

Evaluatie van Intelligente Transport Systemen in de praktijk

Eline Jonkers
TNO
eline.jonkers@tno.nl

Isabel Wilmink
TNO & TrafficQuest
isabel.wilmink@tno.nl

Teije Gorris
TNO
teije.gorris@tno.nl

**Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk
20 en 21 november 2014, Eindhoven**

Samenvatting

Evaluatie van Intelligente Transport Systemen in de praktijk

Voor een succesvolle invoering van Intelligente Transport Systemen en verkeersmanagementmaatregelen is het nuttig om a priori de socio-economische effecten en maatschappelijke kosten en baten te kunnen bepalen. De overheid baseert bijvoorbeeld haar investeringsbeslissingen mede op deze effecten. Praktijkproeven en de evaluatie daarvan spelen een belangrijke rol hierbij.

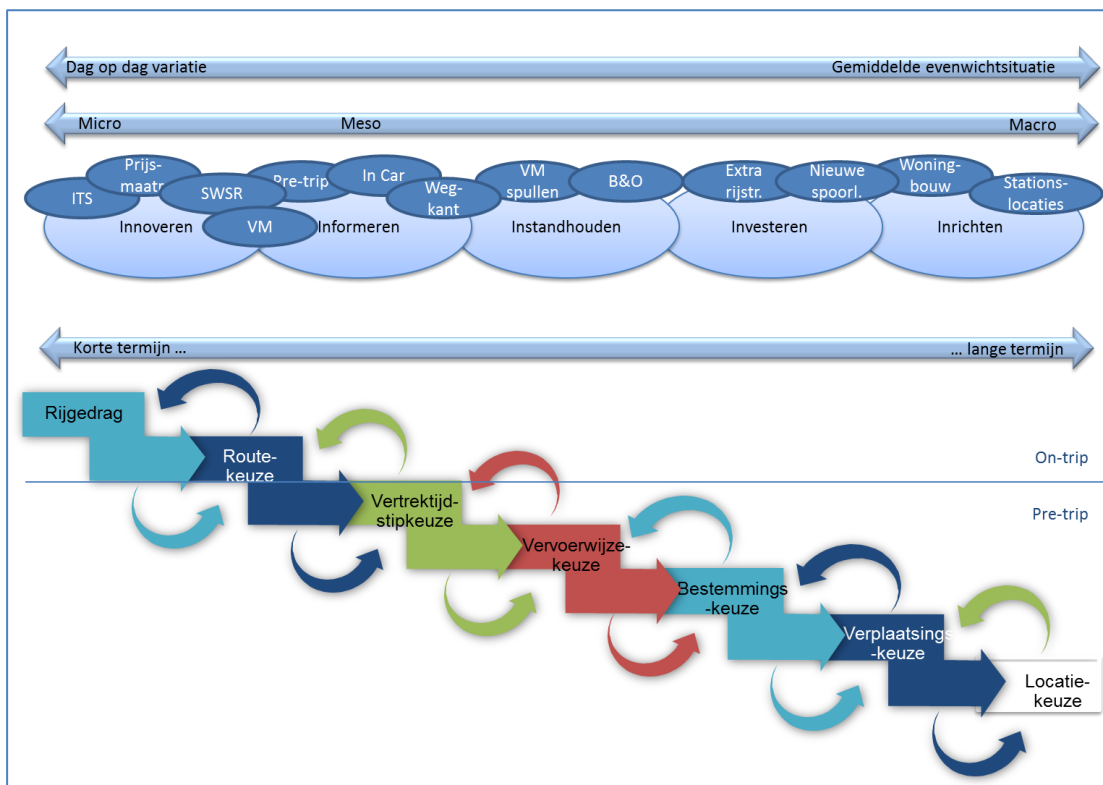
Voor het opzetten van praktijkproeven en het uitvoeren van evaluaties zijn handleidingen opgesteld, zoals de Europese FESTA methodologie en de Leidraad evaluaties benutting. De afgelopen jaren zijn deze methodologieën in verschillende projecten in de praktijk toegepast. Dit artikel geeft aan hoe de methodologieën gebruikt kunnen worden en geeft inzicht in hoe dit er praktisch gezien uitzag in een aantal projecten. De FESTA-methodologie helpt goed bij het structureren van evaluaties (en het opzetten van proeven in zijn geheel). Dit geldt zowel voor de opzettende partij als voor een eventuele beoordelende partij – de checklist die in de Leidraad staat, werkt goed om te bepalen of wel overal aan gedacht is.

De evaluatie van praktijkproeven blijft altijd een uitdaging; deze paper geeft een aantal voorbeelden van hoe de methodologieën in projecten worden toegepast en waar we dan tegen aan lopen. Het loont de moeite om ervaringen aan te dragen - voortschrijdende inzichten worden continu in de methodologieën opgenomen!

1. Inleiding

Het toevoegen van intelligentie en communicatiemiddelen aan voertuigen en infrastructuur (verzamelnaam Intelligente Transport Systemen – ITS) brengt grote beloftes met zich mee die ook door beleidsmakers op allerhande vlakken worden onderkend. Denk bijvoorbeeld aan het ITS Action Plan van de Europese Commissie (EC, 2011) en, in Nederland, de Routekaart Beter Geïnfomeerd op Weg (Connekt, 2013) en de programma's Beter Benutten en Connecting Mobility.

