DOCUMENT DE TRAVAIL

Compteurs sur le toit

Examen statistique des résultats obtenus

CONSOMMATION ET
CORPORATIONS CANADA
MÉTROLOGIE LÉGALE

LKC
TK
301
.R614
1984

IC

Canada

TK301
R6f

Mémoire technique

Examen statistique des résultats obtenus

dans le projet "Compteurs sur le toit".

M. Romanowski

et

E. Green

Industry Canada
Library - Queen
OCT 17 2014
Industrie Canada
Bibliothèque - Queen

Métrologie légale

Consommation et Corporations Canada

DEPARTMENT OF CONSUMER & CORPORATE AFFAIRS
LIBRARY
NOV 8 1984
BIBLIOTHÈQUE
MINISTÈRE DE LA CONSOMMATION ET DES CORPORATIONS

L'analyse et les conclusions contenues dans ce rapport ne représentent pas nécessairement le point de vue du ministère.

© Ministre des Approvisionnement et Services Canada 1984

N° de cat. RG 23-80/1984F

ISBN 0-662-92723-0

Available in English

REMERCIEMENTS

Les auteurs expriment leur gratitude au directeur, M. R.G. Knapp, pour son aide et ses encouragements constants tout au long de la préparation du présent mémoire.

Ils tiennent aussi à remercier LEXICA INC. qui a assuré avec compétence la dactylographie des versions anglaise et française du manuscrit.

TABLE DES MATIÈRES

Préface 1

Description du projet 5

Description de l'analyse de la variance 11

Calculs numériques 17

Autres recherches statistiques 21

Appendice I

Tableau I-B à VIII 27

Appendice II

Les erreurs du disque 37

PRÉFACE

Pour comprendre les raisons qui ont conduit au projet expérimental dont les résultats sont analysés dans le présent rapport, il est utile de souligner les événements qui ont eu lieu au début des années 60 et de revoir les lignes de conduite de la Direction de la métrologie légale en matière de vérification, de scellage et de revérification des compteurs.

La Loi sur l'inspection de l'électricité précisait que tous les appareils de mesure utilisés à des fins commerciales devaient, avant d'être mis en service, être vérifiés en vue d'établir leur exactitude en fonction des limites d'erreur autorisées, puis être scellés par un inspecteur du gouvernement pour les mettre à l'abri de toute falsification. Ces dispositions s'appliquaient aussi, cela va de soi, aux compteurs d'électricité domestiques.

Étant donné qu'on ne peut s'attendre qu'un compteur demeure exact indéfiniment, la Loi prévoyait également que chaque compteur devait être vérifié de nouveau à intervalles fixes. À l'origine, la durée du scellage attestant la vérification avait été fixée, par voie de réglementation, à six ans pour tous les compteurs d'électricité, mais, en ce qui a trait aux compteurs domestiques, cette période a, par la suite, été étendue à huit ans. Ainsi, jusqu'au début des années 60 les services publics avaient l'habitude de retirer du service tous les compteurs dont la période de validité du scellé était expirée, d'effectuer un travail de révision et d'entretien dans un atelier, puis de les présenter à un inspecteur du gouvernement pour qu'il les vérifie et les scelle de nouveau.

Les progrès techniques réalisés dans la conception et la fabrication des compteurs, notamment la mise au point des compteurs à suspension magnétique, ont suscité, tant de la part des services publics que des fabricants, des demandes de prolongation de la durée du scellage, fondées sur l'argument qu'un compteur de qualité supérieure pouvait satisfaire pendant une période plus longue aux tolérances permises.

À la même époque, la Direction de métrologie légale était en train d'élaborer et de parfaire ses plans d'échantillonnage statistique pour la vérification et la revérification des compteurs. Au sein de son effectif, plusieurs ont alors décidé d'entreprendre un certain nombre d'expériences à long terme en vue d'éprouver la validité des solutions en voie d'être adoptées en matière de vérification et d'aider à l'élaboration des lignes directrices de la Direction en ce qui concerne les aspects métrologiques des compteurs watt-heuremètres domestiques. De plus, l'idée qu'un compteur correctement fabriqué pouvait être installé indifféremment à l'intérieur ou à l'extérieur sans que cela n'altère de façon sensible son fonctionnement à long terme, a encouragé l'utilisation généralisée des compteurs à suspension magnétique. Le personnel de la Direction estimait que cette idée, si logique et si attrayante qu'elle fût, devait être vérifiée expérimentalement dans un milieu naturel pendant plusieurs années consécutives. C'est ainsi qu'est né, en 1968, le projet des "Compteurs sur le toit".

Cent quarante-quatre compteurs, soit trente-six provenant de chacun des quatre fabricants de l'époque, furent installés sur une structure construite à cette fin sur le toit de l'immeuble de la Métrologie légale. Avant d'être installé pour le programme d'essai, chaque compteur a fait l'objet d'un étalonnage et d'une mise au point, sur courants de 2,5 et de 25 ampères, pour ramener sa limite d'erreur le plus près possible de zéro. En 1970, l'analyse des observations trimestrielles de la première année complète révélait que les résultats n'étaient pas concluants parce que chacun des trois sous-groupes de compteurs était alimenté par une des phases distinctes d'un système de triphasé. Il a donc été décidé de vérifier les compteurs après les avoir connectés en série sur courant monophasé. On a placé, sur ce même courant, trois compteurs "moniteurs" étalonnés situés dans l'immeuble. Ces compteurs devaient enregistrer, aux fins des calculs statistiques, le nombre considéré "exact" de watts-heures consommés par les compteurs installées sur le toit.

À la suite de cette modification, les observations faites au cours de chaque trimestre de 1971 ont été analysées et décrites dans un rapport technique publié en 1975. Le relevé trimestriel des erreurs d'enregistrement s'est poursuivi de façon ininterrompue jusqu'en octobre 1979, mois qui marquait la fin de la phase expérimentale du projet. Les relevés directs des compteurs,

1. DESCRIPTION DU PROJET

Un bref historique du projet, qui contient la description du montage original, est donné dans l'Introduction rédigée par Richard G. Knapp, directeur de la Métrologie légale.

Au point de vue purement statistique, chacune des années de l'expérience, qui a duré pendant huit années consécutives, est représentée par quatre "échantillons" de EET (erreur d'enregistrement totale). Chaque échantillon est constitué par les relevés effectués sur la totalité des 144 compteurs observés. Chacun des quatre manufacturiers (General Electric, Ferranti-Packard, Westinghouse, Sangamo) est représenté par un groupe de 36 compteurs considérés comme étant nominalement identiques entre eux. Un "groupe" est subdivisé en quatre "cellules": l'une d'elles fait face au sud (S), une à l'est (E), une au nord (N) et une à l'ouest (O). Une cellule (9 compteurs) est ainsi le plus petit ensemble où deux compteurs peuvent être interchangés sans influencer les résultats de l'analyse.

Les enregistrements étaient effectués quatre fois par an: en février, mai, août et novembre. Si aucun des compteurs n'était éliminé, ceci aurait conduit à un total de 4608 EET. En réalité, ce total se trouve quelque peu réduit par le fait que le nombre initial de 144 compteurs a progressivement diminué au cours des années et se trouva finalement réduit à 136.

Une opération d'enregistrement, c.à.d. le relevé des indications de tous les compteurs, représente pour une équipe de deux membres du laboratoire d'électricité deux ou trois journées de travail. La réduction des relevés c.à.d. leur conversion en "pourcentage" (%) demande beaucoup de temps. Par

suite du manque de main d'oeuvre (techniciens et calculateurs), au début de notre analyse seulement les résultats des mesures de novembre ont pu être réduits. L'analyse de la variance était donc appliquée aux résultats des huit années consécutives (novembre à novembre), et le milésime, p.ex. 1972 indique (en %) les EET de la période allant de novembre 1971 à novembre 1972. (Voir table IA.)

Après que quelques calculs aient été effectués, il est devenu évident que pour des analyses autres que celle de la variance, des échantillons plus grands seraient plus appropriés. On a alors pris la décision de réduire aussi les relevés effectués en mai. Les calculs décris dans la section 4 sont ainsi effectués sur des échantillons qui, par consentement général, sont considérés comme "grands": leurs dimensions sont de l'ordre de 2000. Les conclusions qui sont tirées des échantillons de cette grandeur sont considérées comme suffissament probantes pour conduire à des indications pratiques. Avant d'appliquer les méthodes de l'analyse statistique, examinons d'abord la nature de quelques questions générales qui peuvent être posées et auxquelles l'analyse qui suit peut fournir des réponses plus ou moins complètes.

a) Tenue générale et fonctionnement des instruments quand ils sont utilisés dans des conditions particulières et la probabilité qu'ils vont se comporter d'une manière similaire quand ils sont ensuite placés dans les conditions d'utilisation normales. Il est évident qu'il est très important, pour le manufacturier, pour les institutions métrologiques et, cela va de soi, pour le consommateur de l'énergie électrique de déterminer quelles conditions atmosphériques peuvent influencer le fonctionnement d'un compteur. Est-il possible de prédire comment la position d'un compteur sur une maison ou une usine peut influencer la somme d'argent que l'usager peut avoir à payer? Le fait qu'un compteur est placé sur le côté ensoleillé d'une maison favorise-t-il le producteur ou le consommateur de l'énergie électrique?

