

PROJET D'UNE ANALYSE STATISTIQUE  
DE DONNÉES PORTANT SUR LE RÉSEAU ROUTIER  
ET CONTENUES DANS DIVERS FICHIERS INFORMATISÉS

Jean David

Service de la statistique

Direction de la recherche

CANQ

TR

BSM

RE

132

Octobre 1984

## TABLE DES MATIÈRES

|                                                                | Page |
|----------------------------------------------------------------|------|
| Introduction .....                                             | 1    |
| 1. Définition du problème .....                                | 1    |
| 2. Données disponibles                                         |      |
| a) variables circonstancielles .....                           | 2    |
| b) autres variables .....                                      | 3    |
| 3. Analyse des données .....                                   | 3    |
| 4. Perspectives .....                                          | 4    |
| Conclusion .....                                               | 5    |
| Annexe 1: Description détaillée de l'analyse statistique ..... | 6    |

## INTRODUCTION

Différents fichiers informatisés du MTQ comprennent un grand nombre de données sur le réseau routier. Le traitement de ces données est plutôt limité et elles sont rarement synthétisées pour caractériser le réseau routier de façon globale.

Nous allons définir les paramètres de ce problème, déterminer précisément les données disponibles, voir de quelle manière on pourrait analyser puis synthétiser celles-ci et enfin, ouvrir de nouvelles perspectives. L'objectif premier de tout ce travail est de voir quelle sorte d'image du réseau routier on pourrait ainsi obtenir.

### 1. DÉFINITION DU PROBLÈME

Beaucoup de données sur le réseau routier sont actuellement disponibles dans les différents systèmes informatisés au ministère des Transports.

Ces données peuvent nous donner les caractéristiques physiques, l'utilisation et l'état des sous-sections ou sections de routes. On peut également obtenir certaines données relatives aux travaux effectués. Ces données sont, par leur nature, particulières en ce sens qu'elles se réfèrent à des entités plutôt petites. C'est pour cette raison que leur utilisation vise des objectifs dont la portée est assez restreinte.

Il faut donc trouver un moyen de rendre ces données plus disponibles, particulièrement à ceux qui désirent se faire une image plus globale du réseau routier. En effet, il faudrait pouvoir répondre à toutes sortes de questions relatives à l'ensemble du réseau routier, ou à tout le moins, à des entités plus vastes que la section ou la sous-section de route. Par exemple, on peut chercher à connaître les caractéristiques du réseau dans telle ou telle région, à déterminer dans quel état il est, quel est son utilisation, comment il se compare avec d'autres régions, ou encore avec l'ensemble du réseau, comment il évolue dans le temps... On devrait donc pouvoir analyser le réseau routier globalement, c'est-à-dire au niveau d'entités telles que le district, la région, la circonscription électorale ou encore la municipalité régionale de comté. Il faudrait pouvoir dégager les impacts qu'auraient sur le réseau routier un ensemble d'interventions particulières faites, évidemment, sur plusieurs sous-sections ou sections de route.

Beaucoup de données sont disponibles et le problème est de déterminer de quelle façon on obtiendra des statistiques qui nous permettront de répondre à des questions visant l'ensemble du réseau routier, ou à tout le moins, des entités plus larges que la section ou la sous-section de route.

## 2. DONNÉES DISPONIBLES

Il faut déterminer quelles sont les variables disponibles et susceptibles de répondre aux questions posées. Il faut toutefois se rendre compte qu'à ce niveau, tout n'est pas bien défini. Le plus souvent, le temps et l'usage permettent de dégager les variables les plus importantes. Dans un premier temps, il faut donc voir quelles sont les variables théoriquement disponibles. On peut les diviser en deux groupes: celles qui définissent les entités auxquelles se rapportent les données ou variables circonstanciellles, et les autres, celles qui définissent les caractéristiques du réseau.

### a) Variables circonstanciellles

La plus petite unité du réseau routier est la sous-section de route et dans plusieurs systèmes informatisés, elle est l'unité de base. Pour toutes les autres variables circonstanciellles, il faut faire des regroupements, ce qui implique un traitement des données de base. Voici une liste de variables circonstanciellles que nous retrouvons dans un ou plusieurs des systèmes:

- . route
- . tronçon
- . section
- . sous-section
- . région
- . district
- . municipalité
- . circonscription électorale provinciale (CEP)
- . municipalité régionale de comté (MRC)
- . milieu
- . classification fonctionnelle du réseau

Il faut tout d'abord remarquer que par le numéro de route, nous pouvons déterminer quelle est le type de la route, ce qui en quelque sorte nous donne une variable supplémentaire. D'autre part, la classification fonctionnelle du réseau n'est pas encore opérationnelle et pour cette raison les données sont toujours manquantes à ce niveau.

