UN MODELE POUR L'EVALUATION DU TEMPS PERDU PAR LES USAGERS A L'OCCASION DES TRAVAUX D'ENTRETIEN DES CHAUSSEES

F. BRILLET & Ph. LEPERT
Division Gestion de l'Entretien des Routes (GER), LCPC, Nantes, France
Francois.brillet@lcpc.fr philippe.lepert@lcpc.fr

Certains outils, comme le HDM par exemple, ont été développés afin d'optimiser les avantages économiques apportés par l'entretien des routes ; ils comprennent des modèles tenant compte des principaux termes économiques que sont les coûts pour l'usager (temps, sécurité, confort, consommation du véhicule), les coûts pour le maître d'ouvrage (construction, entretien et exploitation de route), et parfois des coûts externes comme les dommages à l'environnement ou la consommation de ressources non renouvelables. Cependant, les études récentes ont prouvé qu'une optimisation complète de l'entretien routier devait tenir compte de toutes les conséquences des travaux, y compris les coûts à l'usager au droit des chantiers, et qui constituent des facteurs importants de la rentabilité économique de l'entretien, particulièrement dans les pays européens où la circulation est souvent proche des niveaux de saturation. Réduire la fréquence des travaux routiers, leur durée et leur impact sur la circulation (ralentissements et bouchons) permettra de diminuer les pertes de temps subies par des usagers dans leurs déplacements, et réduira ainsi le coût économique total de la politique d'entretien. De telles décisions peuvent mener à une augmentation du coût direct des travaux, mais ceci peut être accepté si le bilan global est positif. Le problème principal dans l'évaluation des pertes de temps provoquées par les travaux est que la réponse des conditions de circulation est très sensible à de faibles changements des paramètres comme la demande du trafic et la capacité de route, même si ces changements sont très limités dans le temps. Afin de surmonter cette difficulté, un modèle probabiliste a été mis au point, permettant de prévoir le trafic instantané selon une distribution normale. Les paramètres du modèle peuvent être estimés par une analyse des comptages horaires de trafic réalisés pendant une série de jours qui peuvent être répartis en différentes catégories. Ce modèle sera employé pour calculer, sur réseau routier donné et pour une politique d'entretien donnée, le total des pertes de temps subies par l'usager du fait de l'entretien de la route, et ainsi leur coût et leur impact sur le bénéfice total attribué à la solution proposée.

MOTS-CLÉS : TRAFIC, ENTRETIEN, CHAUSSEE, USAGER DE LA ROUTE, TEMPS D'ATTENTE, MODELE PROBABILISTE

1 OBJECTIF

L'analyse coût bénéfice, appliquée depuis longtemps au choix des investissements routiers (tracés neufs ou aménagements), est encore peu utilisée pour optimiser les stratégies d'entretien des chaussées, du moins dans les pays industrialisés. Il est vrai que cette démarche implique la « monétarisation » de tous les coûts et avantages résultant des travaux envisagés, qu'ils soient directs ou indirects, immédiats ou futurs, subis par la puissance publique, l'usager ou la population en général…

Or, si les dépenses réelles (réalisation des travaux, consommation en carburant, coût matériel des accidents) peuvent être établies de manière objective, d'autres coûts, comme ceux des accidents corporels ou des effets sur l'environnement, reposent sur des hypothèses ou des présupposés qui ne font pas tous l'objet d'un consensus. En France, un travail important de formalisation et d'évaluation de ces coûts a été réalisé dans le cadre du Rapport Boiteux [1].

La nécessité de progresser et converger sur ce sujet au niveau européen a été à l'origine du contrat FORMAT (Fully Optimised Road Maintenance) [2]. Le LCPC y prend une part active, et pilote notamment le volet concernant l'évaluation des retards subis par l'usager du fait des travaux d'entretien.

Les paragraphes qui suivent décrivent une partie du travail réalisé en 2002 dans le cadre de FORMAT, et plus spécialement l'évaluation probabiliste des temps d'attente. Cela a consisté à déterminer, à tout instant d'une journée, la probabilité que le flot de trafic atteigne un niveau donné ; la comparaison entre cette valeur et la capacité résiduelle de la route en chantier permet alors de calculer le nombre probable d'heures perdues par les usagers, puis l'espérance mathématique de ce nombre.

2 ENJEUX

Afin de cerner l'ordre de grandeur du problème, prenons un exemple : sur une route à double chaussée, où la vitesse est limitée à 110 km/h et où l'on compte 20 000 véhicules par jour, un chantier de renouvellement de couche de surface doit avoir lieu sur 5 km pendant 10 jours ; cela nécessite la neutralisation d'une chaussée, avec mise à double sens de l'autre, et limitation de la vitesse à 70 km/h. Le coût direct des travaux est évalué à 30 €/m², soit 1,2 M€ pour l'ensemble.

Si la circulation reste fluide, le seul retard subi sera dû à la réduction de vitesse limite, soit 0,025 h (une minute trente) par véhicule, et donc 5 000 v.h (véhicules-heures) pour l'ensemble du chantier ; à raison de 13,7 € par v.h (valeur préconisée par le Rapport Boiteux pour les trajets à longue distance), cela donne 68 500 €, ce qui représente déjà 6 % du coût des travaux.