TABLE IA (1972)

	GE i = 1	FP i = 2	W i = 3	S i = 4	h = 4
SOUTH	Y_{11} = - 0.01	$Y_{2,1}$ = + 0.72	$Y_{3,1}$ = + 0.62	$Y_{4,1}$ = + 1.24	
	- 0.20	+ 0.56	+ 0.60	+ 1.40	
	+ 0.12	- 0.95	+ 0.09	+ 0.72	
	+ 0.39	+ 0.83	+ 0.76	+ 1.15	
	+ 0.07	+ 0.24	+ 0.14	+ 0.80	
	- 0.17	- 0.40	- 0.72	+ 0.62	
	+ 0.16	+ 0.44	+ 0.73	+ 0.64	
	- 0.01	- 0.91	+ 1.06	+ 1.02	
	Y_{19} = + 0.54	- 0.49	+ 0.87	+ 0.95	
EAST	$Y_{1,10}$ = + 0.31	$Y_{2,10}$ = - 0.13	$Y_{3,10}$ = + 0.05	$Y_{4,10}$ = + 0.70	
	+ 0.30	+ 0.27	- 0.29	+ 0.64	
	+ 0.11	+ 0.43	- 0.21	+ 1.20	
	+ 0.65	+ 0.34	- 0.38	+ 0.57	
	+ 0.36	+ 0.07	+ 1.17	+ 0.89	
	+ 0.58	- 0.07	+ 0.60	+ 0.75	
	+ 0.56	+ 0.40	- 0.14	+ 0.92	
	+ 0.65	+ 0.53	- 0.18	+ 0.87	
	$Y_{1,18}$ = + 0.47	+ 0.05	+ 0.24	+ 0.81	
NORTH	$Y_{1,19}$ = + 0.32	$Y_{2,19}$ = + 0.26	$Y_{3,19}$ = + 0.39	$Y_{4,19}$ = + 0.51	
	+ 0.11	+ 0.69	- 0.07	+ 0.60	
	+ 0.07	+ 0.98	+ 0.68	+ 0.98	
	+ 0.65	+ 0.55	- 0.18	+ 0.44	
	+ 0.72	+ 0.83	+ 0.59	+ 0.72	
	+ 0.69	+ 0.30	- 0.03	+ 0.25	
	+ 0.68	+ 0.62	+ 0.98	+ 0.92	
	+ 0.62	+ 0.31	+ 0.31	+ 0.81	
	+ 0.35	+ 0.42	+ 0.03	+ 1.20	
WEST	$Y_{1,27}$ = + 0.22	$Y_{2,27}$ = + 0.50	$Y_{3,27}$ = + 0.27	$Y_{4,27}$ = + 0.27	
	+ 0.46	+ 0.66	+ 0.40	+ 2.50	
	+ 0.36	+ 0.30	+ 0.03	+ 0.59	
	+ 0.05	+ 0.77	+ 0.28	+ 0.45	
	+ 0.42	- 0.03	+ 0.15	+ 0.47	
	+ 0.99	+ 0.50	- 0.43	+ 0.56	
	+ 0.40	+ 0.04	+ 0.56	+ 1.65	
	+ 0.40	+ 0.20	+ 0.57	+ 0.88	
	$Y_{1,36}$ = + 0.63	$Y_{2,36}$ = + 0.32	$Y_{3,36}$ = - 0.08	$Y_{4,36}$ = + 0.45	
	t_1 = +13.02	t_2 = +10.20	t_3 = + 9.46	t_4 = +30.14	t_0 = 62.82
	t_1^2 = 169.52	t_2^2 = 104.04	t_3^2 = 89.49	t_4^2 = 908.42	
					t_0^2=3946.35

b) Influence de critère "manufacturier".* Quoique tous les compteurs sont supposés être construits "identiquement", il est évident que chaque manufacturier est susceptible d'introduire quelques détails de construction dont la présence peut influencer les résultats.

c) Influence du critère "direction géographique". Nous avons ici l'une des raisons majeures du fait que l'expérience a été entreprise. Il est, en effet, impossible d'affirmer à priori qu'un compteur tourné vers le sud va enregistrer l'électricité de la même manière qu'un compteur tourné vers le nord. L'orientation peut agir non seulement par la température mais aussi par les facteurs tels que vents, humidité, etc.

d) Influence du critère "saison". Ceci pose des problèmes qui sont, du point de vue des calculs statistiques, identiques à ceux posés par le critère "direction".

L'information fournie par l'analyse de ces problèmes est aussi importante par les conséquences sur le plan scientifique et technique que sur le plan économique.

La méthode mathématique utilisée pour comparer le comportement des compteurs en tenant compte d'un certain critère de classification est celle de l'analyse de la variance. Les principes de ce puissant outil mathématique ont été établis par R.A. Fisher et sont, en général, exposés dans tous les manuels de statistique.

---

* En statistique les termes tells que "manufacturier" "direction" etc. désignent des "critères d'après lesquels un échantillon peut-être classifié.

Pour aider le lecteur, les fondements de l'analyse de la variance, tels qu'ils sont utilisés dans le présent rapport, sont décrits brèvement dans la section suivante.

2. DESCRIPTION DE L'ANALYSE DE LA VARIANCE

Dans le présent rapport (comme dans tous les travaux précédents sur les sujets analogues) le symbole Y (lettre capitale munie d'indices appropriés) désigne l'erreur d'enregistrement totale (EET) calculée sur une période annuelle c.à.d. de novembre à novembre. Un nombre tel que Y_k (compris entre 1 et 144) indique le "pourcentage" par lequel l'enregistrement du k^{eme} compteur s'écarte de celui des moniteurs*2). Les moniteurs sont des étalons de référence placés dans un laboratoire (non à ciel ouvert). Ils forment la base pour l'évaluation des enregistrements des compteurs étudiés. Si, par ex., Y_{10} = 0.23 cela signifie que le compteur de rang 10 a enregistré (après avoir été sous tension pendant douze mois) une quantité d'énergie électrique de 0.23% inférieure à celle enregistré par les moniteurs. Cette dernière est parfois appelée, quoique pas toujours correctement, le "vrai" enregistrement, en kWh.

Pour rendre le présent rapport plus facilement accessible aux non-statisticiens, seulement l'analyse de la variance avec un seul critère sera utilisée: les 144 compteurs seront donc subdivisés en quatre classes ou bien d'après le critère "direction" ou le critère "manufacturier". Au début de l'expérience (1972) toutes les classes étaient de la même grandeur (36 compteurs). Ainsi qu'il a déjà été dit plus haut les grandeurs initiales n'ont pas pu être maintenues car au cours des années, certains compteurs ont été (pour diverses raisons) éliminés mais non remplacés.

---

* 2 Les moniteurs sont des étalon de référence placés dans un laboratoire et non à l'extérieur. Ils constituent la base pour l'évaluation de l'enregistrement des étalons examinés.

Si un échantillon de N éléments Y_k (k = 1, 2, - - - N) est considéré dans sa forme non classifiée, la moyenne est donnée par la formule

$$\overline{Y} = \frac{1}{N} \sum_{k=1}^{N} Y_k. \qquad \ldots(1)$$

La relation la plus importante dans tous les calculs qui suivent est celle qui exprime la somme des carrés des déviations des éléments par rapport à n'importe quel nombre A, en fonction des déviations par rapport à la moyenne $\overline{Y}$ de l'échantillon.

$$\sum_k (Y_k - A)^2 = \sum_k (Y_k - \overline{Y})^2 + N(\overline{Y} - A)^2. \qquad \ldots(2a)$$

Lorsque A = 0, ceci se réduit à

$$\sum_k Y_k^2 = \sum_k (Y_k - \overline{Y})^2 + N\overline{Y}^2. \qquad \ldots(2b)$$

Si l'échantillon est distribué en "classes" ces classes seront désignées par l'indice i; le total des éléments de la $i^{ème}$ classe sera désigné par le symbole n_i. Le rang dans la classe sera indiqué par l'index j de sorte que j = 1, 2, ...n_i. Un élément aura ainsi deux indices p.ex. Y_{ij}:

Y_{ij} avec i = 1, 2, ...h (c.à d. h classes; ici h = 4)
j = 1, 2, ...n_i (n_i éléments dans $i^{ème}$ classe)

