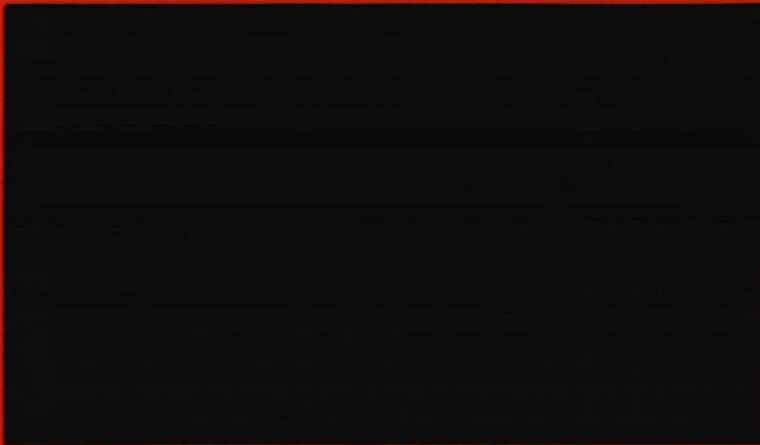


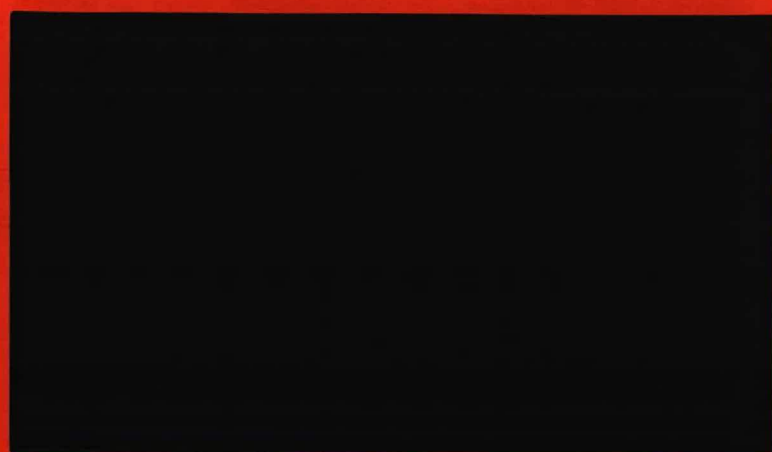
DI: 121544

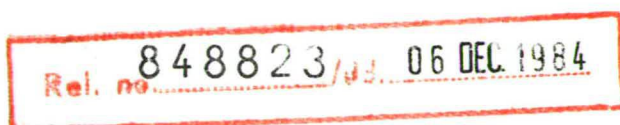
instituut voor zintuigfysiologie tno



Rijkswaterstaat
Adviesdienst Verkeer en Vervoer
Bureau Dokumentatie
Postbus 1031
3000 BA Rotterdam

C 7520





postbus 23
3769 ZG soesterberg

kampweg 5
3769 DE soesterberg

telefoon 03463 - 1444

rapport no. IZF 1984 C-13

ex. no.

Oplage: 60

VALIDATIE VAN ONGEVALSVOORSPELLERS VOOR
KRIMPENERWAARDKRUISPUNTEN

W.H. Janssen en A.R.A. van der Horst

In opdracht van de Dienst Verkeerskunde van de Rijkswaterstaat werd een eerder uitgevoerd onderzoek naar de mogelijkheid om aantallen ongevallen op kruispunten van het zgn. Krimpenerwaardtype te kunnen voorspellen, gevalideerd.

Aantal pag.: 13

I N H O U D

pag.