Dans des conditions plus difficiles (où la moitié des véhicules subirait une attente de 15 minutes, et franchirait le chantier à 30 km/h), le compte s'établit à 12 500 v.h pour l'attente et 19 000 v.h pour le franchissement, soit un total de 430 000 € selon le précédent barème, ce qui représente alors plus du tiers du coût direct. Pourtant ce calcul ignore les conséquences pour les poids lourds, la sécurité, les consommations et l'environnement.

Il existe bien sûr des cas encore plus graves, mais aussi des solutions palliatives (itinéraires de délestage, travail de nuit, information du public, incitations à utiliser d'autres modes de transport) qui elles mêmes peuvent être traitées dans une analyse coût bénéfice, mais tout cela dépasse notre propos…

3 MODELISATION DU TRAFIC

3.1 Les comptages disponibles

Les stations SIREDO (Système Informatisé de Recueil des Données) [3], réparties sur le réseau des routes nationales françaises, assurent en continu les comptages par tranche d'une heure. Sont en outre relevés sur certaines stations le nombre de poids lourds, les poids par essieux et les vitesses. Sur demande préalable, des exploitations telles que le comptage par période de six minutes peuvent être obtenues, mais aucune demande rétroactive n'est possible, les données unitaires n'étant pas archivées.

Les fichiers de comptages ont été mis à notre disposition par le CETE (Centre d'Études Techniques de l'Équipement) de l'Ouest. Quatre stations ont été choisies, correspondant à des niveaux et types de trafic différents :

1. Une route à deux voies en zone rurale, recevant un trafic moyen (N 171).
2. Une autre recevant un fort trafic interrégional, auquel s'ajoutent les relations entre deux grandes villes et la desserte des plages (E60).
3. Un itinéraire Paris province, avec forte incidence saisonnière (E50).
4. Le périphérique d'une métropole régionale (E03).

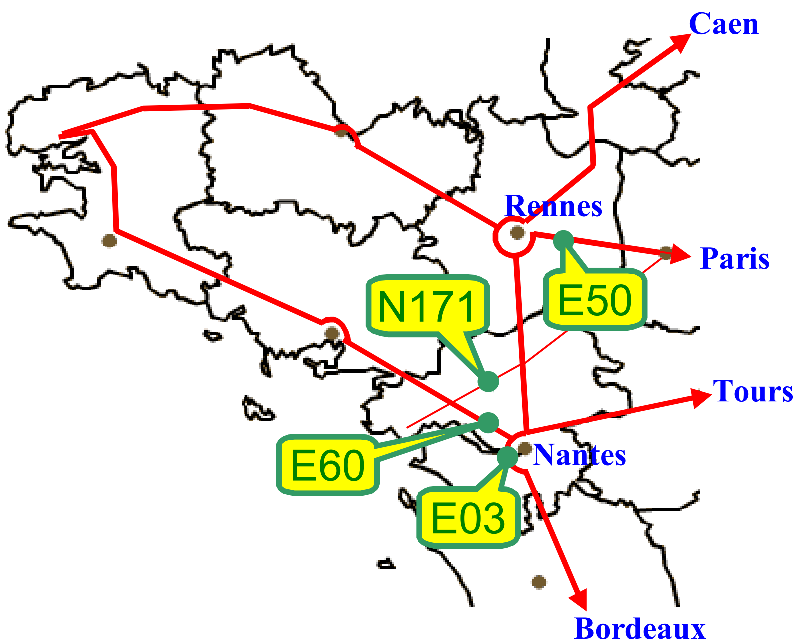


Figure 1 – Localisation des comptages

Ces routes sont, sauf la première, exploitées en voie express à deux fois deux voies.

Seuls les comptages « tous véhicules » ont été exploités ici, les données en poids lourds n'ayant pu être obtenus pour l'ensemble des cas traités. Les périodes considérées sont :

- Pour les comptages horaires, l'ensemble des années 2000 et 2001 ;
- Pour les comptages fins (par tranche de six minutes), la période du 20 au 29 juin 2002.

3.2 Le modèle déterministe de prévision du trafic

Une première tentative de modélisation a consisté à prédire le volume de trafic des 24 tranches horaires à partir d'un nombre limité de comptages : il a été ainsi montré que le comptage sur quatre tranches horaires suffisait à reconstituer avec une bonne précision l'ensemble des 24 tranches. Mais ce principe présentait deux inconvénients : d'abord les tranches ayant le meilleur pouvoir prédictif n'étaient pas les mêmes suivant le jour et la localisation, et ensuite ce modèle ignorait les variations à l'intérieur des tranches horaires.