La $i^{ème}$ classe constitue en elle même un échantillon non classifié de sorte qu'il est possible de lui appliquer la relation (1) dans laquelle a) l'indice k est remplacé par l'indice j lequel prend alors toutes les valeurs de j = 1 à j = n_i et b) $\overline{Y}$ est remplacé par la moyenne $\overline{Y}_i$ de la classe, laquelle est donc égale à:

$$\overline{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}. \qquad \ldots(3)$$

Ceci donne

$$\sum_{j=1}^{n_i}(Y_{ij} - A)^2 = \sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 + n_i(\overline{Y}_i - A)^2.$$

Si dans cette expression on fait A égal à la moyenne générale $\overline{Y}$ on obtient:

$$\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y})^2 = \sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 + n_i\ (\overline{Y}_i - \overline{Y})^2. \qquad \ldots(4)$$

Finalement, la sommation par rapport à toutes les classes h conduit à la formule fondamentale de la <u>décomposition de la somme des carrés:</u>

$$\sum_{i=1}^{h}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y})^2 = \sum_{i=1}^{h}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 + \sum_{i=1}^{h}(\overline{Y}_i - \overline{Y})^2 n_i \qquad \ldots(5)$$

Une très importante propriété de cette égalité réside dans le fait qu'un examen de ses trois termes montre que chacun d'eux peut conduire à une estimation de la variance de l'échantillon (à condition d'évaluer correctement le nombre des degrés de liberté). Pour le calcul de l'estimation, le terme doit être divisé par son propre nombre de degrés de liberté. Les estimations sont non-biaisées c.à.d. qu'elles convergent toutes vers la même limite lorsque la grandeur de l'échantillon tend vers l'infini.

Les expressions de l'estimation sont données plus bas. Les symboles pour les estimations des variances son s^2, s'^2 s''^2 et les degrés de libertés qui leur correspondent sont ν, ν', ν''.

$$s^2 = \frac{1}{N-1}\sum_{ij}(Y_{ij} - \overline{Y})^2\ ; \qquad \nu = N-1. \qquad \ldots(6a)$$

$$s'^2 = \frac{1}{N-h}\sum_{ij}(Y_{ij} - \overline{Y}_i)^2\ ; \qquad \nu' = N-h. \qquad \ldots(6b)$$

$$s''^2 = \frac{1}{h-1}\sum_{ij} n_i(\overline{Y}_i - \overline{Y})^2\ ; \qquad \nu'' = h-1. \qquad \ldots(6c)$$

Les bases de l'analyse de la variance peuvent être présentées comme suit. Si l'univers des éléments (p.ex. un très grand nombre d'observations) dont un échantillon de 144 éléments a été tiré, était homogène, alors aucune classification (d'après un critère formel) ne pourrait introduire une structure quelconque statistiques. De sorte que les trois estimations de la variance ne pourraient être sensiblement différentes l'une de l'autre. Si, cependant, les estimations sont sensiblement différentes, alors cette différence indique qu'il doit exister une raison systématique c.à.d. une vériable cause de variation (terme adopté par les statisticiens). Celles des causes qui pertubent la valeur de s'^2 sont nommées causes des variations à l'intérieur des classes et celles qui pertubent s''^2 sont nommées causes des variations entre les moyennes des classes.

Dans la statistique mathématique, le rapport le plus utilisé est s'^2/s''^2 ou son inverse. Il est représenté par le symbol F en honneur de R.A. Fischer qui a créé et développé l'analyse de la variance. La définition de F est:

$$\text{si } s'^2 > s''^2, \text{ alors } F = \frac{s'^2}{s''^2},$$

$$\text{si } s'^2 < s''^2, \text{ alors } F = \frac{s''^2}{s'^2}. \qquad \ldots(7)$$

Dans la plupart des manuels de statistiques, les tables de F sont sous une forme abrégée. Des tables plus complètes se trouvent dans "CRC-Standard Math. Tables". Dans le présent rapport il y a h = 4 classes, par conséquent $\nu_1 = 4-1 = 3$. La valeur de ν_2 est égale à $\nu_2 = N-4 = 140$. Dans les tables de F, pour $\nu_1 = 3$ et $\nu_2 = 140$ il y a deux valeurs de F: la valeur supérieure appelée "point 5%" est plus grande que la valeur inférieure appelée "point 1%" (souvent imprimée en caractères gras). Par rapport à ces deux "points", égaux ici 2.4 et 3.4 repectivement, les valeurs numériques de F peuvent être

a) inférieures au point 5%,

b) supérieures au point 1%,

c) entre ces deux points.

Supposons que l'hypothèse à tester est que les variances s'^2 et s''^2 des échantillons sont tirées d'une population complètement homogène et normalement (ou presque normalement) distribuée. La valeur a) ci-dessus est celle qui dans un échantillonage répété serait dépassée avec une probabilité égale à 0.05 (c.à.d. 5%). La valeur b) serait, par contre, dépassée avec une probabilité de 0.01 (c.à.d. 1%). Si la valeur de F est en dessous du "point 5%" elle peut être considérée comme insignifiante et ceci serait alors en faveur de l'hypothèse formulée c.à.d. que la population est homogène.

Cette section peut être résumée comme suit:
si un échantillon est classifié d'après le critère "manufacturier" et si cette classification n'a pas d'effet systématique (ou, ce qui revient au même, les compteurs d'un manufacturier ne peuvent être distingués de ceux des autres manufacturiers) les variances s'^2 et s''^2 sont vraisemblablement proches l'une de l'autre. On peut en dire autant au sujet du critère "direction". Si F est grand c.à.d. si s'^2 est en réalité très différent de s''^2 alors le critère a un effet considérable sur le comportement des compteurs. Ainsi qu'il a été dit plus haut, les points de repère commodes son ceux qui sont nommés "point 5%" et "point 1%". Les points situés entre les deux ne peuvent pas être évalués plus exactement et peuvent demander un complément d'investigation.

Finalement, il faut souligner fortement, en tant que la conclusion de cette esquisse théorique de l'analyse de la variance, que cette méthode, quoique très efficace (pour juger la probabilité d'une hypothèse) ne suggère pas pour quelles raisons réelles une hypothèse doit être retenue ou rejetée. Il reste à l'expérience et à l'imagination de l'observateur - calculateur de découvrir les vraies raisons qui produisent les effets observés.

3. CALCULS NUMÉRIQUES

La forme des équations (6) n'est pas bien adaptée aux calculs numériques. Pour indiquer la manière dont ces calculs doivent être effectués, les résultats des opérations faites en novembre 1972 vont être traités en détail. Les principaux échelons des calculs sont aussi reproduits sous une forme plus condensée dans le tableau "sortie" (output) de la machine calculatrice, les symboles utilisés étant partout les mêmes. Les calculs de toutes les autres années sont présentés uniquement sous forme de tableaux "sorties" fournis par la machine (Voir Appendice I).

Les observations de 1972 (EET sous forme de %) figurent dans le Tableau IA. Elles sont classées verticalement d'après le critère "manufacturier": General Electric (GE), Ferranti-Packard (FP), Westinghouse (W), Sangamo (S). Le tableau IA peut aussi être utilisé horizontalement. Dans ce cas on prend come critère la direction géographique: Sud (S), Est (E), Nord (N), Ouest (W). Les calculs qui suivent se rapportent à la classification verticale. Leur suite est donnée ci-dessous:

a) Totaux des classes: $t_i = \sum_{j=1}^{36} Y_{ij}$. Ceci donne t_1, t_2, t_3, t_4. ...(8)

Grand total t_0 : $t_0 = t_1 + t_2 + t_3 + t_4$.

Carrés: t_1^2, t_2^2, t_3^2, t_4^2 et $t_0^2 = (t_1 + t_2 + t_3 + t_4)^2$

b) Somme des carrés S:

$$\Sigma = \sum_{k=1}^{144} Y_k^2 = \sum_{i=1}^{4} \sum_{j=1}^{36} Y_{ij} \; ; \; k = 1. 2, \ldots 144 \qquad \begin{array}{l} i = 1, 2, 3, 4 \\ j = 1, 2 \ldots 36 \end{array} \qquad \ldots(9)$$

c) Carrés moyens des classes μ_i: $\mu_i = \frac{t_i^2}{n_i}$, $\mu_1 = \frac{t_1^2}{36}$, ...$\mu_4 = \frac{t_4^2}{36}$. ...(10)

d) Total des carrés moyens μ_0: $\mu_{0_i} = \Sigma\mu_i = \mu_1 + \mu_2 + \mu_3 + \mu_4$. ...(11)

e) Carré moyen général: $M = \frac{t_0^2}{N} = \frac{(t_1 + t_2 + t_3 + t_4)^2}{144}$ ...(12)

$$\sum_{ij}(Y_{ij} - \overline{Y})^2 = \sum_{ij} Y_{ij}^2 - N\overline{Y}^2 = S - N(\frac{t_0}{N})^2$$

$$= S - \frac{t_0^2}{N} = S - M = [A] \text{ (Sortie)} \qquad ...(13)$$

Ceci s'applique aussi à (6b) excepté que le symbole M est remplacé par μ_0.

$$\sum_{ij}(Y_{ij} - \overline{Y}_i)^2 = S - \mu_0 = [B] \text{ (Sortie)}$$

Enfin, la relation (6c) conduit à

$$\sum_i n_i(\overline{Y} - \overline{Y})^2 = \mu_0 - M = [C] \quad \text{(Sortie)}$$

Les valeurs numériques dans les relations ci dessus sont:

$\mu_1 = 4.71$, $\mu_2 = 2.89$, $\mu_3 = 2.49$, $\mu_4 = 25.23$, $\mu_0 = 35.32$.