Samenvatting / Abstract	3
1. INLEIDING	4
2. SELECTIE VAN LOKATIES	5
3. DE BEOORDELINGSPROCEDURE	5
4. ONGEVALS- EN EXPOSITIEGEGEVENS	7
5. TOETSING VAN DE GEVONDEN RELATIES TUSSEN EXPOSITIE, OORDELEN EN ONGEVALSCIJFERS	7
REFERENTIES	11
BIJLAGE A: Lokaties van de in de validatiesteekproef opgenomen kruispunten	12
BIJLAGE B: Ongevalsgegevens (voorspeld vs. werkelijk) en expositiegegevens voor de in de validatie- steekproef opgenomen punten (1977-1980)	13

Validatie van ongevalsvoorspellers voor Krimpenerwaardkruispunten

W.H. Janssen en A.R.A. van der Horst

SAMENVATTING

Voor een steekproef van Krimpenerwaardkruispunten is nagegaan of de op grond van eerder onderzoek opgestelde regressievergelijkingen in staat zouden zijn om ongevalscijfers met een voldoende mate van nauwkeurigheid te voorspellen. Zulks bleek, binnen de marges van statistische betrouwbaarheid, het geval te zijn. De vergelijkingen kunnen daarmee als gevalideerd worden beschouwd.

Validation of accident-predicting equations for "Krimpenerwaard"-intersections

ABSTRACT

Regression equations predicting accident rates on intersections of the so-called "Krimpenerwaard" type were checked on a new sample of intersections. Predictive power appeared to be adequate, i.e. within the range given by statistical reliability. This result permits the regression equations to be considered as validated.

1. INLEIDING

In twee voorafgaande rapporten (Janssen en Van der Horst, 1978; Janssen, Brinkman en Van der Horst, 1983) is verslag gedaan van in opdracht van de Dienst Verkeerskunde van de Rijkswaterstaat verricht onderzoek op Krimpenerwaardkruispunten.

Het onderhavige rapport beoogt de werkzaamheden dienaangaande tot een afronding te brengen.

Oogmerk van het onderzoek is tweëërlei geweest, namelijk:

- a. Het opstellen van een "checklist" voor kruispunten van het Krimpenerwaardtype, dat wil zeggen een gebruikersvriendelijke lijst van beperkte lengte waarvan validiteit en betrouwbaarheid in het voorspellen van ongevallen genoegzaam aangetoond zouden zijn.
- b. Het leveren van een bijdrage aan de theorievorming over de factoren die het ongevalspotentieel van dit soort kruispunten bepalen.

De gevolgde aanpak is geweest om kwantitatieve oordelen over de kwaliteit van de informatieverwerking op Krimpenerwaardkruispunten, tezamen met een aantal expositie-indices te relateren aan ongevalscijfers.

Het eerste onderzoek (Janssen en Van der Horst, 1978) was exploratief van opzet en beperkte zich tot een steekproef van negen kruispunten. Binnen een aantal categorieën werden kruispuntkenmerken beoordeeld, te weten: (a) de aanwezigheid van het kruispunt; (b) de aanduiding van de route-mogelijkheden op het punt; (c) de geometrie van het punt; (d) de op het punt van toepassing zijnde voorrangsregeling; en (e) de aanwezigheid en de bewegingskenmerken van het overige verkeer. In totaal werden er ± 60 kenmerken beoordeeld, en de gegeven beoordelingen gecorreleerd met ongevalsgegevens. De resultaten konden gekwalificeerd worden als bemoedigend in de zin dat er een aantal significante correlaties aangetoond konden worden tussen oordelen en ongevalscijfers.

Een belangrijke beperking van dit eerste onderzoek lag in de kleine steekproefgrootte. Om die reden was een gedetailleerde statistische analyse van de resultaten niet mogelijk. Het belangrijkste oogmerk van het daarop uitgevoerde vervolgonderzoek (Janssen et al., 1983) is dan ook geweest het opvoeren van de statistische "power"

door een grotere steekproef te hanteren. Er zijn in dat onderzoek van 20 Krimpenerwaards ongevals- en expositiegegevens verzameld en oordelen gegeven over de informatieverwerkingsaspecten van de punten. Op grond daarvan konden (multipele) regressievergelijkingen worden opgesteld waarmee zich ongevalscijfers lieten voorspellen uit de combinatie van expositiegegevens en de scores op een zestal aan de gegeven oordelen ten grondslag liggende oordeelsdimensies.