L'idée est alors venue d'une fonction continue calculable à partir d'un nombre limité de paramètres, et c'est l'examen visuel des courbes qui a donné le chemin à suivre : pour les différentes stations on aperçoit à peu près toujours deux pointes de trafic, ayant la forme d'une courbe « en cloche », dont l'équation est identique à celle de la densité de probabilité de la loi normale :

$$T(t) = \frac{a}{\sqrt{2\pi} s} \exp(-(t-m)^2 / 2s^2)$$

où :

- ***T*** est la valeur prédite du flot de trafic,
- ***t*** est le temps, exprimé en heures décimales à partir de zéro heure pour le jour considéré,
- ***a*** représente le nombre de véhicules cumulé inclus dans une pointe,
- ***m*** correspond à la moyenne, soit la valeur de *t* qui maximise la fonction,
- ***s*** correspond à l'écart type, qui traduit l'étalement du phénomène.

Afin de prendre en compte l'effet des pointes en deçà de zéro heure et au delà de 24 heures, la valeur de *t-m* a été remplacée dans les calculs par *t-m-24* si *t-m > 12* et *t-m+24* si *t-m < -12.*

Le logiciel StatGraphics-Plus® pour Windows®, version 4.1, a été utilisé pour effectuer une régression non linéaire selon ce modèle, pour chaque station et chaque jour pour lesquels des comptages par six minutes étaient disponibles.

Il est rapidement apparu que la convergence n'était pas toujours obtenue, notamment si la période comprise entre les deux pointes présentait un troisième maximum ou un palier à un niveau élevé : il convenait donc de modéliser cette période intermédiaire, et l'ajout d'une troisième courbe en cloche répondait assez bien au problème posé.

Ceci donne la formule :

$$T(t) = a_0 + \sum_{i=1}^{i=3} \frac{a_i}{\sqrt{2\pi} s_i} \exp(-(t-m_i)^2 / 2s_i^{\,2})$$

La valeur prédite du trafic peut ainsi être calculée pour tout instant du jour, en utilisant dix paramètres (la constante a_0 plus les trois m_i, a_i et s_i). C'est ce que l'on appelle la

« composante déterministe » du modèle. La figure 2 montre le principe d'un tel calcul, et la figure 3 un résultat de régression non linéaire.

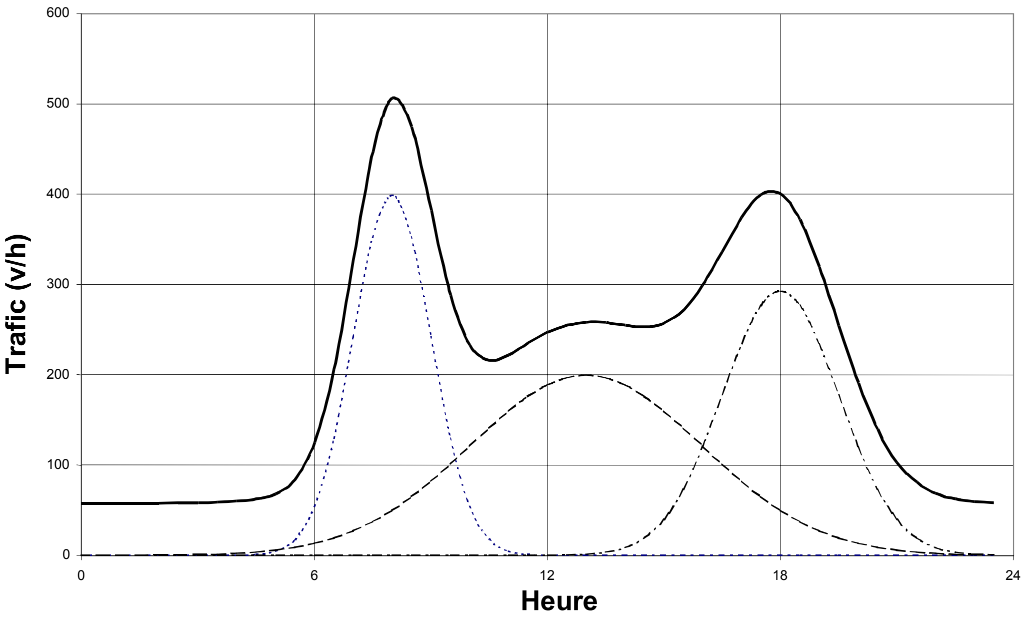
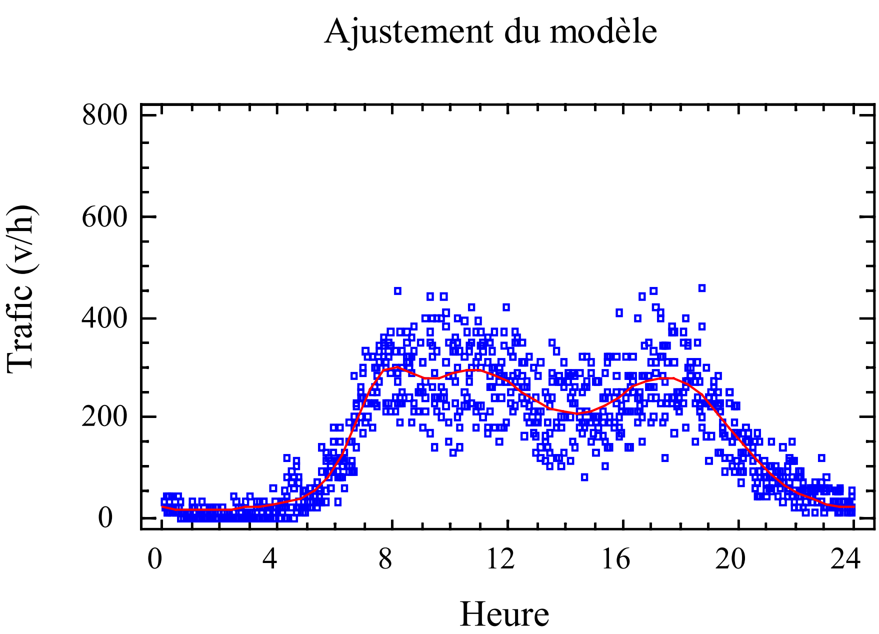


