnementaux se prêtent plus à une présentation sous forme visuelle et géographique que sous forme de séries statistiques, tableaux, etc.
- Les images aériennes sont fournies par des avions, hélicoptères ou drones.
- Les satellites enregistrent l'énergie réfléchie ou émise par la terre pour une large gamme de fréquences électromagnétiques. En fonction de la nature de ces ondes électromagnétiques, les satellites peuvent identifier certaines caractéristiques de la surface terrestre ou de son atmosphère : température, présence d'eau, densité et type de couvert végétal, zones bâties, etc.
- Les données satellitaires peuvent être organisées par capes : une cape regroupe un ensemble de caractéristiques relatives à la zone terrestre étudiée, par exemple le mode d'occupation et d'utilisation des terres ; une autre regroupe les informations sur les caractéristiques physiques des sols, la présence de zones humides, de couvert végétal, etc. ; une autre peut regrouper des informations relatives à la topographie et à l'hydrographie de la zone. Il est alors possible de « superposer » les différentes capes pour obtenir une information multidimensionnelle sur le territoire étudié. Ce type d'information est particulièrement utile pour structurer le territoire et réaliser des stratifications adaptées à la réalisation d'enquêtes ou d'analyses environnementales.
- L'imagerie aérienne et satellitaire permet :
 - d'identifier des unités statistiques au sein d'une population/zone géographique donnée. Elle peut donc être utilisée pour établir des bases de sondage aréolaires. Par exemple : l'ensemble des parcelles agricoles d'une zone donnée, l'ensemble des lacs, zones humides, zones forestières, etc.
 - De fournir directement les données de base pour le calcul d'indicateurs environnementaux, qui seraient difficile/couteux de collecter par des méthodes traditionnelles. Elles permettent par exemple d'évaluer :
 - Les variations de surfaces forestières dans une zone donnée entre deux périodes, et donc l'ampleur de la déforestation, etc.
 - L'ampleur de l'érosion marine, etc.

- L'étendue des zones urbaines
 - L'étendue, la densité et le volume du couvert végétal, variable souvent utilisée dans la prévision des récoltes.
- Lorsque le niveau de résolution le permet et le coût n'est pas prohibitif, l'imagerie aérienne et satellitaire peut permettre de caractériser de manière plus détaillée les unités statistiques en fournissant par exemple des informations sur la densité du couvert végétal ou des informations volumétriques sur les cours d'eau/lacs, forêts, etc. qu'il serait difficile et coûteux de collecter avec des méthodes de terrain.
- Il est toujours important de combiner imagerie aérienne/satellitaire avec visite de terrain afin de contrôler l'exactitude des mesures / de l'information recueillie et d'apporter des compléments d'informations que seul une visite de terrain est à même de fournir.
- Souvent, il n'est pas possible de collecter des informations par voie aérienne et satellitaire pour l'ensemble des unités de la population. Les informations sont alors collectées pour un échantillon de points (voir paragraphe sur les méthodes d'échantillonnage pour des exemples de méthodes de sondage), puis les informations sur les points manquants sont inférées par interpolation spatiale et/ou modélisation, en tirant parti de l'information auxiliaire de nature spatiale disponible par ailleurs. Par exemple, on sait que la densité du couvert végétal dépend (entre autres) du niveau de pluviométrie. Les données pluviométriques étant disponible pour l'ensemble de la zone étudiée, il est alors possible d'estimer la densité de couvert végétal pour les points manquants en fonction de la pluviométrie observée (et d'autres variables telles que les coordonnées géographiques, l'étendue des zones urbaines, etc.).

Études techniques et scientifiques

- Les phénomènes environnementaux ne peuvent souvent pas être observés directement, même au moyen des techniques satellitaires. Par exemple, il n'est pas (encore) possible d'observer directement les caractéristiques biochimiques d'un milieu : des prélèvements *in situ* sont nécessaires et/ou une modélisation à partir de relations physico-chimiques connues. De même, il est difficile de mesurer directement l'ensemble des émissions de gaz à effet de serre générées par un territoire donné sur une période spécifique, notamment lorsque les sources d'émission sont diffuses (celles générées par le transport automobile par exemple). Il faut généralement d'abord mesurer les pressions (ex : consommation d'énergie), puis passer de ces pressions à une mesure des quantités émises au moyen de paramètres appropriés (coefficients d'émission).
- Les différents paramètres permettant de passer des pressions aux impacts proviennent généralement d'études scientifiques, parfois réalisées au niveau international. Par exemple, le GIEC dispose d'un référentiel de facteurs d'émissions de gaz à effet de serre par secteur et type de processus industriel, provenant d'études techniques et scientifiques.