Sauf pour des études de nature très particulière, les regroupements au niveau de la section, du tronçon et de la route n'ont pas un très grand intérêt, ni une très grande signification. Les regroupements de plusieurs routes en fonction du type de route présentent un intérêt certain, puisque le type de route est intimement lié aux programmes du Ministère, principalement pour la construction.

D'autre part, les regroupements au niveau des districts et des régions présentent un intérêt beaucoup plus grand à cause, évidemment, de la structure du Ministère. C'est un objectif important que d'avoir une image du réseau routier et de voir son évolution au

niveau des régions.

Au niveau des municipalités, des CEP et des MRC, l'intérêt vise particulièrement des personnes qui sont extérieures au MTQ, mais qui sont également des intervenants à divers titres. Il y a donc là des intérêts plus particuliers qu'il y a lieu de satisfaire selon les besoins exprimés.

En ce qui regarde le milieu (urbain ou rural) et la classification fonctionnelle, tout comme pour le type de route, l'intérêt est beaucoup plus technique, se situant au niveau de la planification et de la programmation.

#### **b) Autres variables**

Les autres variables disponibles peuvent se diviser en trois groupes. Un premier groupe de cinq variables servent à définir des caractéristiques physiques de la route:

- . longueur de la sous-section
- . largeur de l'accotement de la route
- . largeur de l'emprise
- . nombre de voies
- . largeur de la voie

Le second groupe de variables nous éclairent sur l'utilisation de la route. Il s'agit du jour moyen annuel (jma), du pourcentage de camions et du nombre d'accidents. D'autre part, si nous combinons la longueur de la sous-section avec le jma, nous obtenons une variable qui définit particulièrement bien l'utilisation de la sous-section, soit le nombre de véhicules-kilomètres. Cette variable permettra de comparer et de regrouper les résultats de plusieurs sous-sections.

Enfin, un dernier groupe de variables, soit le coefficient de roulement, la déflexion, le drainage et le nombre de courbes sous-standards nous permettent de connaître l'état de la route.

Les variables, contenues dans les différents systèmes, peuvent donc être séparées en deux groupes: les variables circonstanciellles qui serviront d'unités pour faire différents regroupements, et les autres variables qui définissent le réseau sous divers aspects et qui seront analysées en fonction des premières.

### **3. ANALYSE DES DONNÉES**

L'analyse des données devra se faire en deux grandes étapes. La première étape consistera à déterminer le taux de données manquantes et le taux d'erreurs dans les données. La deuxième étape sera l'analyse statistique de ces données.

Dans un premier temps, il nous faut donc déterminer le pourcentage de données manquantes. Il s'agit là d'une étape assez simple et qui peut se faire rapidement. Ce travail doit être effectué pour chacune des variables. En second lieu, il faudra déterminer le taux d'erreurs pour chacune des variables. Pour ce faire, il n'est pas nécessaire de vérifier chacune des données. En effet, on peut se limiter à un échantillonnage du fichier. Dans le cas des variables circonstanciellles, il s'agit d'un travail assez simple. Pour les autres variables, le travail est plus complexe. Pour certaines de ces variables, on ne peut pas réellement faire de vérifications, puisque les mesures ne peuvent être reprises (déflexion, coefficient de roulement, drainage, etc.). Par contre, pour d'autres variables (les caractéristiques physiques, par exemple), on peut vérifier les données à l'aide de plusieurs fichiers. Ces analyses pourront orienter les futures cueillettes de données dans le sens de compléter ce qui manque et d'améliorer la précision des données là où c'est nécessaire.

La seconde étape sera de faire l'analyse proprement dite des données. Notre objectif principal sera d'analyser chacune des variables en fonction des variables circonstanciellles dans le but de caractériser le réseau routier de façon plus globale, c'est-à-dire au niveau d'ensembles plus larges que ceux de la route, du tronçon ou de la section. Ces analyses devront permettre de faire des comparaisons entre les différentes entités de même niveau, de suivre l'évolution du réseau au niveau de ces entités et ainsi d'être un outil de gestion, de programmation et de planification. Ces analyses seront d'ordre statistique: on calculera des moyennes, des écarts-types, des distributions de fréquences et même, on tentera de mettre en corrélation certaines variables. Tenant compte des taux d'erreurs et de données manquantes, on analysera les variables pour lesquelles on aura des résultats ayant un bon degré de précision. En annexe 1, on donne une description plus détaillée des statistiques qui seront calculées pour chacune des variables.