Figure 2 - Addition de trois courbes en cloche et d'une constante



Fi

Figure 3 – Exemple de résultat de régression non linéaire (N171 sens +, 24 au 27 juin 2002)

3.3 La composante probabiliste

Les comptages par six minutes étaient disponibles pour les quatre stations, sur une période allant du jeudi 20 au samedi 29 juin 2002. Mais pour pouvoir réaliser des prédictions sur la base de ces données, il convenait de calculer les paramètres du modèle sur un nombre suffisant de jours homogènes entre eux.

Pour cela, quatre jours consécutifs ont été pris en compte, allant du lundi 24 au jeudi 27 inclus ; le vendredi 28 a été exclus à cause de l'importante pointe de trafic observée en soirée (c'est l'amorce d'un week-end de départ en vacances). Bien sûr, même si le but recherché est d'obtenir des résultats caractéristiques des « jours de semaine », cette série est insuffisante pour caractériser l'ensemble des jours de l'année, ce d'autant plus que des variations saisonnières importantes ont été constatées lors de l'analyse des données par tranches horaires (qui elles sont disponibles sur deux années complètes).

Les paramètres du modèle ont ainsi été calculés pour la période de quatre jours, et les huit séquences de comptage (les deux sens de chaque poste) ; notons que, pour E60, il existe plusieurs plages de données manquantes, ce qui perturbe les résultats ; pour tous les autres postes, les données sont complètes.

La pertinence du modèle peut être estimée par le R^2, qui est le carré du coefficient de régression et traduit la part de *variance expliquée* (par exemple R^2=0,80 signifie que 80 % de la *variance* du trafic est *expliquée* par le modèle). Les résultats se situent entre 0,82 et 0,95 : il reste donc une part de variance non expliquée, qui constitue l'addition de trois phénomènes : l'imperfection du modèle, les différences entre jours successifs et les variations aléatoires.

Construire un modèle probabiliste revient à modéliser la probabilité des écarts au modèle déterministe : les résidus (différences entre valeur observée et valeur prédite) présentent une moyenne proche de zéro et un écart type égal à *l'erreur type* sur les prédictions ; leur étude a montré qu'ils suivaient une loi normale avec une précision acceptable (voir exemple en figure 4), ce qui autorise à fonder la composante probabiliste du modèle sur ce principe :

$$F(t, p) = \max(0, T(t) + S\, N(p))$$

Où $T(t)$ est la composante déterministe, S l'erreur type sur les prédictions et $N(p)$ la fonction normale inverse centrée réduite de la probabilité p ; la fonction « max » permet d'empêcher le résultat d'être négatif.

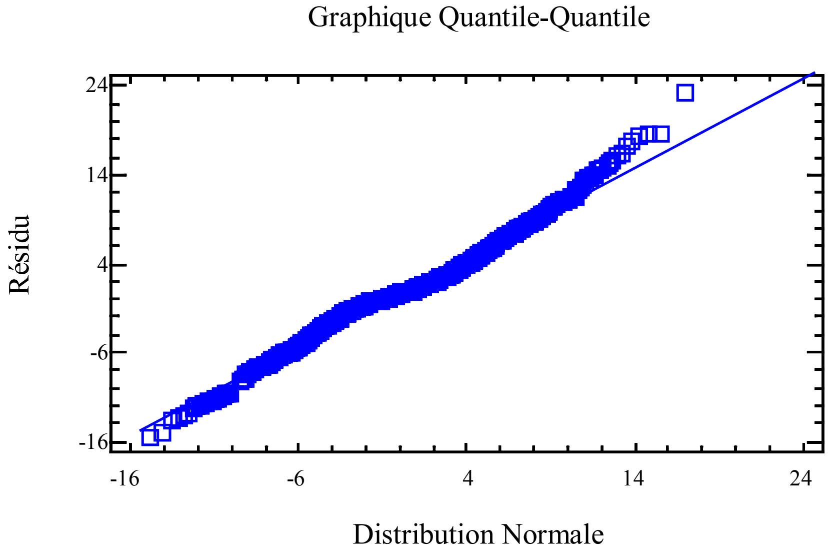


Figure 4 – Répartition des résidus comparée à la loi normale
(N171 sens +, 24 au 27 juin 2002)