Voor beleidsmakers is het interessant om te weten welke bijdragen ITS systemen nu daadwerkelijk leveren aan het realiseren van beleidsdoelstellingen op gebied van bijvoorbeeld doorstroming, verkeersveiligheid en milieu, maar ook aan andere doelstellingen zoals het beter informeren van reizigers of het meer overlaten aan de markt. Daarnaast is ook aandacht nodig voor andere aspecten, zoals risico's en aansprakelijkheden, maar ook potentiële barrières rondom privacy en veiligheid die deze systemen met zich meebrengen. Deze paper gaat in op het inschatten van potentiële effecten van ITS systemen door middel van evaluatie van systemen en maatregelen in praktijkproeven. We richten ons daarbij op individuele maatregelen of bundels van een beperkt aantal maatregelen; evaluaties van complete beleidsprogramma's vallen buiten de scope van deze paper.



Figuur 1: Keuzes die onder invloed van ITS kunnen veranderen.

ITS is een containerbegrip, waar heel veel maatregelen in geplaatst kunnen worden – als er een ICT-component aan zit, is een maatregel al gauw ITS te noemen. Figuur 1 laat zien welke soorten keuzes allemaal gemaakt kunnen worden (waar ITS invloed op kunnen hebben), en hoe zich dat verhoudt tot de verschillende soorten maatregelen waar momenteel aandacht voor is (waarbij ITS maatregelen vooral in de categorieën

innoveren, informeren en instandhouden liggen), ook qua schaalniveau (mate van detail in tijd en ruimte).

Figuur 1 laat zien hoeveel keuzes door ITS beïnvloed kunnen worden, en dat geeft direct een indicatie van hoe complex de evaluatie van dergelijke maatregelen kan worden.

Want hoe houd je rekening met effecten op al die vlakken? De volgende paragraaf geeft daarvoor een aantal handvatten.

2. Evaluatieaanpakken

Er is niet één alles omvattende evaluatiemethodologie waarmee alles kan. Maar er zijn wel diverse aanpakken waarmee belangrijke stappen in evaluaties gezet kunnen worden. We nemen hier de FESTA-methodologie als uitgangspunt, en bespreken daarnaast enkele aanvullende aanpakken.

2.1 De FESTA-methodologie en de Leidraad evaluaties benutten

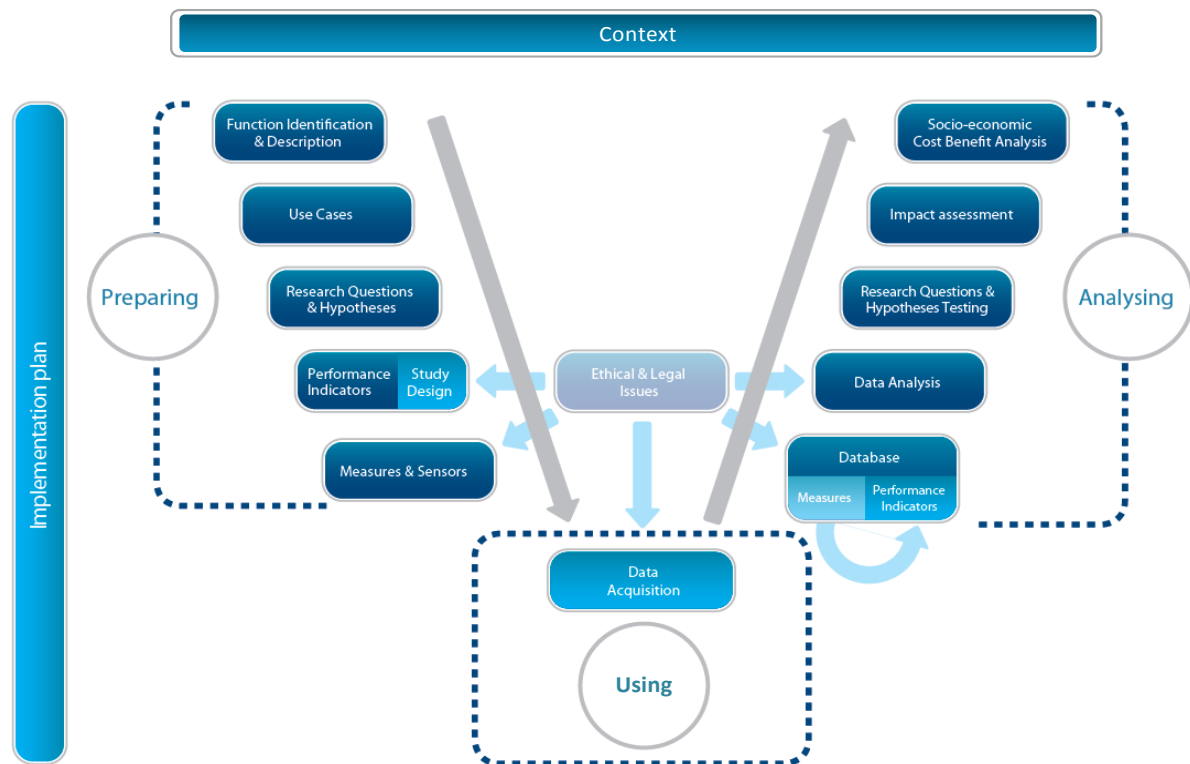
De ervaringen in diverse Europese projecten waarin ITS ontwikkeld en getest werden leidden tot de wens om een gedegen evaluatiemethodologie te ontwikkelen, vooral voor de zogenaamde *Field Operational Tests (FOTs)*, die doorgaans groot van opzet zijn en waarin behoorlijk complexe systemen getest worden. Het FESTA project werd hiervoor opgestart, en in 2008 heeft het FESTA consortium de eerste versie van het FESTA handboek uitgebracht, waarin de FESTA methodologie uiteen wordt gezet. FESTA bleek goed bruikbaar in de praktijk en wordt nu beschouwd als dé (evaluatie)methodologie voor Field Operational Tests in Europa. De methodologie is echter breder inzetbaar.