In het onderhavige rapport zal onderzoek worden beschreven dat is verricht ter validatie van de genoemde regressievergelijkingen op een nieuwe - onafhankelijke - steekproef van Krimpenerwaardkruispunten. Een dergelijke procedure vormt, indien het resultaat positief uitvalt, de laatste stap in de ontwikkeling van een checklist. De theoretische implicaties van de verkregen resultaten zijn elders reeds behandeld (Janssen et al., 1983).

2. SELECTIE VAN LOKATIES

Uit het in de vorige fase van het onderzoek verkregen bestand aan Krimpenerwaardkruispunten, minus de reeds gebruikte punten (zie Janssen et al., 1983 voor een beschrijving van de wijze waarop het bestand is verkregen) zijn 12 punten op aselechte wijze gekozen. Een lijst van lokaties is bij dit rapport gevoegd als Bijlage A.

3. DE BEOORDELINGSPROCEDURE

Op grond van de resultaten van de vorige fase is een lijst van 11 te beoordelen items samengesteld voor gebruik in het huidige onderzoek. Deze items dekken de vijf dimensies - van de zes die de gegeven oordelen bleken te kunnen beschrijven - die in één of meer van de verkregen regressievergelijkingen zijn opgenomen. Strikt genomen zouden er voor het beschrijven van 5 dimensies niet meer dan 5 items nodig hoeven te zijn. Omdat echter géén der items een perfecte dekking geeft van de bijbehorende dimensie is er voor gekozen om meerdere items per dimensie op te nemen (behalve voor de 6e dimensie, H_6 , waarop slechts één item in het voorgaande onderzoek in voldoende mate bleek te laden).

De opgenomen items zijn de volgende:

Voor de eerste dimensie, H_1 :

- routeborden als indicator voor de aanwezigheid van het kruispunt, vanuit de zijwegen (variabele 2 uit de oorspronkelijke verzameling zoals beschreven in Janssen et al., 1983).
- informatie over de te volgen route vanuit de zijweg, voor de richting linksaf (variabele 24).
- idem, vanuit de hoofdweg, voor de richting rechtsaf (variabele 28).

Voor dimensie H_2 :

- informatie over aangeboden plaatsnamen, vanuit de zijwegen, voor de richting rechtsaf (variabele 19).
- plaatsing routeborden t.o.v. beslissingspunt, vanuit de hoofdwegen, voor de richting rechtdoor (variabele 32).
- idem, voor de richting rechtsaf (variabele 34).

Voor dimensie H_3 :

- zicht op baan verkeer van links, van ver, vanuit de zijwegen (variabele 68).
- schatting positie en snelheid verkeer van rechts, vanuit de hoofdwegen (variabele 107).

Voor dimensie H_5 :

- duidelijkheid te volgen baan vóór kruispuntvlak, vanuit de zijwegen, voor de richting rechtsaf (variabele 52).
- duidelijkheid te volgen baan op kruispuntvlak, vanuit de hoofdwegen, voor de richting rechtdoor (variabele 59).

Voor dimensie H_6 :

- zicht op baan achteropkomend verkeer, vanuit de zijwegen (variabele 78).

Aangezien de dimensie H_4 niet in de gevonden vergelijkingen bleek voor te komen zijn er géén die dimensie beschrijvende items in de lijst opgenomen.

Kwantitatieve oordelen op elk van de genoemde items zijn, op een 10-punts schaal, na inspectie op lokatie gegeven door de beide auteurs, en wel onafhankelijk van elkaar. De precieze wijze van

beoordeling is elders beschreven (Janssen et al., 1983). Het beoordelen vond plaats in voorjaar en zomer van 1983.