$S = 58.23$. $M = 27.41$.

$[A] = 30.82$, $[B] = 22.91$, $[C] = 7.91$.

Table des formules finales:

$$s^2 = \frac{1}{N-1}\sum_{ij}(Y_{ij} - \overline{Y})^2 = \frac{[A]}{N-1} = \frac{30.82}{143} = 0.22.$$

$$s'^2 = \frac{1}{N-h}\sum_{ij}(Y_{ij} - \overline{Y}_i)^2 = \frac{[B]}{N-h} = \frac{22.91}{140} = 0.16.$$

$$s''^2 = \frac{1}{h-1} \sum_i n_i(\overline{Y}_i - \overline{Y})^2 = \frac{[C]}{h-1} = \frac{7.91}{3} = 2.64.$$

L'index de Fisher F est:

$$F = \frac{s''^2}{s'^2} = \frac{2.64}{0.16} = \underline{16.2}$$

D'après le sens de l'index F, ceci montre que la valeur de F est très significative car, d'après les tables, le "point 5%" est égal à 2.4 (pour $\nu_1 = 140$, $\nu_2 = 3$). Cette valeur est donc considérée comme étant très défavorable à l'hypothèse que l'échantillon est homogène par rapport au critère "manufacturier".

Les calculs ci-dessus sont aussi résumés dans la partie supérieure du Tableau IB. La partie inférieure montre les résultats des calculs concernant le critère "direction". Ici, F = 1.56 ce qui est très petit par rapport au "point 1%" (3.4): l'hypothèse que l'échantillon est homogène par rapport à l'orientation géographique a donc beaucoup de chance d'être correcte.

L'examen des huit tables (I à VIII) confirme la conclusion tirée de l'année 1972. Le lecteur est autorisé à se demander si une conclusion similaire pourrait être tirée d'un examen direct des données numériques contenues dans les tables. Considérons les valeurs des totaux

	t_1	t_2	t_3	t_4
Manuf.	13.02	10.20	9.46	30.14
Direct.	13.62	14.09	18.35	16.76

Le total t_4 (Manuf), égal à 30.14, est très différent de ceux qui sont situés sur la même ligne horizontale; on ne peut pas en dire autant de la seconde ligne horizontale (Directions). Ainsi, en classant les compteurs d'après le critère "manufacturier" on trouve, parmi les quatres classes de 36 compteurs, une classe où $t_4 = 30.14$. Ce total est donc fortement différent des trois autres totaux. Par contre, la classification d'après les "directions" conduit

à quatre classes (de 36 chacun) qui sont très semblables l'une à l'autre (seconde ligne du tableau).

Les conclusions d'ordre général auxquelles conduit l'analyse de la variance sont amalgamées avec celles qui résultent de l'analyse effectuée dans la section suivante.

4. AUTRES RECHERCHES STATISTIQUES

Dans l'analyse de la variance, un échantillon est représenteé par sa moyenne et sa variance. On ne tient pas compte de la distibution intérieure de ses fréquences. Il faut, tout de même, vérifier que cette distribution n'est pas trop loin de la normalité (courbe en cloche): le fondement même des tables de Fisher est basé sur le postulat que cette distribution est normale ou presque normale. Le lecteur peut aisément vérifier que tel est le cas en examinant les Tables I à VIII. Dans la présente section, c'est la structure intérieure des échantillons qui va jouer un rôle important, en particulier en relation avec la structure des "mélanges" des sous-échantillons dans lesquels la population totale peut être subdivisée.

Avant de commencer cette analyse, il faut d'abord souligner le fait qu'un échantillon de 144 éléments est, à vrai dire, un "petit" échantillon. Si on le subdivise encore, les parties deviennent si petites que les calculs statistiques conduisent à des conclusions fort incertaines. Afin de pouvoir manipuler des échantillons plus considérables, les observations faites pendant les huit années consécutives ont été jointes afin de ne former qu'un seul échantillon. En plus il a été décidé de recalculer toutes les EER mais sur la base demi-annuelle. Ceci a permis de constituer une population totale de 2252 éléments qui, à cause des éliminations, est un peu plus petite que le maximum théorique de 2 X 144 X 8 = 2304.

La perte de 52 EET est due à l'élimination de sept compteurs, marqués dans les tableaux par des astériques. La population totale se compose donc de:

a) Huit "tables de mai" (analogues à IA) contenant 1128 EET calculées pour les périodes de six mois: novembre à mai.

b) Huit "tables de novembre", contenant 1124 EET calculées pour les périodes de six mois: mai à novembre.

Les erreurs relevées en mai proviennent de la période qui contient les mois d'hiver de sorte que cette période est appelée "saison froide" (en anglais "cold season") et est désignée par l'indice inférieur "c". La période de mai à novembre contient les mois d'été et est appelée "saison chaude" (en anglais "warm season") et est désignée par l'indice inférieur "w". Les valeurs individuelles des 2252 erreurs d'enregistrement totales ne sont pas citées dans le présent mémo; elles sont cependant donnée sous une forme classifiée dans la Fig 1 et Fig 2: dans la Fig 1 le total (2252) est considéré comme un "mélange" de deux sous-échantillons, partie "saison froide" et partie "saison chaude"; dans la Fig 2, le total est un mélange de la partie "direction froide" avec la partie "direction chaude". Les statistiques de la subdivision en "saisons" sont les suivantes:

Saison froide:	$N_c = 1128$,	$\overline{m}_c = 0.57$,	écart type	$s_c = 0.52$.
Saison chaude:	$N_w = 1124$,	$\overline{m}_w = 0.22$,	"	$s_w = 0.59$.
Total:	$N_t = 2252$,	$\overline{m}_t = 0.39$,	"	$s_t = 0.58$.

En ce qui concerne "direction", le nord et l'est (ensemble) constituent la direction froide, tandis que sud et ouest (ensemble constituent la direction chaude:

Direction froide (N + E) :	$N_c = 1129$,	$\overline{m}_c = 0.45$,	écart type	$s_c = 0.54$.
Direction chaude (S + W) :	$N_w = 1123$,	$\overline{m}_w = 0.33$,	" "	$s_w = 0.62$.
Total:	$N_t = 2252$,	$\overline{m}_t = 0.39$,	" "	$s_t = 0.58$.

Avant de tirer les conclusions des Fig 1 et 2, il faut mentionner de nouveau qu'au cours des huit années d'expérience sept compteurs ont été éliminés pour défauts de fonctionnement. La manière dont les relevés ont été

accomplis ne permet pas de se rendre compte quand les compteurs éliminés ont commencé à mal fonctionner. Dans la moitié des cas, le compteur se trouve simplement et définitivement stoppé. Dans d'autres cas, le chiffre indiqué par le compteur n'avait aucun sens. Seulement un manufacturier expériementé et un expert en mesures électriques peut évaluer la qualité des compteurs qui ont été testés de la manière qui est décrite dans le présent rapport. Si le nombre des cas où les compteurs ont mal fonctionné n'est pas considéré comme très significatif, on peut dire que les compteurs ont fonctionné d'une manière satisfaisante. Un compteur semble être, dans les limites de la tolérance admise, un instrument auquel on peut se fier malgré le dur climat de la région d'Ottawa.

Au sujet de la distribution des fréquences on peut faire la remarque générale suivante; les cinq diagrammes observés sont raisonablement réguliers, symétriques et en forme de cloche mais quatre d'entre eux ne sont pas vraiment bien représentés par la courbe de Gauss. Ils sont leptokurtiques avec différents degrés de leptokurtose.* Dans la théorie moderne la présence de la leptokurtose dans les diagrammes des erreurs est traitée simplement comme un caractère additionnel sans influence spécifique sur la qualité des observations.