Het FESTA handboek is een levend document. Ervaringen met het gebruiken van FESTA in de praktijk in verschillende projecten zijn gebruikt om het FESTA handboek te verbeteren in de Europese Support Action FOT-Net. In 2011 is een eerste update van het handboek verschenen en begin 2014 een tweede update. De meest recente versie stamt uit 2014 (FOT-Net consortium, 2014-1).

Het FESTA handboek biedt praktische handreikingen voor het uitvoeren van een FOT. Het beschrijft op een systematische manier het totale proces van plannen, voorbereiden, uitvoeren, analyseren en rapporteren van een FOT. Ook geeft het handboek informatie over administratieve, logistieke, juridische en ethische aspecten. Een ander doel van het handboek is het mogelijk maken van standaardisatie van (aspecten van) FOTs. Dit helpt bij het vergelijken van verschillende FOTs.

Het FESTA handboek kan gebruikt worden voor de evaluatie van zowel autonome systemen zoals ADAS (Advanced Driver Assistance Systems) en in-car informatiesystemen als coöperatieve systemen (waarin communicatie tussen voertuigen of tussen voertuigen en infrastructuur gebruikt wordt). Hiernaast kan het ook gebruikt worden voor de evaluatie van 'nomadic devices' en Naturalistic Driving Studies. Veel van de activiteiten zoals beschreven in het handboek zijn ook nodig voor de evaluatie van voertuig- en infrastructuur gebaseerde ICT technologie, en daarom is de FESTA methodologie toepasbaar op allerlei soorten veldtesten, (praktijk)proeven en pilots. FESTA werd gebruikt in de Europese FOTs euroFOT, TeleFOT, DRIVE C2X en FOTSis en in het project eCoMove.

In de FESTA-methodologie is een FOT beschreven in een aantal stappen. Deze stappen zijn op een overzichtelijke manier weergegeven in de FESTA-V, zie Figuur 2.



Figuur 2: De FESTA-V.

De eerste stappen, waaronder het formuleren van een doel en het opzetten van een (onderzoeks)team, en de laatste stappen – analyse van het geteste systeem en socio-economische evaluatie – hebben te maken met meer algemene aspecten van een FOT en het aggregeren van resultaten. Hoe 'dieper' in de FESTA-V, hoe meer de focus ligt op onderdelen met een hoog detailniveau, zoals welke indicatoren te kiezen en hoe de database op te zetten. Evaluatie vindt plaats aan het einde van de FOT of proef (rechtsboven in de V), maar eerder in het proces of project moeten de analisten al betrokken worden, omdat evaluatie van veel zaken afhangt die eerder in het project plaatsvonden, zoals experimentopzet, deelnemersselectie, meten van data en piloting.

Specifiek voor de evaluatie van benuttingsmaatregelen in Nederland is in 2011 de Leidraad evaluaties benutting (Wilmink et al., 2011) verschenen (in het vervolg ook wel Leidraad genoemd), een adaptatie van de FESTA methodologie. Het houdt rekening met verschillende niveaus van inspanning en (dus) het soort evaluatie dat uitgevoerd wordt. Naast de al genoemde onderwerpen, behandelt de Leidraad ook expliciet de evaluatie van verkeersmanagementmaatregelen. Voor de Leidraad is er ook een Nederlandse versie van de FESTA-V ontwikkeld. Ook in Nederland worden FESTA en de Leidraad gebruikt, bijvoorbeeld voor de PraktijkProef Amsterdam In-Car.

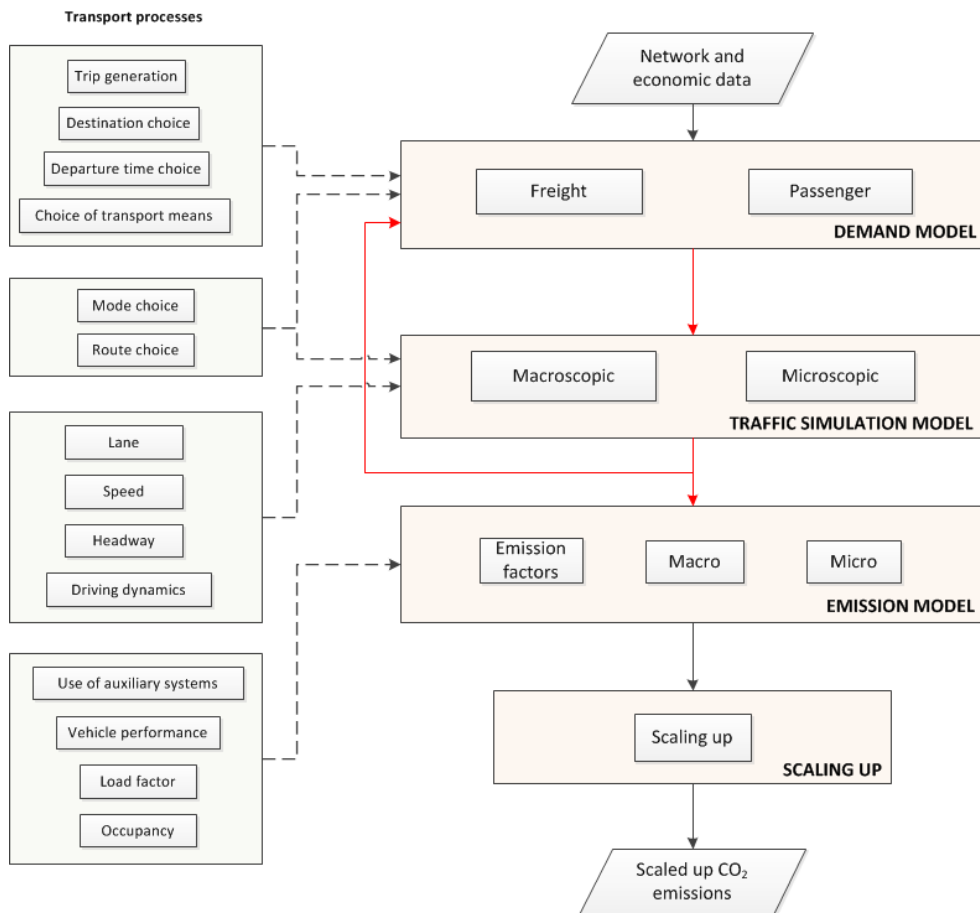
2.2 Aanvullende aanpakken: Amitran en ECOSTAND