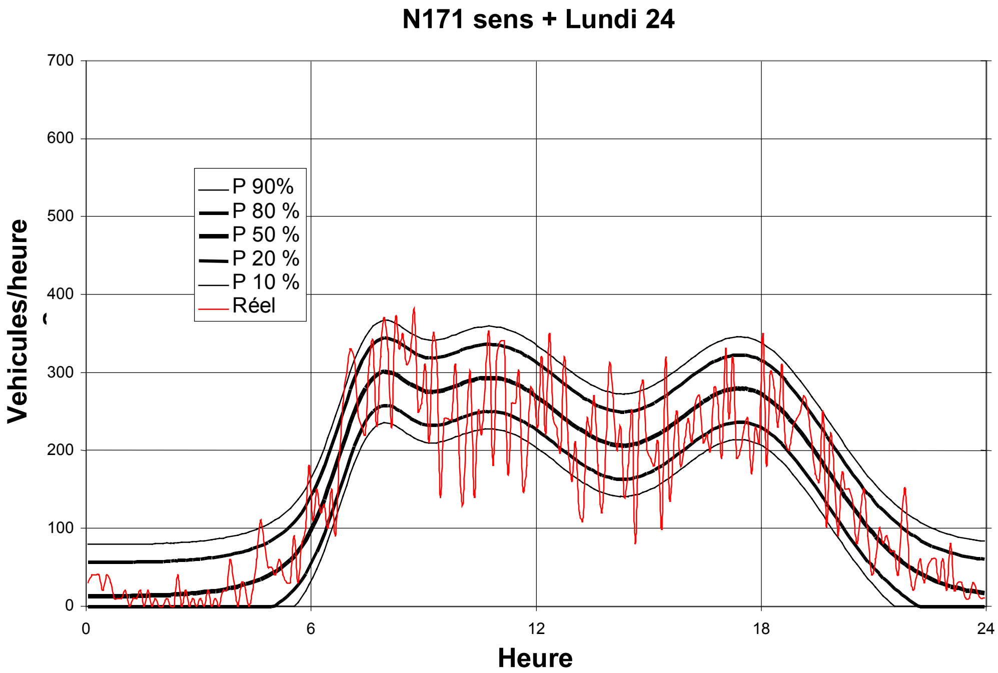
Cependant, deux difficultés sont apparues :

- d'une part, une corrélation, faible mais significative, existe entre la valeur prédite et la valeur absolue du résidu (en d'autres termes, plus le trafic risque d'être élevé, plus l'incertitude sur la valeur prédite est grande), mais cette tendance est trop dispersée pour qu'il soit intéressant de l'inclure dans le modèle ;

- D'autre part, une interdépendance existe entre des valeurs qui se suivent, ce qui se traduit par une *autocorrélation* positive pour un écart de temps inférieur à deux heures (en d'autres termes, si pour un instant *t* le trafic est supérieur à la normale, il y a une tendance significative pour qu'il en soit même dans l'intervalle $t \pm 2h$). La prise en compte de ce type d'interaction aurait nécessité une modélisation beaucoup plus complexe (du type de ce qui se fait en prévisions météorologiques), et a été, pour le moment, abandonnée.

Notons enfin que, si les prédictions probabilistes fondées sur une périodicité de six minutes sont bien en accord avec les comptages réels, le calcul à partir de comptages horaires élimine une partie de la variance, et donc diminue le terme probabiliste, qui devrait alors être corrigé à la hausse, selon un coefficient qui reste à déterminer.

La figure 5 illustre l'application de ces principes (calcul du trafic aux probabilités de 10, 20, 50, 80 et 90 % et données réelles).



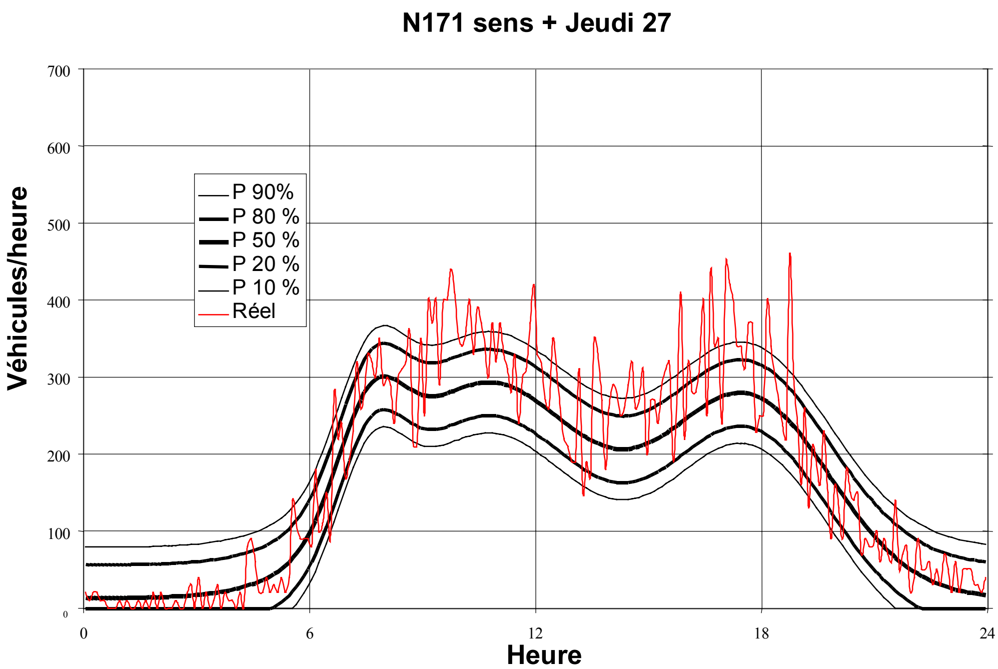


Figure 5 – Comptages réels (lundi 24 et jeudi 27 juin 2002) et prédictions

3.4 Simplifications possibles

Le modèle défini préalablement présente plusieurs inconvénients, tant théoriques que pratiques :