L'impression générale (suggérée par un examen superficiel) produite par la Fig 1 est celle de la "dispersion" tandis que celle produite par la Fig 2 est celle de la "cohésion". Ceci suggère l'idée que le critère "saison" agit sur le comportement des compteurs plus fortement que le critère "direction". Un examen plus profond montre que dans les deux figures les moyennes avec l'index w sont plus proches de 0 (zéro) que les moyennes avec l'indice c; il semble par conséquent que dans les conditions plus chaudes (sud et été) les erreurs d'enregistrement ont la tendance de décroître tandis que

---

* Note: un diagramme de fréquences est dit "leptokurtique" (terme de K. Pearson) si, dans sa portion centrale, les fréquences observées (représentées par les points , dans les deux totaux) sont plus grandes que les correspondantes fréquences indiquées par la courbe de Gauss (courbe qui représente le diagramme). Dans la courbe qui représente le total (2252) il y a neuf de ces points. Cet excès au centre est compensé par un abaissement correspondant des points situés des deux côtés de la partie centrale.

dans les conditions plus froides la tendance est dans le sens contraire. En plus, les moyennes des sous-échantillons sont plus éloignées l'une de l'autre dans la Fig 1 que dans la Fig 2. Et la différence dans Fig 1 est loin d'être non significative:

$$\overline{m}_c - \overline{m}_w = 0.35.$$

Comme s_c et s_w sont tous les deux de l'ordre de 0.55, la différence est donc de l'ordre de 0.6 écart type. Elle n'est donc pas négligeable et on peut dire: il est à l'avantage du consommateur de placer le compteur sur le côté sud du bâtiment et de consommer de l'électricité surtout pendant la période chaude. Cependant, son gain ne sera que "probable" et, de toute façon, seulement léger!

NOTE

L'échantillon d'Erreurs d'enregistrement totales (EET), analysé ci-dessus, est aussi utilisé pour illustrer la théorie des "Mélange", en particulier quand les composantes sont du type normal modulé. L'article sur ce sujet est publié dans le Bulletin Géodésique 53, No. 1, 1983, p. 62 et peut être obtenu des auteurs (MR ou EG) en s'addressant à NRC, Division de Physique, Ottawa, Ontario, Canada K1A 0R6.

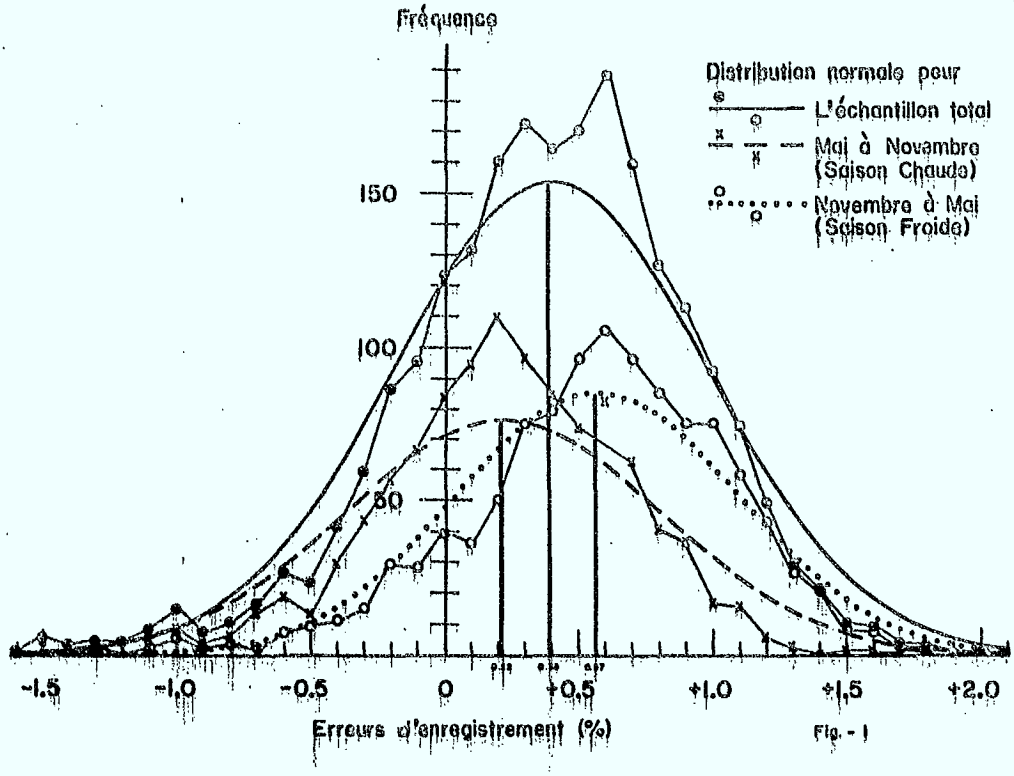
Fréquence
Distribution normale pour
L'échantillon total
Mai à Novembre (Saison Chaude)
Novembre à Mai (Saison Froide)
150
100
50
-1.5
-1.0
-0.5
0
+0.5
+1.0
+1.5
+2.0
Erreurs d'enregistrement (%)

Fig. - 1

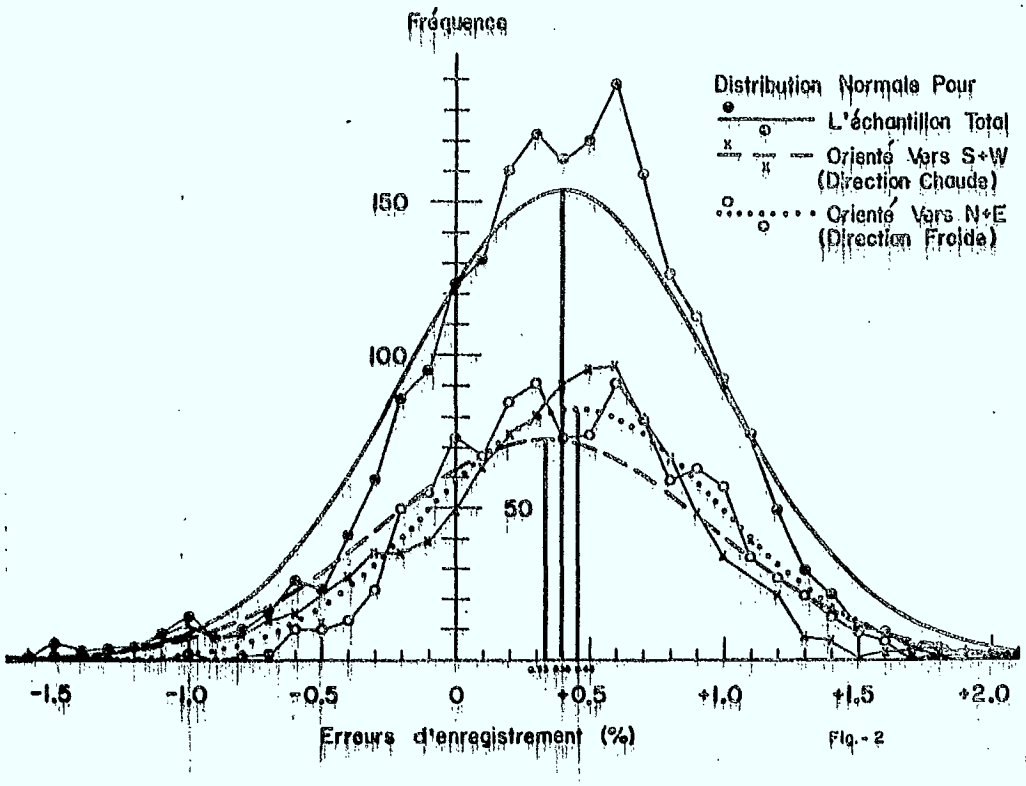
Fréquence
Distribution Normale Pour
L'échantillon Total
Orienté Vers S+W (Direction Chaude)
Orienté Vers N+E (Direction Froide)
150
100
50
-1.5
-1.0
-0.5
0
+0.5
+1.0
+1.5
+2.0
Erreurs d'enregistrement (%)

Fig. - 2

Appendice I

Cet appendice contient les tables des données observées et les "sorties" (outputs) pour chacune des années 1972 - 1979. Dans ces tables les données observées sont présentées comme dans la Table IA.

Le lecteur remarquera que les résultats numériques, quoique un peu différents d'une année à l'autre, conduisent aux mêmes conclusions que celles présentées à la fin de la Section 2.