In het Europese project Amitran is een evaluatiemethodologie voor het bepalen van de effecten van ITS maatregelen op CO₂ emissies ontwikkeld, inclusief opschaling. Deze methodologie kan worden gezien als een verbijzondering van bepaalde aspecten van FESTA. De methodologie is minder breed dan FESTA (het gaat niet over het opzetten van een FOT of veldtest), maar gaat dieper in op de manier waarop ITS maatregelen CO₂ emissies beïnvloeden en het gebruik van modellen om de CO₂ effecten te bepalen. De methodologie is generiek: alle potentiële effecten van ITS maatregelen worden beschouwd en de methodologie is toepasbaar op alle soorten ITS maatregelen, modaliteiten en settings, en hangt niet af van specifieke software die gebruikt wordt. De evaluatiemethodologie in Amitran werkt met behulp van effectketens. ITS maatregelen hebben op allerlei manieren effect op CO₂ emissies. Er zijn bijvoorbeeld directe en indirecte effecten, korte- en lange termijn effecten, effecten op zaken die met de verkeersvraag te maken hebben en effecten die met rijgedrag te maken hebben. Afhankelijk van de type effecten die ITS maatregelen hebben, zijn verschillende (combinaties van) modellen nodig om de effecten op CO₂ emissies uit te rekenen. Een overzicht hiervan is te zien in Figuur 3.

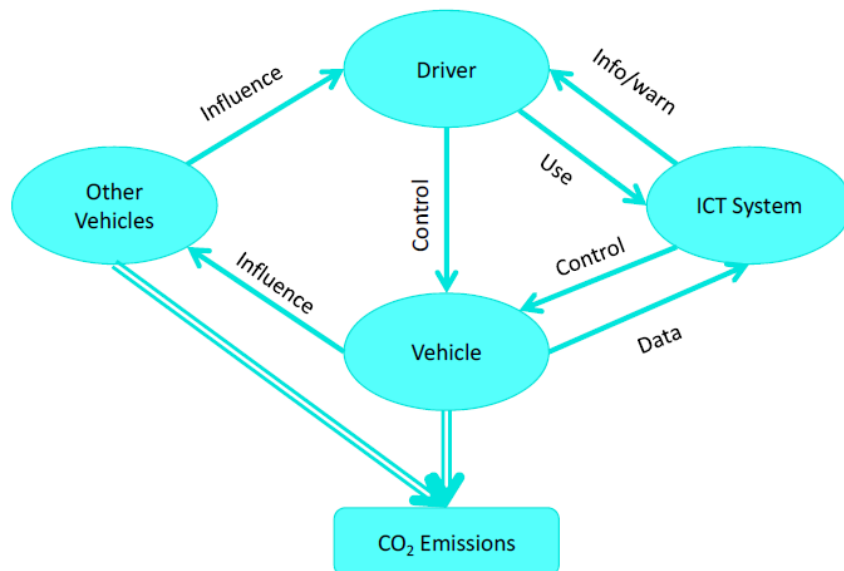
Het Amitran project heeft publieke, generieke interfaces ontwikkeld om verschillende typen modellen te koppelen (bijvoorbeeld een macro verkeersmodel aan een macro emissiemodel of een micro verkeersmodel aan een micro emissiemodel). Meer informatie over het Amitran project kan gevonden worden op de website (Amitran, 2014-1) en op de openbare Wiki die voor het project ontwikkeld is (Amitran, 2014-2). Op deze Wiki zijn onder andere de publieke interfaces tussen modellen, data bruikbaar voor opschaling, methodologie en voorbeelden te vinden. Meer specifieke informatie over de methodologie kan gevonden worden in Deliverable 4.1 (Amitran, 2014-3).

In het Europese project ECOSTAND is, in samenwerking met het Japanse ITS-Energy project, uitgewerkt hoe ITS maatregelen goed in beeld gebracht kunnen worden, en hoe ze gemodelleerd kunnen worden (vastgelegd in een gezamenlijk Europees-Japans-Amerikaans rapport (Energy ITS & ECOSTAND, 2013). Dit biedt ook aanknopingspunten voor de opzet van een evaluatie van (veld)testen van dit soort maatregelen. In het rapport worden "reference models" gegeven voor vijf categorieën ITS. Daarbij wordt onderscheid gemaakt naar "category models" en "instance models" (zie Figuur 4 en Figuur 5). Een reference model geeft aan welke mechanismen een rol spelen bij bepaalde ITS, en biedt zo inzicht in welke mechanismen ook in de evaluatie een rol zullen spelen. Een "instance model" is een specifieke uitwerking ervan voor een bepaalde maatregel. Voor evaluaties met behulp van verkeersmodellen bieden de diagrammen ook inzicht in welk type verkeersmodel het meest geschikt is (waarbij een check plaats moet vinden of het in te zetten verkeersmodel de belangrijke mechanismen uit de diagrammen –ze zijn niet per se allemaal altijd van belang, of er bestaat al inzicht in zodat aannames gedaan kunnen worden– wel goed representeert).