- Les phénomènes environnementaux sont parfois tellement complexes et spécifiques à un espace ou une zone écologique donnée (qui peut être de très petite taille), que la réalisation d'enquêtes ou d'études de grande échelle s'avèrent impossible d'un point de vue technique et financier. Il existe cependant un nombre important d'études portant par exemple sur un ou un nombre réduit d'habitats, permettant de fournir des informations environnementales intéressantes (diversité biologique, état, etc.). Ces informations étant spécifiques à ce ou ces milieux, elles sont difficilement généralisables ou extrapolables à un territoire administratif (région, pays, etc.). Dès lors, compte-tenu de leur spécificité, il devient difficile de les inclure dans un système d'information environnemental. Elles peuvent néanmoins servir de base à des travaux de modélisation permettant de réaliser des inférences et des extrapolations.
- Une source potentielle de données peut également provenir des études d'impact environnementales (EIA). Beaucoup de pays (développés et en développement) disposent d'une législation imposant la réalisation d'EIA préalablement à la réalisation de projets importants, notamment dans le domaine des infrastructures. Le cahier des charges est variable suivant les législations, mais ces études comportent généralement l'étude des impacts du projet en terme d'émissions de gaz à effet de serre, impacts sur les écosystèmes, etc. Les données produites dans le cadre de ces études constituent une base de données dont l'utilité croît avec le nombre d'études réalisées : plus la couverture est importante en termes de zones géographiques, secteurs ou phénomènes environnementaux, plus les données gagnent en robustesse et peuvent être à même d'être incluses dans un système d'information environnemental.

3.4. En pratique : l'utilisation combinée de différentes méthodes

- La diversité des sources de données environnementales est le reflet de l'étendu des phénomènes environnementaux, de leur transversalité et de leur complexité.
- Il est rare qu'une seule source de données, qu'une seule méthode de recueil de l'information soit suffisante pour apporter une information complète sur un phénomène environnemental.
- En règle générale, différentes sources de données sont combinées pour produire des statistiques/indicateurs environnementaux.
- Exemples de sources utiles à la production de statistiques sur les émissions de gaz à effet de serre :
 - Une enquête sur la consommation d'énergie des ménages et des entreprises ;
 - Une revue de littérature scientifique permettant d'identifier les coefficients d'émission permettant de passer des consommations d'énergie aux émissions de gaz à effet de serre ;
 - Une revue de littérature scientifique permettant d'identifier les coefficients d'émission associés aux différents procédés industriels ;
 - Données administratives permettant de couvrir certains gros émetteurs.

- Les données sur les émissions de gaz à effet de serre sont utiles dans la perspective d'en mesurer l'impact sur les changements climatiques. Pour cela, il faut :
 - Des modèles biophysiques permettant d'estimer l'impact des émissions sur la concentration de l'atmosphère en gaz à effet de serre ;
 - Des modèles climatiques permettant d'identifier la gamme des changements climatiques et leur ampleur à partir du niveau de concentration de l'atmosphère en gaz à effet de serre ;

4. Indicateurs statistiques et systèmes d'information environnementaux

4.1. Construction d'indicateurs statistiques : éléments de méthode

Définition 6 : Indicateur environnemental

- Un indicateur fournit une information synthétique sur un phénomène donné, est généralement disponible à intervalles réguliers et répond à des critères de cohérence ;
- Il permet de suivre dans le temps un phénomène donné, comme la concentration en gaz à effet de serre de l'atmosphère, les émissions de gaz à effet de serre annuelles d'un pays, etc. ;
- Un indicateur est construit à partir d'une statistique, généralement au moyen de transformations supplémentaires. Par exemple :
 - En reprenant le premier exemple de la définition 5, $I_t = 100 \cdot y_{S,t} / y_{S,t_0}$ est un indicateur mesurant l'évolution du nombre total d'espèces comptabilisées dans les différentes zones de recensement par rapport à la période de base t_0 ;
 - En reprenant le second exemple, $\hat{Y}_t = \frac{1}{N(t)} \sum_{i=1}^I \pi_{i,t}^{-1} y_{i,t}$ est un indicateur mesurant la consommation annuelle (t) moyenne d'électricité des ménages d'un pays. Les $\pi_{i,t}$ représentent les poids de sondage associés à chaque ménage enquêté.
- On parle d'indicateur synthétique ou composite lorsqu'un indicateur est construit à partir d'autres indicateurs, fournissant alors une information encore plus agrégée et synthétique. Un exemple d'indicateur synthétique est l'Indicateur de Développement Humain (IDH), qui est une moyenne pondérée d'un ensemble d'indicateurs tels que la richesse économique d'un pays, son niveau d'éducation, etc.
 - Un indicateur composite fournit une information plus synthétique. La contrepartie est que l'information est davantage diluée, ce qui peut parfois limiter son interprétation et sa portée analytique et/ou décisionnelle.