Comme les tables on été faites par ordinateur, le texte n'est disponible qu'en anglais. Voice la liste des termes utilisés;

Date of Data: Date de relevé des indications
Criterion-Manufacturer: Critère-Manufacturier
Criterion-Direction: Critère-Direction
Square: Carré
East: Est
West: Ouest
North: Nord
South: Sud

Date of Data is 1972 NOV

	GE	FP	W	S
SOUTH	-.01	.72	.62	1.24
	-.20	.56	.60	1.46
	.12	-.95	.09	.72
	.39	.83	.76	1.15
	.07	.24	.14	.80
	-.17	-.40	-.72	.62
	.16	.44	.73	.64
	-.01	-.91	1.06	1.02
	.54	-.49	.87	.95
EAST	.31	-.13	.05	.70
	.30	.27	-.29	.64
	.11	.43	-.21	1.20
	.65	.34	-.38	.57
	.36	.07	1.17	.89
	.58	-.07	.60	.75
	.56	.40	-.14	.92
	.65	.53	-.18	.87
	.47	.05	.24	.81
NORTH	.32	.26	.39	.51
	.11	.69	-.07	.60
	.07	.98	.68	.98
	.65	.55	-.18	.44
	.72	.88	.59	.72
	.69	.30	-.03	.25
	.68	.62	.98	.92
	.62	.31	.31	.81
	.35	.42	.03	1.20
WEST	.22	.50	.27	.27
	.46	.66	.40	2.50
	.36	.30	.03	.59
	.05	.77	.28	.45
	.42	-.03	.15	.47
	.99	.50	-.43	.56
	.40	.04	.56	1.65
	.40	.20	.57	.88
	.63	.32	-.08	.45
	t_1	t_2	t_3	t_4

Date of Data is 1972 NOV

TABLE I-B

Criterion- MANUFACTURER

S= 58.23 N= 144.00
M= 27.41 h=4

Manufacturer	GE	FP	W	S
Totals (t)	13.02	10.20	9.46	30.14
Totals squared	169.52	104.04	89.49	908.42
Mu	4.71	2.89	2.49	25.23

t zero = 62.82 N-h= 140.00
t zero squr.= 3946.35 h-1=3
Mu zero = 35.32
A = 30.82
B = 22.91
C = 7.91
s' square = .16
S'' square = 2.64
F = 16.12
(v' 3, v''= 140.00)

Criterion- DIRECTION

S= 58.23 N= 144.00
M= 27.41 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	13.62	14.09	18.35	16.76
Totals squared	185.50	198.53	336.72	280.90
Mu	5.15	5.51	9.35	7.80

t zero = 62.82 N-h= 140.00
t zero squr.= 3946.35 h-1=3
Mu zero = 27.82
A = 30.82
B = 30.40
C = .42
s' square = .22
s'' square = .14
F = 1.56
(v'=3 v''= 140.00)

Date of Data is 1973 NOV

.18	.93	1.46	1.39
-.09	.64	.82	.87
.18	-.52	.25	.94
.55	.88	.89	1.00
.14	.48	.18	1.10
-.03	-.37	-.53	.68
.36	.53	.66	.93
.07	-.88	1.16	1.07
.68	-.56	.90	1.00

.37	.01	.23	1.04
.36	.51	-.10	.72
.21	.77	-.11	1.21
.72	.43	-.19	.67
.43	.12	1.27	.94
0.00	.96	.64	.80
.63	.56	-.03	.94
.73	.62	-.16	1.12
.51	.04	.33	.81

.31	.35	.41	.53
.13	1.14	-.04	.56
.04	1.09	.67	1.04
.65	.56	.02	.52
.78	.91	.60	.73
.71	.43	-.07	.27
.70	.79	.99	1.01
.59	.38	.31	1.03
.38	.35	.21	1.29

.23	.59	.35	.39
.50	.07	.57	-1.15
.34	1.40	.08	.65
.07	.85	.35	.50
-.06	.05	.20	.52
1.04	.59	-.32	.55
.47	.05	.61	-.07
.47	.19	.67	.95
.71	.43	-.01	.35

Date of Data is 1973 NOV

TABLE II

Criterion- MANUFACTURER

S= 63.06 N= 144.00
M= 33.64 h=4

Manufacturer	GE	FP	W	S
Totals (t)	14.06	15.37	13.27	26.90
Totals squared	197.68	236.24	176.09	723.61
Mu	5.49	6.56	4.89	20.10

t zero	=	69.60	N-h= 140.00
t zero squr.	=	4844.16	h-1=3
Mu zero	=	37.05	
A	=	29.42	
B	=	26.01	
C	=	3.41	
s' square	=	.19	
S'' square	=	1.14	
F	=	6.11	
(v' 3, v''=		140.00)	

Criterion- DIRECTION

S= 63.06 N= 144.00
M= 33.64 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	17.94	18.11	20.37	13.18
Totals squared	321.84	327.97	414.94	173.71
Mu	8.94	9.11	11.53	4.83

t zero	=	69.60	N-h= 140.00
t zero squr.	=	4844.16	h-1=3
Mu zero	=	34.40	
A	=	29.42	
B	=	28.66	
C	=	.76	
s' square	=	.20	
s'' square	=	.25	
F	=	.81	
(v'=3 v''=		140.00)	

Date of Data is 1974 NOV

-.10	.60	.88	.84
-.27	.29	.63	.56
.03	-1.26	0.00	.56
.35	.42	.64	.81
-.13	.40	.06	.80
-.48	-.81	1.02	.47
.05	.34	.51	.56
-.61	-1.20	.08	.85
.44	-.76	.63	.77
.04	-.31	.15	.99
.19	.29	.01	.55
.06	.50	-.07	1.09
.57	.23	.28	.43
.30	-.02	1.38	.79
-.17	.83	.54	.63
.52	.38	-.13	.90
.57	.46	-.26	.87
.60	-1.01	.25	.70
.19	.09	.28	.39
0.00	.81	-.18	.38
-.06	.86	.46	.86
.53	.50	-.12	.39
.61	.76	.61	.60
.60	.19	-.16	.13
.50	.61	1.00	.84
.42	.15	1.11	.95
.23	.43	.12	1.22
.06	.29	.26	.22
.32	1.22	.44	.50
.22	1.03	-.18	.43
-.10	.65	.16	.29
***	-.23	0.00	.30
.87	.26	-.46	.39
.25	-.25	1.46	.15
.21	-.08	.49	.72
.50	.33	.33	.37

Date of Data is 1974 NOV

TABLE III

Criterion- MANUFACTURER

S= 47.27 N= 143.00
M= 16.67 h=4

Manufacturer	GE	FP	W	S
Totals (t)	7.31	6.99	12.22	22.30
Totals squared	53.44	48.86	149.33	497.29
Mu	1.53	1.36	4.15	13.81

t zero	=	48.82	N-h= 139.00
t zero squr.	=	2383.39	h-1=3
Mu zero	=	20.85	
A	=	30.60	
B	=	26.42	
C	=	4.18	
s' square	=	.19	
S'' square	=	1.39	
F	=	7.33	
(v' 3, v''=		139.00)	

Criterion- DIRECTION

S= 47.27 N= 143.00
M= 16.67 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	7.97	13.13	16.30	11.42
Totals squared	63.52	172.40	265.69	130.42
Mu	1.76	4.79	7.38	3.73

t zero	=	48.82	N-h= 139.00
t zero squr.	=	2383.39	h-1=3
Mu zero	=	17.66	
A	=	30.60	
B	=	29.61	
C	=	.99	
s' square	=	.21	
s'' square	=	.33	
F	=	.64	
(v'=3 v''=		139.00)	

Date of Data is 1975 NOV

-.03	.76	.96	.70
-.19	.21	.72	.62
.06	-.16	.05	.59
.42	.37	.79	.81
-.07	.58	***	1.02
-.42	-.72	.34	.47
.08	.29	.63	.62
-.05	-.27	1.02	.84
.49	-.94	.62	.78

-.51	.29	.35	1.16
.24	.42	.11	.61
.11	.50	.09	1.09
.55	.25	-.14	.46
.31	-.01	1.55	.80
.37	.83	.65	.67
.49	.41	-.02	.96
.53	.49	-.11	.89
.44	-.14	.19	.83

.13	.05	.30	.39
-.06	.74	-.10	.31
-.08	.78	.44	.83
.40	.50	-.01	.36
.49	.79	.67	.61
.47	.13	.15	.09
.39	.59	-1.00	.81
.36	.14	.26	.90
.21	.51	.24	1.16

.02	.28	.41	.17
.28	.37	.53	.53
.18	-.83	-.11	.37
-.18	.46	.26	.25
***	-.29	.03	.3
.72	.05	-.26	.33
.22	-.36	.44	.60
.10	-.15	.59	.66
.47	-.02	***	.51

Date of Data is 1975 NOV

TABLE IV

Criterion- MANUFACTURER

S= 40.12 N= 141.00
M= 16.10 h=4

Manufacturer	GE	FP	W	S
Totals (t)	6.94	6.90	10.64	23.16
Totals squared	48.16	47.61	113.21	536.39
Mu	1.38	1.32	3.33	14.90

t zero	=	47.64	N-h= 137.00
t zero squr.	=	2269.57	h-1=3
Mu zero	=	20.93	
A	=	24.03	
B	=	19.19	
C	=	4.83	
s' square	=	.14	
S'' square	=	1.61	
F	=	11.50	
(v' 3, v''=		137.00)	