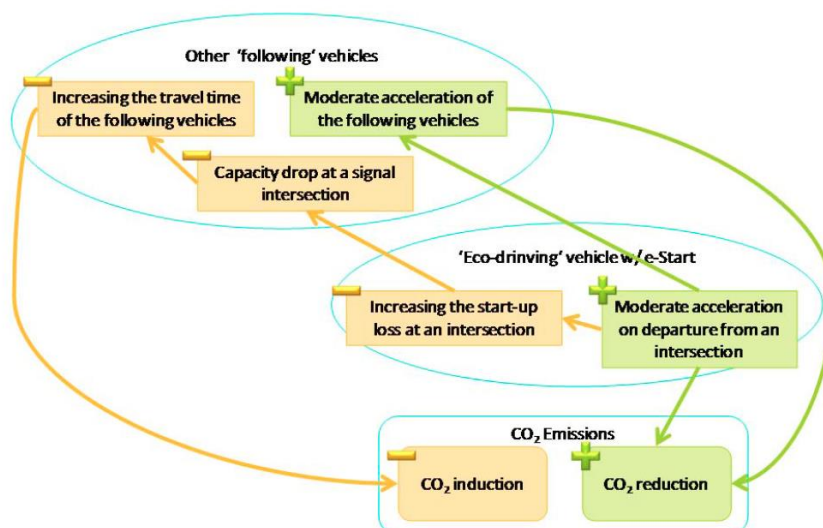
Dit soort diagrammen horen bij de bovenkant van de FESTA-V: aan de linkerkant worden ze gebruikt om de maatregel/functie goed te beschrijven, en de stappen naar beneden te zetten. Aan de rechterbovenkant van de V kunnen de diagrammen er weer bij gepakt worden om de effecten zo compleet mogelijk te beschrijven. Hoewel deze diagrammen vooral voor het evalueren van de effecten op CO₂ emissies zijn gemaakt, kunnen ze ook breder toegepast worden (of beperkt worden tot de verkeerseffecten).



Figuur 3: Amitran Framework Architecture.



Figuur 4: Category model voor ITS uit de categorie "Improving Driving Behavior" (Energy ITS & ECOSTAND, 2013).



Figuur 5: Instance model voor de applicatie "e-Start" (Energy ITS & ECOSTAND, 2013).

2.3 Gebruik van de methodologieën in Nederland

Hierboven is al aangegeven dat de FESTA-methodologie ook gebruikt wordt in Nederland – waarbij vooral gebruik wordt gemaakt van de Leidraad. Naast deze Leidraad Evaluaties benutting is er ook het rapport "Richtlijnen ex-post beleidsevaluaties" (MuConsult, 2011). Samen geven deze rapporten een compleet overzicht hoe evaluaties van benuttingsmaatregelen en -pakketten moeten worden opgezet en uitgevoerd. De rapporten zijn opgesteld in het kader van het project 'Beleidsafweging systematiek Benutten', een project met als doel het ontwikkelen van ex-ante en ex-post evaluatiemethoden ten behoeve van een verantwoorde kosten-batenanalyse voor de (brede) inzet/uitrol van benuttingsmaatregelen.

Speciaal voor het evalueren van regelscenario's is een verder uitwerking gemaakt van de Leidraad. Bij deze uitwerking is veel aandacht gegeven aan het opzetten van een evaluatie, de selectie van data en de analyse. Het rapport "Leidraad evaluatie regelscenario's" (Witteveen & Bos, 2012) bevat veel tips en aandachtspunten om een evaluatie van regelscenario's goed op te zetten en uit te voeren, en is vooral bedoeld om de verkeerskundigen in de verkeerscentrales te ondersteunen bij hun taak om regelscenario's te ontwerpen en te operationaliseren. Verder is er nog het Kader verkeerskundige evaluatie regelscenario's (Arcadis, 2012), die goed in combinatie met de Leidraad evaluatie regelscenario's ingezet kan worden.

3. Ervaringen

Met het beschrijven van de evaluatiemethodologieën in hoofdstuk 2 wordt wellicht de indruk gegeven dat evalueren een kwestie is van het volgen van de juiste methodologie, en dat dan de resultaten vanzelf volgen. Helaas is het niet zo eenvoudig en zijn de methodologieën geen kookboeken waarvan eenvoudigweg het 'recept' gevolgd kan worden. Ook met de juiste methode is evalueren nog steeds een flinke klus waarbij

kennis en praktijkervaring nodig is van verschillende aspecten, zoals de te testen maatregel, de omgeving waarin getest wordt, dataverzameling, modellen, etc. etc. In dit hoofdstuk wordt een overzicht gegeven van een aantal problemen en oplossingen bij evalueren in de praktijk. Deze zaken zijn niet alleen door de auteurs ondervonden, maar ook verzameld in een werkgroep over 'Impact Assessment and Scaling up' van het FOT-Net project. De inzichten die in de werkgroep zijn opgedaan zijn onder andere verzameld door het houden van interviews, webinars en workshops met experts. De volledige uitkomsten van de FOT-Net werkgroep kunnen gevonden worden in (FOT-Net consortium, 2014-2). Aanvullend zijn de ervaringen van TNO in Nederlandse projecten (zoals de evaluatie van de dynamische snelheidslimieten) gebruikt voor dit artikel.

In de meeste FOTs en praktijkproeven worden de FESTA methodologie en/of Leidraad niet heel strikt gevolgd. Niet omdat ze niet goed of niet bruikbaar zijn, maar omdat de methodologieën niet gedetailleerd en specifiek genoeg zijn om voor een zeer diverse range aan projecten als kant-en-klare methode gebruikt te kunnen worden. De methodologie wordt meestal als basis gebruikt en aangepast voor de specifieke doelstellingen van een project. Dit is in lijn met hoe de methodologieën zijn opgezet en bedoeld. Een aanbeveling vanuit veel projecten is dat, vanuit oogpunt van analyse, genoeg tijd moet worden besteed aan het opzetten van een methodologisch goed onderbouwde evaluatie.