4. ONGEVALS- EN EXPOSITIEGEGEVENS

De relevante ongevals- en expositiegegevens zijn op dezelfde wijze verkregen als voor de oorspronkelijke steekproef, namelijk door opgave daarvan te vragen bij de betreffende provincies (Friesland, Zeeland en Limburg). Er is geïnformeerd naar ongevals- en intensiteitsgegevens (stromentellingen) over de periode 1977 t/m 1980, dezelfde als die voor de oorspronkelijke steekproef was aangehouden. Ook is geïnformeerd naar eventueel uitgevoerde grote en kleine reconstructies in of na die periode.

Om de ongevalsgegevens zo onbewerkt mogelijk ter beschikking te krijgen is verzocht kopieën van de VOR-formulieren te verstrekken, na onleesbaarmaking van personalia.

Over de analyseperiode (1977 t/m 1980) zijn per kruispunt expositie-indices berekend volgens Surti (1969), en wel zowel apart voor 1e- en 2e-rijbaanongevallen (S_1 , respectievelijk S_2) als voor het totale punt (S). De indices zijn berekend op uurbasis, d.w.z., de stromentellingen zijn per uur beschouwd en de relevante stromen per uur verwerkt. Naar jaren waarover wél ongevalsgegevens beschikbaar waren, maar geen tellingen, is geëxtrapoleerd met behulp van de CBS-verkeersindex-cijfers voor secundaire en tertiaire wegen over de relevante jaren (CBS, 1980). Omrekenen naar etmaalindices, teneinde op basis daarvan de expositie-indices over een jaar te berekenen, is gebeurd op basis van het "Onderzoek Verplaatsingsgedrag 1978" van het CBS, waarin de verkeersprestatie naar tijdperiode (van de dag) per kalendermaand is uitgesplitst. Per kruispunt resulteerden uiteindelijk dus: (a) aantallen ongevallen en (b) expositie-indices beide over de periode 1977 t/m 1980. Voor alle verdere bewerkingen is gerekend met het gemiddelde over de jaren 1977-1980; zie Bijlage B.

5. TOETSING VAN DE GEVONDEN RELATIES TUSSEN EXPOSITIE, OORDELEN EN ONGEVALSCIJFERS

Na het bij elkaar brengen van de benodigde gegevens is het mogelijk na te gaan in hoeverre de uit de oorspronkelijke steekproef afgeleide

regressievergelijkingen opgaan voor de validatiesteekproef. Daarbij dient bedacht te worden dat het voorspellend vermogen van deze vergelijkingen voor de validatiesteekproef veelal lager zal zijn dan voor de oorspronkelijke steekproef. Bij het opstellen van de vergelijkingen heeft men immers ook gekapitaliseerd op toevallig aanwezige verbanden (error) in de steekproefgegevens. Een gevonden regressievergelijking zal over het algemeen dus een overschatting geven van R^2 in de populatie en in een volgende steekproef. Kwantitatieve uitdrukkingen om de mate van te verwachten "krimp" ("shrinkage") in R^2 te berekenen zijn beschikbaar en ook toegepast op de betreffende vergelijkingen (Janssen et al., 1983).

De voor de oorspronkelijke steekproef gevonden vergelijkingen luiden:

$$\begin{aligned} \text{gemiddeld aantal ongevallen per jaar} = \\ .641 S + .950 H_6 - .634 H_5 + 2.56; \quad R^2 = .808 [1a] \end{aligned}$$

$$\begin{aligned} \text{gemiddeld aantal 1e-rijbaanongevallen per jaar} = \\ .448 S_1 + .45; \quad R^2 = .348 [1b] \end{aligned}$$

$$\begin{aligned} \text{gemiddeld aantal 2e-rijbaanongevallen per jaar} = \\ .433 S_2 - .733 H_1 + .486 H_6 - .453 H_2 - .293 H_5 + 1.98; \quad R^2 = .762 [1c] \end{aligned}$$

$$\begin{aligned} \text{gemiddeld aantal overige ongevallen per jaar} = \\ .356 S + .356 H_1 + .263 H_3 + .10; \quad R^2 = .866 [1d] \end{aligned}$$

Hierin zijn: S , S_1 en S_2 = expositie-indices volgens Surti ($\times 10^8$);
 $H_1 - H_6$ = scores op met de procedure Homals geëxtraheerde oordeelsdimensies;
 R^2 = gekwadrateerde correlatie ("coefficient of determination") tussen predictoren en criterium voor de betreffende regressievergelijking.