Criterion- DIRECTION

S= 40.12 N= 141.00
M= 16.10 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	11.99	15.71	12.95	6.99
Totals squared	143.76	246.80	167.70	48.86
Mu	4.11	6.86	4.66	1.44

t zero	=	47.64	N-h= 137.00
t zero squr.	=	2269.57	h-1=3
Mu zero	=	17.06	
A	=	24.03	
B	=	23.06	
C	=	.96	
s' square	=	.17	
s'' square	=	.3	
F	=	.52	
(v'=3 v''=		137.00)	

Date of Data is 1976 NOV

.19	1.05	1.23	.85
-.05	.37	.74	.69
.18	-.59	.23	.75
.56	.61	.86	.87
-.01	.80	-.08	1.07
-.21	-.09	.33	.61
.32	.35	.61	.79
.02	-1.16	1.05	.95
.64	-.95	.61	.92

***	-.15	.67	1.18
.36	.55	.02	.70
.24	.68	.11	1.07
.69	.36	-.18	.52
.37	.11	1.50	.86
***	.91	.72	.78
.61	.45	-.05	1.10
.66	.44	-.16	1.11
.57	-.16	.21	1.21

.29	.25	.46	.49
.09	.95	.05	.46
.03	.98	.67	1.01
.53	.49	.05	.45
.68	.93	.74	.68
.66	.31	.26	.21
.60	.72	.82	1.00
.54	.21	.39	1.15
.28	.60	.30	1.31

.14	.46	.57	.37
.45	.76	.70	.57
.27	1.33	.08	.62
.01	.70	.29	.44
***	-.21	.36	.52
.97	.33	-.23	.55
.36	-.30	.35	.86
.34	-.17	.60	.88
.67	.13	***	.82

Date of Data is 1976 NOV

TABLE V

Criterion- MANUFACTURER

S= 58.24 N= 140.00
M= 32.45 h=4

Manufacturer	GE	FP	W	S
Totals (t)	12.05	12.05	14.88	28.42
Totals squared	145.20	145.20	221.41	807.70
Mu	4.40	4.03	6.33	22.44

t zero	=	67.40	N-h= 136.00
t zero squr.	=	4542.76	h-1=3
Mu zero	=	37.20	
A	=	25.79	
B	=	21.05	
C	=	4.75	
s' square	=	.15	
S'' square	=	1.58	
F	=	10.23	
(v' 3, v''=		136.00)	

Criterion- DIRECTION

S= 58.24 N= 140.00
M= 32.45 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	15.11	18.06	19.64	14.59
Totals squared	228.31	326.16	385.73	212.87
Mu	6.34	9.59	10.71	6.26

t zero	=	67.40	N-h= 136.00
t zero squr.	=	4542.76	h-1=3
Mu zero	=	32.91	
A	=	25.79	
B	=	25.33	
C	=	.46	
s' square	=	.19	
s'' square	=	.15	
F	=	1.21	
(v'=3 v''=		136.00)	

Date of Data is 1977 NOV

.19	1.12	1.31	.86
-.05	.33	.84	.66
.04	-.54	.27	.78
.55	.58	.90	.89
-.03	.87	-.04	.89
-.17	0.00	.39	.64
.28	.38	.72	.82
.03	-1.20	1.11	.96
.64	-.98	.77	.91
***	-.16	.78	1.22
.30	.55	.05	.73
.22	.66	.23	1.31
.69	.40	-.14	.50
.54	-.12	1.55	.93
***	.95	.76	.81
.59	.44	.02	1.13
.65	.42	-.12	1.16
.59	-.20	.25	1.39
.30	.24	.54	.51
.10	.97	.13	.48
.04	.99	.70	1.09
***	.50	.13	.46
.66	1.00	.79	.65
.66	.32	.30	.24
.61	.77	1.22	1.02
.52	.15	.47	1.20
.30	.62	.30	1.36
.05	.50	.60	.38
.45	.75	.78	.59
.28	1.42	.19	.66
.01	.73	.42	.47
***	-.20	.49	.59
.97	.35	0.00	.58
.36	-.34	.41	.89
.33	-.18	.67	.85
.69	.13	***	.81

Date of Data is 1977 NOV

TABLE VI

Criterion- MANUFACTURER

S= 63.97 N= 139.00
M= 36.08 h=4

Manufacturer	GE	FP	W	S
Totals (t)	11.39	12.22	17.79	29.42
Totals squared	129.73	149.33	316.48	865.54
Mu	4.05	4.15	9.04	24.04

t zero	=	70.82	N-h= 135.00
t zero squr.	=	5015.47	h-1=3
Mu zero	=	41.29	
A	=	27.89	
B	=	22.68	
C	=	5.20	
s' square	=	.17	
S'' square	=	1.73	
F	=	10.33	
(v' 3, v''=		135.00)	

Criterion- DIRECTION

S= 63.97 N= 139.00
M= 36.08 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	15.72	19.08	20.34	15.68
Totals squared	247.12	364.05	413.72	245.86
Mu	6.86	10.71	11.82	7.23

t zero	=	70.82	N-h= 135.00
t zero squr.	=	5015.47	h-1=3
Mu zero	=	36.62	
A	=	27.89	
B	=	27.35	
C	=	.54	
s' square	=	.20	
s'' square	=	.18	
F	=	1.12	
(v'=3 v''=		135.00)	

Date of Data is 1978 NOV

-.41	.60	.82	.41
-.52	-.16	.47	.27
***	-1.35	-.18	.36
.02	.12	.42	.54
-.49	.43	-.18	.43
-.75	-.64	.04	.24
-.30	0.00	.44	.36
-.41	-1.77	.73	.53
.16	-1.36	.48	.46

***	-.79	.43	.73
-.10	-.01	-.14	.39
-.23	.04	-.05	.8
.29	-.04	-.48	.12
.06	-.28	1.31	.67
***	.60	.35	.46
.22	-.03	-.29	.79
.21	.09	-.39	.80
.19	-.46	-.19	1.06

***	-.17	.27	.22
-.23	.52	-.14	.21
-.28	.60	.42	.79
***	.18	-.08	.20
.33	.71	.43	.43
.36	-.04	-.03	-.01
.22	.42	.75	.71
.19	-.23	.19	.91
-.04	.12	.06	1.05

-.13	.10	.31	.06
.06	.27	.43	.20
-.06	.93	-.15	.29
-.41	.32	.18	.15
***	-.59	.12	.30
.62	-.01	.32	.20
.01	-.75	.12	.57
-.06	-.54	.37	.50
.36	-.32	***	.53

Date of Data is 1978 NOV

TABLE VII

Criterion- MANUFACTURER

S= 32.69 N= 137.00
M= 2.73 h=4

Manufacturer	GE	FP	W	S
Totals (t)	-1.12	-3.49	7.16	16.79
Totals squared	1.25	12.18	51.27	281.90
Mu	.04	.34	1.46	7.83

t zero	=	19.34	N-h= 133.00
t zero squr.	=	374.04	h-1=3
Mu zero	=	9.68	
A	=	29.9	
B	=	23.01	
C	=	6.95	
s' square	=	.17	
S'' square	=	2.32	
F	=	13.38	
(v' 3, v''=		133.00)	

Criterion- DIRECTION

S= 32.69 N= 137.00
M= 2.73 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	-.19	6.19	9.04	4.30
Totals squared	.04	38.32	81.72	18.49
Mu	.00	1.13	2.40	.54

t zero	=	19.34	N-h= 133.00
t zero squr.	=	374.04	h-1=3
Mu zero	=	4.08	
A	=	29.96	
B	=	28.61	
C	=	1.35	
s' square	=	.22	
s'' square	=	.45	
F	=	.48	
(v'=3 v''=		133.00)	

Date of Data is 1979 NOV

.09	1.16	1.25	.71
-.14	.21	.80	.56
***	-.69	.27	.70
.54	.45	.92	.84
-.11	.83	0.00	.77
-.96	-.03	.40	.57
.18	.38	.75	.79
.02	-1.46	1.09	.8
.63	-1.15	.60	.85
***	-.48	.87	1.01
.37	.53	.22	.73
.21	.63	.34	1.19
.67	.31	-.12	.44
.52	.07	1.64	1.02
***	1.00	.82	.78
.60	.36	.05	1.16
.63	.39	-.10	1.14
.59	-.28	.26	1.38
.34	.22	.54	.58
.10	.94	.15	.47
.03	.95	.73	1.11
***	.49	.08	.43
.63	1.08	.81	.69
.64	.25	.17	.25
.56	.73	1.09	.98
.51	.04	.50	1.21
.28	.58	.32	1.32
.12	.54	.61	.34
.36	.71	.90	.51
.25	1.38	.19	.61
-.06	.66	.50	.41
***	-.28	.38	.58
.93	.29	-.03	.51
.36	-.48	.43	.89
.28	-.24	.70	.85
.69	-.03	***	.85

Date of Data is 1979 NOV

TABLE VIII

Criterion- MANUFACTURER

S= 63.47 N= 138.00
M= 31.70 h=4

Manufacturer	GE	FP	W	S
Totals (t)	9.86	10.06	18.13	28.09
Totals squared	97.22	101.20	328.70	789.05
Mu	3.14	2.81	9.39	21.92

t zero	=	66.14	N-h= 134.00
t zero squr.	=	4374.50	h-1=3
Mu zero	=	37.26	
A	=	31.77	
B	=	26.21	
C	=	5.56	
s' square	=	.20	
S'' square	=	1.85	
F	=	9.47	
(v' 3, v''=		134.00)	

Criterion- DIRECTION

S= 63.47 N= 138.00
M= 31.70 h=4

Direction	SOUTH	EAST	NORTH	WEST
Totals (t)	12.69	18.95	19.80	14.70
Totals squared	161.04	359.10	392.04	216.09
Mu	4.60	10.56	11.20	6.36

t zero	=	66.14	N-h= 134.00
t zero squr.	=	4374.50	h-1=3
Mu zero	=	32.72	
A	=	31.77	
B	=	30.75	
C	=	1.02	
s' square	=	.23	
s'' square	=	.34	
F	=	.67	
(v'=3 v''=		134.00)	

Appendice II

Note sur la correlation entre l'erreur d'enregistrement totale (EET) et l'erreur du disque dans le projet "Compteurs sur le toit".