Een steeds terugkerend aspect bij het uitvoeren en evalueren van proeven is de timing. Evaluatie vindt plaats aan het eind van een project, en is afhankelijk van alle eerdere stappen in het project (zoals experimentopzet, deelnemersselectie en data-acquisitie). Alle problemen en vertragingen die eerder in het project ontstonden, hebben invloed op de evaluatie. In veel gevallen wordt de tijd die er is om de data-analyse en evaluatie uit te voeren ingekort. Alle FOTs die input voor de FOT-Net werkgroep gaven zeiden dat ze te weinig tijd hadden om de evaluatie (goed) uit te voeren.

Beleidsmakers verwachten concrete resultaten van praktijkproeven, zoals een reductie in aantal ongevallen door het gebruik van een bepaald systeem of maatregel, en kosten en baten op opgeschaald niveau (bijvoorbeeld voor een land of heel Europa). Deze resultaten worden vaak ook beloofd in offertes. Echter, het is erg lastig deze getallen ook daadwerkelijk op te leveren en er kunnen verschillende redenen zijn waarom een FOT of proef deze resultaten uiteindelijk niet oplevert. Het managen van verwachtingen is zeer belangrijk, gedurende de duur van het hele project.

3.1 Pilotfase

In een praktijkproef zit meestal een pilotfase om te testen of aan het eind van het project de onderzoeksvragen beantwoord kunnen worden. De pilotfase is een soort generale repetitie voordat de praktijkproef echt uitgevoerd wordt. Het is belangrijk dat in de pilotfase niet alleen de techniek getest wordt (om te testen of het systeem goed werkt), maar dat het hele proces van dataketen en dataverwerking wordt getest, tot en met de evaluatie. Tijdens de pilotfase kunnen een aantal specifieke zaken getest en gecontroleerd worden, zoals de communicatie tussen de verschillende teams die betrokken zijn bij de uitvoering van de proef, de (robustheid van de) tools voor

dataverzameling en data-overdracht en de (robustheid van de) evaluatietools. Daarom is het belangrijk alle projectpartners te betrekken bij de pilotfase.

In de praktijk wordt meestal een uitgebreid plan opgesteld voor de pilotfase. Echter, de tijd die nodig is voor een goede pilotfase wordt vaak onderschat en door tijdgebrek wordt de pilot vaak niet volledig uitgevoerd, waardoor juist de laatste fase (het testen van het evaluatieproces) niet gedaan wordt of gedaan wordt als de proef al gestart is en aanpassingen (bijvoorbeeld op gebied van experimentopzet en dataverzameling) niet meer mogelijk zijn. De ervaring in FOTs is dat er tijdens de pilotfase altijd problemen aan de oppervlakte komen die moeten worden opgelost voordat de echte praktijkproef start. De projecten euroFOT en TeleFOT hadden een pilotfase van drie tot zes maanden, en dat bleek te kort. De ISA-UK FOT had een pilot van één jaar, en zelfs zij hadden toen de proef echt startte nog steeds enkele problemen. Vanuit evaluatieoogpunt is een korte pilotfase niet acceptabel, vanwege het risico dat niet de hele dataketen getest kan worden. Het wordt aanbevolen een lange periode te reserveren voor de pilotfase en expliciete feedback loops te hebben, zodat (negatieve) uitkomsten in de pilotfase leiden tot aanpassingen van eerdere stappen in het proces.

Om de pilotfase inhoudelijk te versterken, heeft het EU-project TeleFOT het volgende in praktijk gebracht: het zelf piloten. Dit is specifiek aan te bevelen voor mensen die aan de subjectieve dataverzameling en data-analyse werken. Laat mensen die aan het project werken de vragenlijsten beantwoorden, reisdagboeken invullen, etc. Laat ze alles doen wat ook van 'echte' deelnemers wordt verwacht, voordat de echte deelnemers starten. Op deze manier kan worden getest of hetgeen van deelnemers wordt gevraagd en verwacht, realistisch is.

3.2 Dataverzameling

Om effecten die gevonden worden tijdens een proef (bijvoorbeeld een verandering in reistijd) (helpen) uit te leggen, is meer nodig dan alleen de 'primaire' indicatoren. Ook verklarende variabelen (zoals de omstandigheden waaronder gereden of gereisd is) zijn zeer belangrijk. Idealiter worden tijdens een proef zoveel mogelijk in-depth data verzameld, zoals:

- Videodata, waarbij het verwerken van deze data zoveel mogelijk geautomatiseerd wordt
- Vragenlijsten en reisdagboeken (op bepaalde momenten voor, tijdens en na de proef), bijvoorbeeld over multimodaal reizen om mobiliteitsanalyses te doen. De drempel om vragenlijsten en dagboeken in te vullen moet zo laag mogelijk zijn, bijvoorbeeld door gebruik te maken van vooraf gespecificeerde antwoordmogelijkheden en meerkeuzevragen.
- Data over de omgeving, zoals voertuigen in de omgeving (volgtijden) en verkeersintensiteiten.
- Meta data: goede beschrijvingen van de data en testen, zoals wie in welk voertuig reed, welke systemen getest zijn in bepaalde auto's en ritten, omstandigheden tijdens het rijden, andere systemen die in de auto aanwezig waren, datum en tijdstip van de test, etc.

- Audio data, bijvoorbeeld om bestuurders van de voertuigen de optie te geven mondeling toe te lichten wat er gebeurde tijdens het rijden en wat hun ervaringen waren. Dit is een laagdrempelige manier om feedback te geven.
- Loggen van het systeem: de status (aan, uit), de informatie die gepresenteerd wordt aan de bestuurder van het voertuig.