- la notion de « courbe de Gauss » est utilisée hors de son contexte, et il ne peut donc être fait légitimement appel à ses propriétés ;
- la régression non linéaire n'est pas d'un maniement simple (elle fait appel à un logiciel de statistiques, nécessite une évaluation préalable des paramètres, et n'aboutit pas toujours) ;

- la formule comporte dix paramètres à évaluer.

Pour corriger le premier point, la notion de courbe de Gauss est remplacée par celle de « courbe en cloche », ce qui ne change rien au problème, mise à part la définition des coefficients ; d'autre part, l'effet des « queues de courbe » en dehors de la période de 24 heures s'est avéré négligeable. La formule devient alors :

$$T(t) = \alpha_0 + \sum_{i=1}^{i=3} \alpha_i \exp(-\lambda_i (t - m_i)^2)$$

où T, t et m_i conservent leur signification, mais où α_i est le flot de trafic maximum attribué au pic et λ_i un coefficient d'étalement du pic (plus il est fort, plus le pic est concentré).

Une deuxième simplification a consisté à diminuer le nombre de coefficients ; il est apparu en effet que, dans la grande majorité des cas, il existait un pic centré autour de midi et un autre autour de 18 heures, avec des coefficients λ_i relativement constants (autour de 0,12) ; un autre pic existe pour les jours ouvrables autour de 8 heures, avec un λ_i plus fort (de l'ordre de 0,6).

Si les λ_i et les m_i sont fixés à des valeurs constantes, il ne reste plus que les quatre coefficients α_i à évaluer, et ceci peut se faire par une régression linéaire multiple, dont l'emploi est beaucoup plus commode que celui de la régression non linéaire (c'est une fonction standard de Microsoft Excel®).

Cela ne change rien au terme probabiliste du modèle ; cependant ce terme sera légèrement surévalué, la variance expliquée par le modèle étant alors plus faible.

4 MODELISATION DES BOUCHONS

4.1 Calcul des longueurs de queues

Connaissant le débit de trafic Q à l'instant t avec une probabilité p, on peut en déduire le nombre de véhicules présents dans la queue qui se forme lorsque Q dépasse la capacité C de la route (Q et C sont exprimés en véhicules par heure et par sens), qui elle-même peut être une variable aléatoire (mais, dans le cas présent, nous la traiterons comme une constante).

Le nombre de véhicules immobilisés résulte d'un calcul itératif sur des périodes de temps successives (de six minutes dans l'exemple), qui considère que :

- si au cours de la période précédente il n'y a pas de queue et si la capacité n'est pas dépassée, alors la longueur de queue reste nulle ;
- dans les autres cas, la différence entre Q et C s'ajoute algébriquement au nombre de véhicules présents dans la queue, et si le résultat est négatif la queue disparaît.

Ce calcul suppose bien sûr qu'aucune déviation ne soit possible et que l'existence de la queue n'ait aucune influence sur le nombre de véhicules qui y entrent : dans la réalité, lorsqu'une queue importante se forme, des itinéraires de délestage se mettent en place soit de façon organisée, soit spontanément, et le débit de trafic s'en trouve alors réduit. Cette redistribution du trafic devra être prise en compte, sous peine de déboucher sur une estimation très excessive de la longueur de queue.

La figure 6 donne un exemple de calcul, et le tableau qui suit donne les résultats chiffrés. La probabilité à *p* % signifie ici que le volume de trafic a une probabilité de *p* % de se situer au-dessous de la valeur indiquée.