De op grond van deze uitdrukkingen te verwachten waarden van R^2 in een nieuwe steekproef, aangeduid als $\hat{\beta}_0^2$, bedragen volgens Lord (1950) en Nicholson (1960):

$$\hat{\rho}_c^2 = 1 - \frac{n-1}{n-p-1} \cdot \frac{n+p+1}{n} (1 - R^2) .$$

waarin n het aantal leden van de steekproef is en p het aantal predictoren in de multi-pele regressievergelijking. Als waarden voor $\hat{\rho}_c^2$ volgen hieruit (bij $n = 12$):

$$\begin{aligned} \hat{\rho}_c^2 &= .648 \text{ voor het totale aantal ongevallen;} \\ &.163 \text{ voor het aantal 1e-rijbaan ongevallen;} \\ &.345 \text{ voor het aantal 2e-rijbaan ongevallen;} \\ &.754 \text{ voor het aantal overige ongevallen.} \end{aligned}$$

Uiteraard zullen ook deze gekrompen waarden met een zekere onnauwkeurigheid zijn behept. Een exacte uitdrukking daarvoor is niet bekend. Op grond van de resultaten van een Monte Carlo simulatie valt de standaarddeviatie van $\hat{\rho}_c^2$ af te schatten als in de orde van .18 - .20 (Schmitt et al., 1977, Tabel III).

Het invoeren van expositiegegevens en scores op de Homals-dimensies in de vergelijkingen [1a] t/m [1d] resulteert voor de validatiesteekproef in waarden van R^2 zoals gegeven in Tabel I. Voorspelde aantallen ongevallen zijn opgenomen in Bijlage B.

De gevonden waarden van R^2 liggen binnen de marges van wat op grond van de voor "krimp" aangepaste regressievergelijkingen, en op grond van de betrouwbaarheid in de ongevalsgegevens zelf, verwacht mocht worden. Uitzondering hierop zijn de 1e-rijbaanongevallen, waar R^2 zeer veel hoger is dan verwacht mocht worden (.850 versus .163). Een nadere inspectie van de gegevens leert dat deze hoge waarde voor rekening komt van kruispunt nummer 12: zou dit punt niet worden meegerekend dan zou een R^2 resulteren van niet meer dan .182. Dit punt is overigens het punt met verreweg de hoogste expositiewaarden uit de gehele steekproef: blijkbaar is de regressievergelijking (waarin uitsluitend de expositiewaarde S_1 als predictor voorkomt) gevoelig genoeg om grote verschillen in expositiewaarden tot uiting te laten komen, maar slaagt ze daar minder in wanneer er gediscrimineerd moet worden tussen kleinere waarden.

Voor de overige ongevalscategorieën is het niet zo dat één enkel

punt wezenlijk invloed uitoefent op de hoogte van R^2 : de regressievergelijkingen zijn daarbij dus homogener over de range van waarden die de predictoren aannemen.

Tabel I. Waarden van R^2 in validatiesteekproef na toepassen van oorspronkelijke regressievergelijkingen.

=====		
ongevalscategorie	gevonden R^2	verwachte R^2 na "krimp" (β_c^2)

totaal	.757	.648
1e-rijbaan	.850 (.182*)	.163
2e-rijbaan	.459	.345
overige	.618	.754

=====

* indien punt nr. 12 wordt veronachtzaamd.

Er kan worden geconcludeerd dat de gehanteerde vergelijkingen in bevredigende mate van toepassing zijn op de validatiesteekproef. De uitdrukkingen [1a]-[1d] kunnen daardoor als gevalideerd worden beschouwd. Tevens is daarmee vastgesteld dat het voorspellend vermogen van deze vergelijkingen op grond van een beperkt aantal items (11) verwezenlijkt kan worden. Deze set van items kan daarom als checklist toegepast worden om Krimpenerwaardkruispunten te evalueren: een handleiding daartoe verschijnt separaat.