Het verzamelen van al deze data is kostbaar en de analyses van de data kosten veel tijd. Een mogelijke oplossing (die bijvoorbeeld is gehanteerd in het TeleFOT project) is om maar voor een deel van de proef (bijvoorbeeld een deel van de voertuigen of een deel van de tijd) in-depth data te verzamelen.

Naast het opstellen van een goed plan voor het verzamelen en verwerken van data met een lijst maten en indicatoren en specificaties van hoe data gemeten gaan worden, is het aan te bevelen alternatieve databronnen te identificeren voor de belangrijkste maten, en om parallelle back-up systemen te hebben (mogelijk met lagere nauwkeurigheid), zodat er terugvalopties zijn en de proef niet op één manier van data inwinnen vertrouwt. Om dit te doen is het handig te identificeren welke prioriteiten de data hebben ('must have data' versus 'nice to have data'). Mogelijk zijn niet voor alle gevallen even gedetailleerde data nodig.

3.3 Integratie van resultaten

Bij het uitvoeren en evalueren van praktijkproeven is het belangrijk dat er genoeg tijd is voor het integreren van resultaten. Aan het eind van de analysefase zijn er losse resultaten op het gebied van doorstroming, gedrag, veiligheid, milieu, acceptatie, etc. Uiteindelijk moeten er overall conclusies worden getrokken over een systeem of maatregel, waarin de verschillende impactgebieden geïntegreerd zijn. Hiervoor is het nodig de resultaten te laten bezinken en te bediscussiëren met experts. De tijd die dit kost is niet altijd beschikbaar. Wat in elk geval helpt, is het betrekken van de analisten vanaf de start van het project en het laten samenwerken van analisten, mensen die de data verzamelen, de test opzetten, etc., om de integratie van resultaten later in het project te bevorderen.

Het is aan te bevelen om vroeg in het project een template te maken voor het rapporteren van resultaten (zowel intern als extern), zodat de analisten weten wat er van ze verwacht wordt. Een heldere en systematische rapportage helpt bij het integreren van de resultaten en de overall maatregel/systeemanalyse. Een voorbeeldtemplate zou moeten worden ingevuld door de persoon die uiteindelijk verantwoordelijk is, zodat degenen die uiteindelijk de template invullen weten welk detailniveau er van hen verwacht wordt.

4. Praktijkcases

4.1 *PraktijkProef Amsterdam Evaluatieplan*

De praktijkproef Amsterdam (PPA) richt zich op het verminderen van files in de regio Amsterdam, via een brede, innovatieve aanpak. In het in-car spoor¹ krijgen weggebruikers persoonlijke reisinformatie in het voertuig om zelf de beste reiskeuze te maken. Het gaat dan bijvoorbeeld om routeadvies (pre-trip en on-trip) en vertrektijdstipadvies (pre-trip). TNO is in één van de twee consortia die aan het in-car spoor werken verantwoordelijk voor de evaluatie. Op het moment van schrijven zijn de evaluatieplannen gereed, en wordt de data-logging opgezet, maar loopt de proef nog (net) niet. In het programma van eisen was aangegeven dat de evaluatieplannen (voor de evaluaties van het gebruik van de in-car app voor reguliere omstandigheden en bij evenementen) opgezet dienden te worden conform de FESTA-V. Dit leverde weliswaar vrij lange evaluatieplannen op, maar deze blijken bij de uitvoering van de evaluatie heel bruikbaar als naslagwerk. Zo zijn bijvoorbeeld de enquêtes die onder de deelnemers uitgezet zullen worden eenvoudig naast de hypothesen uit het evaluatieplan te leggen, om te kijken of we met de in de enquêtes opgenomen vragen en mogelijke antwoorden in staat zullen zijn de relevante hypothesen (met name over gebruikerservaring en inschatting van verandering van eigen reisgedrag) te toetsen.

Met het opstellen van het evaluatieplan zijn we er natuurlijk nog niet. Het plan blijft op veel plekken nog een plan op hoofdlijnen, omdat toen het geschreven werd nog niet precies duidelijk was hoe de geplande diensten precies aan de deelnemers aangeboden zouden worden en hoe het reisgedrag gemeten zou worden. Daarom was het bijvoorbeeld al wel mogelijk om een hypothese te formuleren als "Een meerderheid van de deelnemers volgt het routeadvies regelmatig (in meer dan 60% van de gevallen) op", maar nog niet wat nou precies werd verstaan onder "opvolgen". De uitwerking hiervan ging gelijk op met de ontwikkeling van de app-software.

4.2 *DRIVE C2X Field Operational Tests*

In het Europese project DRIVE C2X zijn verschillende coöperatieve systemen (voertuig-voertuig en voertuig-infrastructuur) in de praktijk getest en geëvalueerd op test sites in zeven landen (waaronder de technische tests in Helmond). De systemen waren vooral veiligheidssystemen, zoals Traffic Jam Ahead Warning, Weather Warning, Road Works Warning en Obstacle Warning. Doel van het project was het evalueren van coöperatieve systemen door middel van het uitvoeren van een FOT, en het promoten van coöperatieve systemen. De FOT en evaluatie zijn opgezet volgens het FESTA handboek. De lastigheid in dit project voor de evaluatie zat hem in het feit dat het om het testen van *coöperatieve* systemen ging. Om deze in de praktijk te testen moeten de uitgeruste voertuigen elkaar tegenkomen of de uitgeruste voertuigen langs de uitgeruste infrastructuur rijden, en dan moet er ook nog een 'event' zijn (een situatie waarin het systeem actief is, bijvoorbeeld file, wegwerkzaamheden of slecht weer). Bij een 'naturalistic test' (waarbij gewone autobestuurders hun normale ritten maken) is de kans dat dit gebeurt zeer klein, en dit zou te weinig data opleveren om een evaluatie uit te