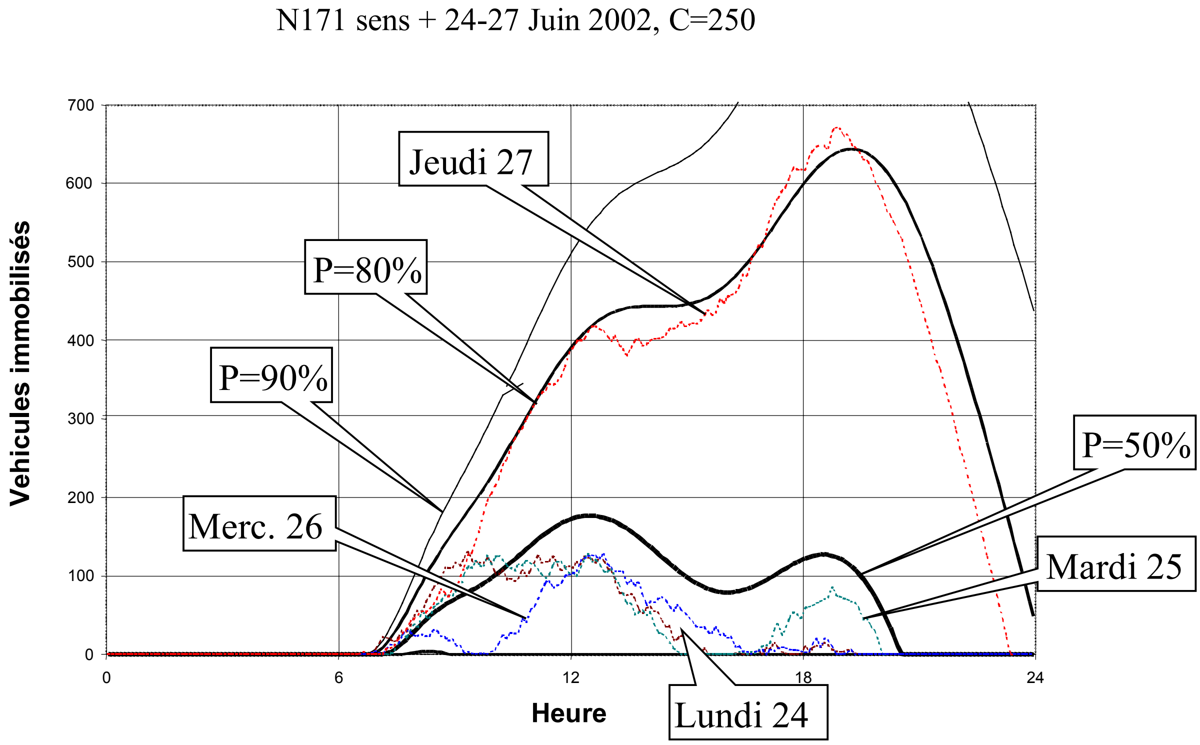


Figure 6 – Exemple d'estimation de la longueur de queue (N171 sens +, C=250 v/h)

Tableau 1 - Caractéristiques calculées des bouchons pour deux valeurs de la capacité résiduelle *C*

Données	Trafic total journalier	*C* = 250 v/h			*C* = 300 v/h		
		Durée de bouchon	Longueur maxi (*)	Heures perdues	Durée de bouchon	Longueur maxi (*)	Heures perdues
Lundi 24	3735	11h00	130	693	3h30	31	34
Mardi 25	3683	11h00	127	762	3h50	18	28
Merc. 26	3701	11h45	127	488	3h20	28	27
Jeudi 27	4416	16h55	671	5923	12h05	159	847
Prob 10%	3549	0h00	0	0	0	0	0
Prob 20%	3663	1h25	4	3	0	0	0
Prob 50%	3881	13h20	177	1389	0	0	0
Prob 80%	4099	17h15	644	6557	12h30	141	820
Prob 90%	4213	>17h30	931	>9935	14h00	311	2988

(*) Nombre maximal de véhicules simultanément dans la queue

Dans cet exemple, les trois premiers jours se situent à un niveau de probabilité proche de 50 %, bien que légèrement au-dessous ; le jeudi en revanche se rapproche de 80 %.

4.2 Espérance mathématique du nombre d'heures perdues

Les calculs précédents permettent l'évaluation de la longueur de file d'attente *L* à un moment donné *T*, pour une capacité donnée de route *C* et une probabilité donnée *p* du volume de trafic ; on peut alors additionner les valeurs successives, afin d'obtenir le total du temps perdu pendant la période de temps donnée, faite de *N* périodes unitaires Δt :

$$H(C, p) = \Delta t \sum_{i=1}^{i=N} L_p(i\Delta t, C)$$

Les calculs qui suivent portent sur un jour, avec une capacité *C* constante. La longueur de queue est initialisée à zéro pour *T = 0*, et le calcul s'arrête à *T = 24 h*, ce qui revient à ignorer tout reliquat du jour précédent, et tout transfert sur le jour suivant.

L'étape suivante consiste à calculer le temps perdu moyen par jour sur un nombre élevé de cas semblables ; cette valeur est *l'espérance mathématique E*, qui peut résulter d'un calcul intégral du temps perdu *H(C, p)* depuis la probabilité zéro jusqu'à la probabilité 1 :

$$E(C) = \int_{p=0}^{p=1} H(C, p)\,dp$$

Le calcul intégral ne pouvant pas se faire d'une façon simple, on peut en réaliser une approximation selon la formule :

$$E(C) = \sum_{i=1}^{i=n} k_i H(C, p_i)$$

où k_i est la fraction de probabilité attribuée à p_i :

- la solution « triviale » consiste à prendre $n = 1$, $k_1 = 1$, $p_1 = 0{,}5$, mais cela revient à ignorer totalement l'aspect probabiliste ;
- le calcul par *centile* prend n = 100, $k_i=0{,}01$ et $p_i = 0{,}01\ i - 0{,}005$;
- k_i et p_i peuvent être optimisés, de façon à limiter le nombre de tranches tout en améliorant la précision du résultat.