- La difficulté principale réside souvent dans la détermination d'un système de pondération adéquat. En l'absence de pondérations objectives (provenant de phénomènes quantifiés directement associé aux indicateurs – comme par exemple le nombre de ménages, le nombre de milieux écologiques d'un certain type, le pouvoir de réchauffement associé aux différents gaz à effet de serre, etc.), des pondérations subjectives sont souvent utilisées. Celles-ci posent question car elles reflètent des opinions/choix/orientations et influent donc de manière déterminante sur l'indicateur composite. C'est le cas de l'IDH : quelle pondération affecter aux indicateurs reflétant la dimension éducation par rapport à ceux reflétant la dimension économique ? Comment hiérarchiser ces différentes dimensions, juger de l'importance relative de chacune ? L'IDH est une moyenne géométrique simple de sous-indices (i.e. chaque sous-indice a la même pondération), ce qui signifie implicitement qu'on donne à chacune des dimensions de l'indice (richesse, éducation, etc.) la même importance.

Propriétés et caractéristiques souhaitables des indicateurs

Elles sont regroupées sous l'acronyme SMART pour :

- **S**pécifique
- **M**esurable
- **A**cceptable
- **P**ertinent (**R**elevant en anglais)
- **T**emporellement défini

4.2. Diffusion et présentation des statistiques environnementales

- De par le caractère spatial des phénomènes environnementaux, les statistiques environnementales se prêtent en général bien à des représentations spatiales.
- Leur représentation par l'intermédiaire de cadres analytiques et comptables (DPSIR/SEEA) permet de relier les phénomènes environnementaux entre eux, d'identifier et de quantifier leurs interactions avec le système humain (notamment avec les activités économiques). Cela augmente la portée analytique des indicateurs et statistiques, notamment dans une optique de prise de décision (par exemple, quel est l'impact de telle mesure sectorielle sur tel impact environnemental ?).
- Les phénomènes environnementaux qui se prêtent le plus à une présentation comptable sont ceux qui font l'objet de flux physiques, parfois valorisables (utilisation/extraction de ressources naturelles, par exemple), comportant des impacts en aval sur différents milieux écologiques (air, eau, etc.).
- Les statistiques sur l'état d'un milieu ou d'une ressource (diversité biologique, composition chimique d'un milieu, etc.) se prêtent mal à une présentation comptable. Elles se prêtent très bien en revanche à une présentation spatiale.

4.3. Le Système de comptabilité environnementale

Définition 7 : Le système de comptabilité environnemental (SEEA – acronymes en anglais)

Il consiste en un ensemble de comptes décrivant les relations entre l'environnement et l'économie, selon le même cadre conceptuel que celui utilisé pour la comptabilité économique traditionnelle (Eurostat, 2015).

Principales caractéristiques du SEEA

- Ce système permet notamment de comptabiliser les prélèvements effectués sur l'environnement, leur utilisation par le système économique et les rejets générés par celui-ci et versés dans l'environnement. Il comptabilise donc les rétroactions entre le système économique et environnemental, ce dernier étant à la fois une ressource pour le premier et un réceptacle pour ses rejets.
- Ce système est adapté aux phénomènes pouvant être décrits par des flux (extraction/utilisation, résidus/émissions, etc.), contrairement aux dimensions environnementales caractérisant l'état d'un milieu écologique et qui reflètent l'impact de ces flux sur le milieu.
- Dans la mesure où elle explique d'où proviennent et comment sont utilisées les ressources environnementales, la comptabilité environnementale s'intéresse également au stock de ces ressources.
- Ces comptes peuvent être présentés en flux physiques (l'unité naturelle en ce qui concerne l'environnement) et en flux monétaires (l'unité des comptes économiques). La valorisation économique des flux physiques permet de faire le lien entre la sphère économique et la sphère physique.

Analyses permises par le SEEA

Le lien comptable établi entre la sphère environnementale et économique, via l'utilisation des ressources naturelles et la quantification des rejets du système économique dans l'environnement, permet la réalisation de multiples analyses, notamment :

- Productivité et efficacité des ressources : quantité de matière utilisée par le système économique pour produire une unité monétaire de valeur ajouté.
- Évaluation de l'empreinte environnementale des différentes composantes de la demande (consommation des ménages, etc.) au moyen d'analyses input-output.
- Évaluation de l'élasticité de certains impacts environnementaux (émissions de GES, etc.) à une variation d'une des composantes de la demande.