REFERENTIES

- CBS, Statistisch Zakboek, 1980.
- Janssen, W.H., A. Brinkman en A.R.A. van der Horst. De relatie tussen informatieverwerkingsaspecten, intensiteit en onveiligheid op Krimpenerwaardkruispunten. Rapport IZF 1983 C-21, Soesterberg, 1983.
- Janssen, W.H. en A.R.A. van der Horst. Onderzoek naar het rijgedrag op Krimpenerwaardkruispunten. Rapport IZF 1978 C-29, Soesterberg, 1978.
- Lord, F.M. Efficiency of prediction when a regression equation from one sample is used in a new sample. ETS RB50-40, Princeton, N.Y. Educational Testing Service, 1950.
- Nicholson, G.E. Prediction in future samples. In: I. Olkin et al. (Eds.), Contributions to probability and statistics. Stanford, Calif.: Stanford University Press, 1960.
- Schmitt, N., B.W. Coyle en J. Rauschenberger. A Monte Carlo evaluation of three formula estimates of cross-validated multiple correlation. Psychol. Bull., 1977 (84), 751-758.
- Surti, V.H. Accident exposure and intersection safety for at-grade, unsignalized intersections. Highway Res. Rec., 1969 (286), 81-94.

BIJLAGE A:

Lokaties van de in de validatiesteekproef opgenomen kruispunten

nummer	lokatie	provincie	nummer v.d. hoofdweg	hectometer- aanduiding
1	Hemelum	F	S9	19.6
2	Witmarsum	F	S9	44.1
3	Burgwerd	F	S9	46.5
4	Edens	F	S9	51.9
5	Deersum	F	S9	15.8
6	Birdaard	F	S2	26.7
7	Rinsumageest	F	S2	23.4
8	Zwaagwesteinde	F	S4	6.6
9	Heille	Z	RW58	19.2
10	Sas van Gent	Z	RW61	2.1
11	Zuiddorpe	Z	RW61	4.9
12	Wittern	L	RW278	17.8

BIJLAGE B:

Ongevalsgegevens (voorspeld vs. werkelijk) en expositiegegevens voor de in de validatiesteekproef opgenomen punten (1977-1980)

kruispunt nummer	ongevalscategorie								expositie- index (x 10 ⁸)	
	totaal		1e-rijbaan		2e-rijbaan		overige		S ₁	S ₂
	voorspeld	werkelijk	voorspeld	werkelijk	voorspeld	werkelijk	voorspeld	werkelijk		
1	3.53	2.25	0.73	1.25	0.39	0.25	0.20	0.75	0.63	0.56
2	3.72	2.75	1.19	0.25	0.60	1.50	0.47	1.00	1.65	1.04
3	3.51	1.25	0.71	0.50	0.37	0.75	0.20	0.00	0.57	0.52
4	3.62	2.75	0.96	1.50	0.47	0.75	0.34	0.50	1.13	0.74
5	5.38	2.75	1.93	0.00	1.00	2.75	0.95	0.00	3.31	1.93
6	3.51	1.00	0.78	0.25	0.31	0.75	0.20	0.00	0.73	0.39
7	3.65	2.25	1.04	0.00	0.55	1.75	0.42	0.50	1.31	0.92
8	4.78	5.25	3.16	1.25	2.48	2.00	2.04	2.00	6.04	5.24
9	4.06	4.25	1.55	1.00	1.49	2.00	1.01	1.25	2.46	3.02
10	4.37	3.25	2.15	0.75	2.04	1.50	1.45	1.00	3.79	4.26
11	4.19	4.25	1.99	1.75	1.52	1.00	1.20	1.50	3.43	3.09
12	7.56	11.25	8.82	6.75	6.58	3.00	1.54	1.50	18.69	15.02