La figure 7 montre un exemple de ce type de calcul.

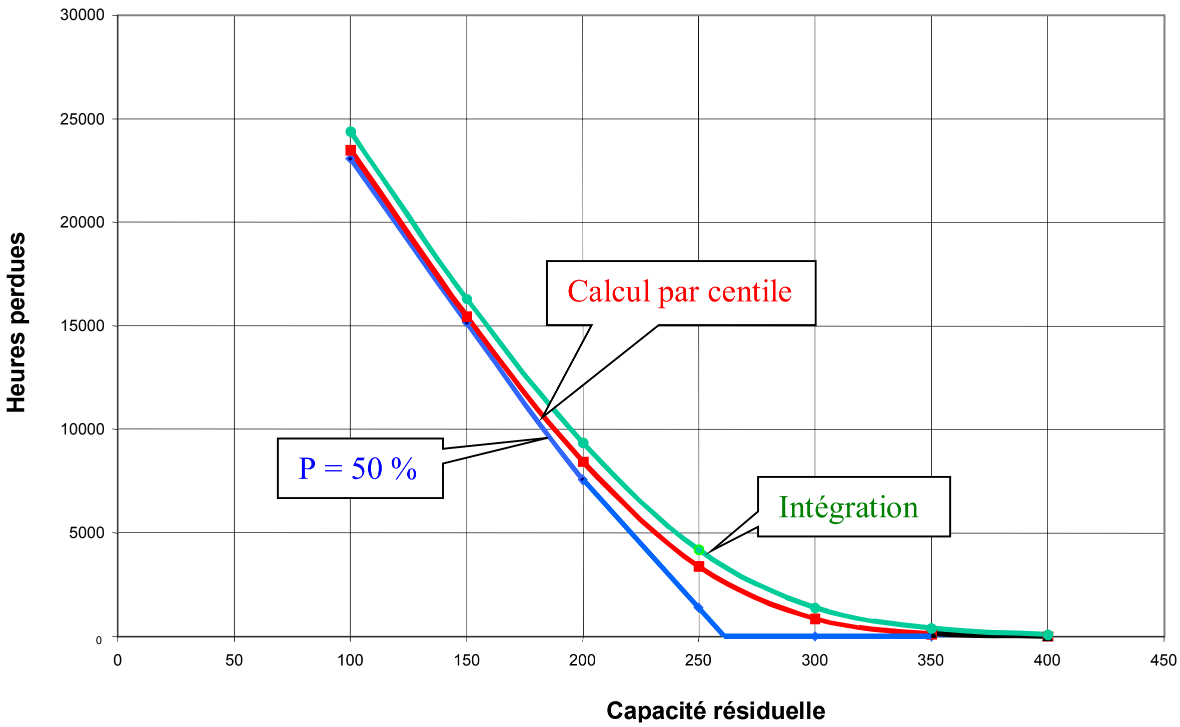


Figure 7 – Exemple d'estimation de l'espérance mathématique du temps perdu (N171 sens +)

5 CONCLUSIONS

Le modèle proposé permet de calculer pour tout temps *t* de la journée une probabilité *p* d'atteindre un débit de trafic *Q*, et donc, connaissant la capacité de la route *C* au même instant *t*, de déterminer la quantité de véhicules retenus et donc les caractéristiques des bouchons.

Ce modèle assimile l'évolution journalière du débit à la somme de trois pointes, ainsi que d'une constante et d'un terme représentant le caractère aléatoire du trafic.

Les coefficients peuvent être calculés par le biais d'une régression non linéaire fondée sur des comptages réels effectués par tranche d'une heure sur un nombre suffisant de journées représentatives du cas considéré ; une variante simplifiée consiste à calculer seulement la hauteur des pointes et le niveau de base, les autres paramètres étant alors constants, ce qui permet d'utiliser une régression linéaire multiple.

Un modèle simple a été appliqué pour calculer le temps perdu dans les files d'attente pour un jour et une probabilité donnés ; ensuite l'intégration des résultats en fonction de la probabilité permet le calcul de « l'espérance mathématique » du temps perdu.

REMERCIEMENTS

Nous citerons particulièrement :

- Sébastien SIAUDEAU, étudiant de quatrième année à l'Institut de Mathématiques Appliquées (IMA) de l'Université Catholique de l'Ouest (UCO) à Angers, pour la contribution apportée dans le cadre de son stage effectué au LCPC d'avril à juillet 2002.
- Frédéric PESTEIL, du CETE du sud-ouest (DTCES/CER), auteur du logiciel ECCU dont certains éléments ont été repris dans la modélisation des queues.
- Robert ABADIE et Guylaine FILY, du CETE de l'ouest, pour la fourniture des données de trafic.

RÉFÉRENCES

[1] Boiteux et al. (2000) Transport : pour un meilleur choix des investissements. Commissariat Général du Plan, Paris

[2] FORMAT *Inception report* (2002), Commission Européenne, programme GROWTH, Bruxelles.

[3] SIREDO (2002) Document E9655, SETRA, France