La présente Note a été préparée en coopération avec B. Fulsom, chef du Laboratoire d'Electricité et J. Samuel, étudiant, engagé temporaire pendant l'été 1982.

LES ERREURS DU DISQUE (ED) DANS LE PROJET "COMPTEURS SUR LE TOIT"

1. Introduction

La principale caractéristique des "erreurs d'enregistrement totales" (EET) est qu'elles demandent que les mesures consécutives soient séparées l'une de l'autre par des intervalles de temps considérables, au moins de plusieurs mois. Ceci entraîne la consommation de grandes quantités d'énergie électrique et le maintien, pendant de longues périodes, de la circuiterie et des installations appropriées. L'une des raisons pour cet état de choses est qu'un compteur ne fournit une lecture bien définie que s'il indique une quantité relativement considérable de watt-heures. Il va de soi que pour le manufacturier et pour l'usager qui consomme une grande quantité d'énergie électrique, les EET sont d'une importance capitale et doivent être surveillées.

Dans la pratique, cependant, l'une des opérations courantes est la détermination de l'erreur d'enregistrement sur une période très courte, p.ex. de l'ordre de quelques minutes. Il est évident que, dans ce cas, une détermination ne peut pas être aussi sûre que celle de l'EET mais, pour des raisons évidentes, elle doit aussi pouvoir être effectuée. Lorsqu'un usager trouve que son compteur demande à être vérifié, ce compteur est acheminé vers un laboratoire dans lequel l'erreur du disque (ED) est aussitôt déterminée.

La propriété fondamentale d'un compteur à induction est qu'un tour de son disque tournant correspond à un nombre constant d'unités d'énergie (joule ou watt-heure). Cette constante, désignée par le symbole "K_h", est un peu

perturbée par les conditions dans lesquelles la mesure est effectuée. La détermination de EET bénéficie de l'effet égalisateur dû au grand nombre de tous qui ont lieu pendant de long intervals de temps.

Un autre point important qui doit être mentionné concerne le courant utilisé pendant les déterminations du ED. L'influence de son intensité sur K_h n'est pas simple, et en tout cas pas linéaire. Dans la pratique courante, K_h est déterminé avec des courants égaux à 2.5 A et 25A. Tous les manufacturiers s'efforcent à produire des compteurs qui possèdent des valeurs prescrites pour les constantes du disque et un compteur est considéré (par le manufacturier) comme acceptable dans les installations domestiques si son erreur du disque (ED) n'est pas supérieure à 1%. En fait, cette limite est plus étroite que la tolérance légale (2%) car la vérification des noueaux compteurs est basée sur la méthode "par échantillonage" et non sur la vérification de chaque instrument individuel.

La référence pour les ED est un instrument appelé "Étalon de transfert". Il produit des battements, le temps entre deux battements consécutifs corresondant à une quantité constante d'énergie électrique. Il est utilisé dans le laboratoire de Métrologie légale pour déterminer la constante K_h du "Compteur de Référence". La constante K_h est ajustée de façon qu'un watt-heure corresponde à trois révolutions du disque. En d'autres termes, K_h = 1/3 et, légalement, un compteur est acceptable si son erreur ne dépasse pas 2%.

2. Travail expérimental et calcul des corrélations

Il est évident que même sur un compteur de haute qualité la relation linéaire entre ses EET et ED ne sera pas absolument parfaite. Dans le langage statistique nous pouvons tout de même prévoir qu'entre ces deux sortes d'erreur la corrélation sera, sinon "parfaite", en tout cas très étroite. Au

cours de l'expérience décrite ci-dessus, qui a duré huit années consécutives, un certain nombre de déterminations des erreurs du disque ont été effectuées. Toutefois, pour des raisons pratiques, l'ensemble des ED ainsi obtenues n'a pas pu être traité dans sa totalité; seulement trois périodes ont pu être calculées et statistiquement analysées. La première et la troisième période s'étendent sur une année chacune; la seconde période occupe deux années consécutives. L'erreur du disque pour une période est la moyenne des observations effectuées au commencement, au milieu et à la fin de la période. Cette moyenne est désignée par le symbole X_i de sorte que dans chaque période, à une valeur Y_i correspond une valeur X_i. Les tables des valeurs individuelles X_i ne sont pas citées dans le présent memorandum et les données concernant les périodes sont résumées comme suit.

Première période	1 Nov 1971 à 11 Oct 1972,	N = 144.
Deuxième "	15 Oct 1974 à 15 Oct 1976,	N = 139.
Troisième "	12 Oct 1975 à 15 Oct 1979,	N = 138.

La suite des opérations conduisant à la valeur de l'index de corrélation r est:

A) Calcul des moyennes $\overline{Y}$ and $\overline{X}$:

$$\overline{Y} = \frac{1}{N}\sum_{i-1}^{N}\overline{Y}_i \; ; \; \overline{X} = \frac{1}{N}\sum_{i-1}^{N}X_i$$

B) Calcul des deviations y_i et x_i:

$$y_i = Y_i - \overline{Y} \; ; \; x_i = X_i - \overline{X}.$$

C) Calcul des variances $s_y{}^2$ et $s_x{}^2$:

$$s_y{}^2 = \frac{1}{N}\sum_{i=1}^{N} y_i{}^2 \; ; \; s_x{}^2 = \frac{1}{N}\sum_{i=1}^{N} x_i{}^2.$$

D) Calcul de la covariance:

$$\mu_{11} = \frac{1}{N}\sum_{i=1}^{N} y_i x_i.$$

E) L'index de corrélation r, qui résulte des valeurs ci dessus est:

$$\mu r = \frac{11}{s_y s_x}$$

Pour se rendre compte à quel point ce "r" est significatif il faut se rapporter à la valeur correspondante du variate t de Student. La relation entre t et r est:

$$t = \frac{r}{1-r^2} (N - 2)$$

Une forte correlation (à laquelle on doit s'attendre) doit conduire à des valeurs de r qui sont proches de l'unité; ceci, à son tour, conduit aux grandes valeurs de t. Par exemple, pour r = 0.5 (N = 144), t = 6.9. Ce t est donc très grand et cela indique que la corrélation est très forte.

Les valeurs numériques des diverses quantités mentionnées ci dessus sont présentées dans la table suivante:

PREMIÈRE PÉRIODE (1972)	DEUXIÈME PÉRIODE (1976)	TROISIÈME PÉRIODE
μ_{11} = 0.170	0.180	0.250
s_y^2 = 0.215	0.185	0.236
s_y = 0.464	0.430	0.486
s_x^2 = 0.301	0.324	0.361
s_x = 0.549	0.569	0.601
r = 0.667	0.736	0.855

Il faut remarquer que toutes les valeurs de r sont supérieures à 0.5. Pour r = 0.667, t est plus grand que 10. Les corrélations entre X et Y dans les trois périodes sont donc très fortes. La conclusion générale de

l'expérience et de l'analyse statistique est qu'il n'y a pas de raison d'invalider l'usage traditionnel de l'erreur du disque pour vérifier l'exactitude d'un compteur électrique.

Si les circonstances le permettent (c.à.d. si l'on dispose d'un personnel approprié) il sera peut-être possible d'appliquer l'analyse de la variance directement aux erreurs du disque. Il serait intéressant de voir comment cette analyse s'accorde, dans ses conclusions, avec les résultats de l'analyse de la variance appliquée aux erreurs d'enregistrement totales (EET).

CACC / CCAC
18669

LKC
TK 301 .R614 1984
Romanowski, M
Mémoire technique examen statistique de
résultats obtenus dans le projet
"compteurs sur le toit"

DATE DUE
DATE DE RETOUR

ARR McLEAN 38-296