- Analyses d'impact de politiques publiques : impact d'une augmentation de l'efficacité énergétique dans un secteur économique donné sur l'utilisation de matière et les rejets dans l'environnement.

4.4. Le cadre analytique DPSIR

- Le cadre DPSIR est un cadre analytique qui place les phénomènes environnementaux au sein d'une chaîne de causalité, partant des pressions exercées sur l'environnement par des facteurs structurels aux réponses apportées pour lutter contre les effets néfastes de l'activité humaine sur l'environnement.
- DPSIR :
 - **Driving forces** -> Forces motrices
 - **Pressures** -> Pressions
 - **State** -> État
 - **Impacts** -> Impacts
 - **Responses** -> Réponses
- Exemple : les changements climatiques
 - Forces motrices : démographie, urbanisation, hausse de la consommation, modification des régimes alimentaires dans les pays en développement (hausse de la part des protéines animales) etc. Ces forces sont nombreuses car elles regroupent toutes celles qui impactent directement ou indirectement la consommation d'énergie des ménages et des entreprises (principale source d'émission de CO₂), les émissions de CO₂ générées par la déforestation, celles de méthane (CH₄) générées par l'élevage. Il faudrait également y ajouter les facteurs affectant les procédés industriels à l'origine d'émissions directes de GES, etc.
 - Pressions : elles concernent les émissions de GES dans l'atmosphère, qui influent sur la concentration de GES (en ppm) de l'atmosphère, variable déterminante pour les changements climatiques (notamment le réchauffement).
 - État : on s'intéresse ici aux changements climatiques et donc à l'état du climat au sens large, c'est-à-dire en particulier : aux températures moyennes (réchauffement climatique), à la pluviométrie, au nombre et à l'intensité des événements climatiques extrêmes et à la variabilité climatique (amplitude des températures, de la pluviométrie, etc. d'une année à l'autre). A ce stade, il est important de définir l'échelle temporelle (plutôt long-terme pour les changements climatiques) et spatiale (région, pays, continent, etc.) sur laquelle porte l'analyse de l'état.
 - Impacts : il s'agit ici de déterminer et mesurer, dans la mesure du possible, l'impact des changements climatiques sur le bien-être humain. Celui-ci ne se résume pas à sa dimension économique, mais inclut également les dimensions éthiques (par exemple, solidarité intergénérationnelle) ou sociales (inégalités). Des exemples

d'impacts sont : les coûts économiques des changements climatiques, via notamment l'impact des aléas climatiques sur les infrastructures de transport, sur les établissements humains, etc. ; l'impact de la hausse de la variabilité climatique sur les rendements agricoles et donc sur les revenus des agriculteurs et la sécurité alimentaires ; l'impact de la pluviométrie sur les maladies hydriques ; etc.

- Réponses : il s'agit de l'ensemble des réponses des systèmes humains afin de lutter contre les impacts des changements climatiques, mais aussi afin d'améliorer directement l'état de certains milieux et de réduire les pressions. On exclut donc les réponses spontanées ou naturelles des milieux écologiques (adaptation spontanée), qui s'adaptent continuellement à leur environnement et au climat (migration d'espèces, mutations génétiques, etc.). Des exemples de réponses :
 - Réduction des impacts : meilleure intégration du risque climat dans la construction d'infrastructures de transport, révision des plans locaux d'urbanisme afin de tenir compte de la montée du niveau des eaux et du nombre et de l'intensité croissante des phénomènes d'inondation, etc.
 - Amélioration de l'état de certains milieux écologiques : augmentation des surfaces forestières afin de réduire les émissions de gaz carbonique (les forêts constituent un puits de carbone), restauration/création de zones naturelles (zones humides, buissons, forêts) permettant de protéger les établissements humains contre les inondations, glissements de terrain, l'érosion côtière, etc.
 - Réduction des pressions : mise en place de stratégies visant à réduire la consommation d'énergie (incitations fiscales, politique d'efficacité énergétique, transports publics, etc.), à améliorer les procédés industriels, à réorienter la consommation vers des produits moins intensifs en carbone, etc.
- Il est important, lorsque l'on réalise une analyse DPSIR, de bien cadrer le phénomène à analyser. Par exemple, il faudrait décomposer la thématique des changements climatiques en ces différentes composantes – température, variabilité climatique, phénomènes extrêmes – car ceux-ci n'ont pas forcément les mêmes impacts et ne font pas forcément référence aux mêmes pressions.