L'analyse des données se fait donc en deux étapes bien déterminées: calcul des taux d'erreurs et de données manquantes et analyse statistique des variables. L'objectif visé est de caractériser le réseau routier de façon plus globale, tant du point de vue de ses caractéristiques physiques que de son utilisation et de son état. L'analyse ainsi faite devra permettre de suivre l'évolution du réseau routier de façon globale.

#### **4. PERSPECTIVES**

Dans notre analyse du réseau routier jusqu'à maintenant, nous avons parlé de ses caractéristiques, de son utilisation et de son état. Chaque année cependant, beaucoup de travaux sont effectués sur le réseau et des sommes d'argent importantes y sont consacrées. Il faudrait donc également analyser le réseau routier sous ces deux aspects.

Actuellement, nous possédons des données sur la programmation des travaux de construction et d'entretien et certaines données sur les dépenses effectuées au cours des dernières années par le biais du système 0129. Nous pouvons donc avoir une idée des sommes dépensées par sous-section pour l'amélioration du réseau routier. Pour une image complète, il nous faudrait obtenir le total des coûts annuels d'entretien et de construction pour chaque sous-section et cela par type de travaux ou d'activités.

Toutes ces données ne sont pas encore disponibles. Cependant, il faudra faire un effort pour les obtenir, car elles représenteraient un formidable outil de gestion. En effet, on pourrait mettre en parallèle tous ces coûts et l'état du réseau, comparer les coûts entre les différentes variantes des variables circonstanciées, suivre l'évolution des coûts et établir une relation entre les coûts et toutes les caractéristiques du réseau routier.

Il ne fait donc nul doute que la variable "coût des travaux" est l'une des variables les plus intéressantes. Il faut cependant reconnaître que la véritable difficulté est précisément d'établir ces coûts. Il faudra cependant, à un moment donné, faire ce travail si l'on veut devenir véritablement efficace.

## CONCLUSION

L'analyse statistique des données existantes dans les systèmes informatisés pourra nous donner une connaissance globale des caractéristiques physiques, de l'utilisation et de l'état du réseau routier. Les données sont présentement disponibles au niveau des plus petites entités du réseau routier. Des variables circonstanciées pourront servir de base pour des regroupements utiles aux différents intervenants. L'analyse statistique de toutes ces données déterminera le taux de précision et l'image globale du réseau routier sous tous les aspects. Dans l'avenir, on pourra même y ajouter une nouvelle dimension: celle des coûts. A ce moment, tous ces systèmes informatisés auront donc atteint leur véritable objectif: servir à la prise de décision à tous les niveaux de gestion.

**ANNEXE 1**

**DESCRIPTION DÉTAILLÉE DE L'ANALYSE STATISTIQUE**

## Description détaillée de l'analyse statistique

Nous avons déjà déterminé que les variables peuvent se diviser en deux grands groupes: les variables circonstanciellles et les autres variables. Ces dernières se divisent elles-mêmes en trois sous-groupes selon leur objet: les caractéristiques physiques, l'utilisation et l'état du réseau. Ces variables sont principalement analysées selon des regroupements faits à partir des variables circonstanciellles. Notre description de l'analyse statistique se fera en trois parties selon l'objet des variables.

### 1. Caractéristiques physiques

Il s'agit de décrire le réseau routier pour chaque district et chaque région. Selon les demandes, on peut également le faire pour les MRC, les CEP et les municipalités. Il faut d'autre part, considérer que les caractéristiques physiques sont fortement dépendantes du type de route; alors, il faut décrire en fonction de cette variable et des variables circonstanciellles déjà nommées.

La première statistique à obtenir est la longueur du réseau routier par type de route. Il s'agit là d'une statistique de base. Il faut remarquer que cette statistique est déjà disponible. Toutefois, elle servira principalement à vérifier l'exactitude du fichier que nous allons constituer pour faire l'ensemble du travail. Cette statistique pourra également être compilée à la fois en fonction du type de routes et du nombre de voies. Il serait également intéressant de calculer une longueur moyenne des sous-sections pour des fins de comparaison entre les diverses valeurs des variables circonstanciellles (par exemple, entre les régions). Cette statistique est une aide lorsque l'on veut évaluer la valeur relative de d'autres statistiques: on veut savoir si les données sont vraiment comparables.