Jusqu'ici nous nous sommes plutôt concentrés sur les aspects techniques des statistiques environnementales. Cependant, à la complexité de certains phénomènes environnementaux viennent s'ajouter des défis de nature institutionnelle et organisationnelle qu'il est nécessaire de surmonter dans l'optique de la mise en place d'un système d'information sur l'environnement durable et de qualité.

5. Le partage des responsabilités et la mise en place d'un cadre institutionnel adapté

Cadre institutionnel

- La nature transversale des phénomènes environnementaux, leur complexité et la relative nouveauté de la demande en statistiques environnementales constituent des défis pour la mise en place d'un système de production statistique viable et efficace.
- Plusieurs organismes publics sont à même de produire (ou produisent déjà) des statistiques environnementales au sens large, c'est-à-dire en incluant les données de base sur les pressions en plus des indicateurs sur l'état et les impacts. On retrouve notamment :
 - Le ministère en charge de l'environnement
 - Le ministère en charge des eaux
 - Le ministère en charge de l'assainissement
 - Le ministère en charge des forêts et des espaces naturels
 - L'institut national de statistiques
 - Le ministère en charge de l'énergie, du pétrole et des mines
 - Agences nationales et régionales de l'environnement
 - Les agences et organismes internationaux
- Compte-tenu de la multiplicité des phénomènes et de leur forte spécificité géographique, les organismes de la société civile sont également à même de produire des données sur l'environnement, notamment :
 - Les associations de protection de l'environnement, de la faune et de la flore.
 - Les organisations non-gouvernementales nationales ou internationales.
- Le fait que la production de statistiques sur l'environnement soit décentralisée est naturel compte-tenu des caractéristiques des phénomènes environnementaux. Il n'est pas souhaitable que cela soit remis en cause. En revanche, il est important qu'il existe une entité coordinatrice qui soit garante de la qualité des méthodes statistiques utilisées, de la robustesse des résultats et de leur diffusion de manière coordonnée. Il est également nécessaire que cette entité puisse favoriser un partage des responsabilités approprié, basé sur les compétences respectives de chacun des acteurs, puisse délimiter le champ des statistiques environnementales et assurer la diffusion des statistiques de manière coordonnée.
- Il est également important que l'organisme en charge de la coordination des statistiques de l'environnement puisse créer les conditions d'un dialogue constant entre la communauté scientifique (les experts dans les différentes thématiques environnementales), les représentants de la société civile et les producteurs de données. Cela en vue de garantir la pertinence des données collectées et des méthodes utilisées au regard des besoins des utilisateurs.

La question de la formation

- La mesure des phénomènes environnementaux requiert des capacités techniques et analytiques qui vont au-delà des compétences classiques des statisticiens publics. Sans devoir pour autant être experts de son domaine, le statisticien doit notamment pouvoir comprendre les phénomènes physiques, chimiques et biologiques qui déterminent l'état

d'un milieu écologique. Par exemple, le statisticien doit pouvoir comprendre d'où proviennent les émissions de GES, par quels mécanismes la concentration de ces gaz dans l'atmosphère influe sur le climat et quels sont les principales caractéristiques du changement climatique.

- Les statistiques de l'environnement constituent une thématique relativement nouvelle, à laquelle par conséquent peu de personnes sont formées.
- Il est donc essentiel d'inclure la formation aux statistiques de l'environnement dans les cursus de formation initiale et continue des statisticiens publics. La formation continue est déterminante compte-tenu du peu de personnes actuellement formées et des constantes évolutions des techniques de collecte et de traitement de l'information. A ce titre, des thématiques qui devraient faire l'objet de formations sont notamment :
 - la collecte de données satellitaires ou aériennes et leur traitement par des logiciels appropriés ;
 - L'utilisation des systèmes d'information géographiques ;
 - Les statistiques et l'économétrie spatiale, qui permettent d'identifier et de quantifier les liaisons spatiales entre les données.

6. Un cas pratique : la construction de statistiques sur les émissions de gaz à effet de serre

Bibliographie indicative

Articles

- **Marker and Stevens Jr.** (2009), Sampling and Inference in Environmental surveys, in *Sample Surveys: Design, Methods and Applications, Vol 29A*, Chapter 19, pp. 487 - 512

Sites internet

- Eurostat / Statistiques de l'environnement : <http://ec.europa.eu/eurostat/web/environment>
- Division des statistiques des Nations Unies : <http://unstats.un.org/unsd/environment/>