Pour les autres variables, soit la largeur de l'accotement, la largeur de l'emprise et la largeur de la voie, des moyennes et également des écarts-types seront calculés en fonction des types de route et des variables circonstanciellles. De plus, pour chacune de ces variables, on pourra créer des classes de largeurs significatives et établir des distributions de fréquences. Ces moyennes et ces distributions pourront être comparées et analysées de différentes manières. A plus long terme, on pourra déterminer s'il y a une certaine évolution au cours des ans.

Les statistiques ainsi obtenues nous fourniraient une description globale du réseau routier. Les comparaisons entre les diverses valeurs des variables circonstanciellles permettraient de constater les écarts pouvant exister et même d'apercevoir certaines déficiences si tel est le cas.

## 2. Utilisation du réseau

Nous avons trois variables principales qui décrivent l'utilisation du réseau routier. Nous avons le jour moyen annuel ou jma, le pourcentage de camions et le nombre d'accidents. En ce qui concerne les deux premières, il serait assez simple de calculer des moyennes et des écarts-type. Ces moyennes seraient des moyennes pondérées par le kilométrage et seraient évidemment calculées selon les types de route et en fonction des variables circonstanciées. Il serait également possible de faire des classes de jma et de pourcentage de camions et d'établir des distributions de fréquences. Avec les mises à jour de ces variables ou encore des actualisations de celles-ci, il serait alors possible de suivre leur évolution.

D'autre part, le nombre de véhicules-kilomètres qui est le produit du jma par le nombre de kilomètres et par le nombre de jours dans une année, est une variable qui définit assez précisément l'utilisation que l'on fait d'une sous-section de route. Cette variable est tout à fait comparable d'une sous-section à l'autre et par conséquent, elle permet aisément tous les genres de regroupements. La statistique voulue serait donc un nombre total de véhicules-kilomètres et peut-être un estimé valable du nombre de kilomètres parcourus dans une entité quelconque. Cette variable a un autre usage très important. En effet, elle permet de calculer un taux d'accidents par nombre de véhicules-kilomètres et rend le nombre d'accidents comparables entre les sous-sections. Elle permet également l'établissement d'un taux moyen d'accidents et les comparaisons entre diverses entités. Finalement, elle permet de suivre l'évolution du nombre d'accidents. Enfin, par voies de comparaison, on peut en arriver à déterminer les endroits les plus dangereux.

La variable "nombre de véhicules-kilomètres" permet d'établir assez bien l'utilisation du réseau routier et rend possible les comparaisons entre diverses entités. On peut également suivre l'évolution de l'utilisation du réseau en termes de jma, pourcentage de camions et nombre d'accidents.

## 3. État du réseau routier

Nous voulons déterminer l'état global du réseau routier. Les statistiques connues sont le coefficient de roulement, la déflexion, le drainage, le nombre de courbes sous-standards et de pentes critiques. Ces données sont disponibles par sous-section. Il s'agit ici de rendre ces résultats disponibles pour des regroupement par type de route, région, district, etc. En faisant une moyenne pondérée par la longueur des sous-sections, il serait possible d'obtenir un coefficient de roulement moyen et une déflexion moyenne pour des regroupements de sous-sections. Grâce à une formule d'actualisation, il est possible d'obtenir des résultats pour la plupart des sous-sections. Les résultats seraient ainsi plus significatifs et l'on pourrait également, grâce également à des mises à jour, suivre l'évolution de l'état du réseau.

Quant au drainage qui est défini par des cotes et le nombre de courbes sous-standards, la statistique la plus significative serait sans doute une distribution de fréquences. Cela permettrait d'obtenir le nombre de sous-sections par cote de drainage et par nombre de courbes sous-standards. Encore là, il serait possible de suivre l'évolution du réseau.

Il apparaît donc qu'il est possible de connaître l'état du réseau de façon globale et de voir comment il évolue.

Il ressort de cette description assez détaillée que certaines statistiques simples permettraient de connaître le réseau routier de façon globale et même de suivre l'évolution de celui-ci, autant dans ses caractéristiques physiques que dans son utilisation et son état.

MINISTÈRE DES TRANSPORTS



QTR A 093 780