¹ Er is ook een wegkantspoor; zie http://www.rijkswaterstaat.nl/zakelijk/innovatie/doorstroming_door_innovatie/lopende_projecten/praktijkproef_amsterdam/

kunnen voeren. Daarom is er vooral gewerkt met zogeheten gecontroleerde testen, waarbij autobestuurders bijvoorbeeld een vooraf gedefinieerd traject moesten rijden of in een bepaalde omgeving moesten rondrijden. Ook werden sommige events geforceerd (bijvoorbeeld een kapotte auto langs de weg). Een andere moeilijkheid in dit project was de opschaling. Vanwege de specifieke gevallen waarin de systemen effectief zijn, moesten ook voor deze specifieke gevallen data op Europees niveau gevonden worden. De FESTA methode biedt wel een beschrijving van hoe deze opschaling uitgevoerd kan worden, maar het zoeken van de data moest door het project gedaan worden. Meer informatie over het project DRIVE C2X en de resultaten kan gevonden worden in (DRIVE C2X, 2014).

4.3 eCoMove Validatieplan

Het eCoMove project (2010-2014) richtte zich op "Cooperative Mobility Systems and Services for Energy Efficiency". De vraag die in eCoMove centraal stond, is hoe coöperatieve systemen zó ingezet kunnen worden dat de energiebelasting en CO₂ uitstoot van verplaatsingen over de weg significant – doelstelling: 20% – verminderen. eCoMove was een omvangrijk project (budget: 22,5 miljoen euro), waarin tientallen applicaties en ondersteunende technologieën ontwikkeld werden en vervolgens getest op de weg, in rijssimulators of in verkeerssimulaties. Om ervoor te zorgen dat aan het eind van het project een inschatting gedaan kon worden of een 20% reductie inderdaad bereikt kan worden, dienden alle applicaties op een zo uniform mogelijke manier geëvalueerd te worden. De evaluatie werd opgezet volgens het FESTA handboek, waarbij de FESTA-V veel houvast bood – alle partners werd gevraagd informatie over hun tests op dezelfde manier aan te leveren, in door de coördinerende partner aangeleverde templates. Ook de door het FESTA-consortium opgestelde spreadsheet met voorbeeldindicatoren, -maten en -sensoren kwam goed van pas (maar behoefde wel aanvulling, met name voor de tests die niet op de weg uitgevoerd werden – FESTA is tenslotte opgezet voor Field Operational Tests). Zo werden gemeenschappelijke indicatoren gedefinieerd in drie categorieën: milieu (CO₂ emissies, brandstofverbruik), mobiliteit (reistijd, vertragingen, aantal stops), en rijgedrag (o.a. opvolging, surrogate safety measures zoals time to collision, gebruikersacceptatie, werklast).

Om niet totaal de weg kwijt te raken in het grote aantal voorgestelde tests, is een structuur gehanteerd met individuele test set-ups (die aangeleverd werden door de partners die de tests uitvoerden), die geclusterd werden in test cases en vervolgens in validatiescenario's (die uiteindelijk ook weer geclusterd werden). In totaal zijn tientallen test set-ups op uniforme wijze gedocumenteerd, en vervolgens zijn de resultaten van deze tests ook weer op uniforme wijze gerapporteerd. Dat maakte het nog steeds niet eenvoudig een eindoordeel te geven over of de 20% reductie gehaald kan worden, omdat de test set-ups daarvoor teveel van elkaar verschilden: resultaten van tests met één voertuig op de weg stonden naast resultaten van tests met veertig proefpersonen in de rijssimulator en simulaties van series kruispunten met honderden voertuigen². Uiteindelijk is middels expert judgment geconcludeerd dat 20% niet haalbaar lijkt met de geteste

² Het Programma Beter Benutten kent een soortgelijke uitdaging – ook daar moeten grote aantallen verschillende typen maatregelen samen beoordeeld worden.

applicaties, maar dat in stedelijke omgeving wel een reductie van meer dan 10% haalbaar lijkt (eCoMove, 2014).

4.4 ecoDriver aanpak

In het lopende Europese project ecoDriver worden in-car applicaties ontwikkeld die feedback geven aan de autobestuurder en hem/haar ondersteunen om efficiënter te rijden (door het uitstoten van minder emissies en het gebruiken van minder brandstof). Deze applicaties worden getest op de weg en daarna geëvalueerd. De effecten die gevonden worden, worden gemodelleerd en vervolgens vindt opschaling en een kosten-batenanalyse plaats. In het project is de FESTA methodologie gebruikt, en zijn de evaluerende partijen vanaf het begin betrokken geweest bij het opzetten van de experimenten (hypothesen, indicatoren en databehoeftedefiniëren) en wordt voortdurend informatie uitgewisseld over hoe de systemen nu echt in de praktijk gaan werken. Het project is dus op een goede manier gestart, maar moet in de gaten blijven houden of alle stappen goed aan bod kunnen komen. Een eerste teken dat goed tussen de oren zit dat de evaluatie tijd nodig heeft, is dat inmiddels, wegens vertragingen vroeg in het project, aangevraagd is of de projectduur verlengd kan worden.

5. Conclusies en aanbevelingen

Onze ervaring is dat de FESTA-V goed helpt bij het structureren van evaluaties (en het opzetten van proeven in zijn geheel). Dit geldt zowel voor de opzettende partij als voor een eventuele beoordelende partij – de checklist die in de Leidraad staat, werkt goed om te bepalen of wel overal aan gedacht is.

Binnen zowel Europese als nationale projecten zijn ervaringen opgedaan met betrekking tot evaluatie. Deze paper beschrijft een selectie van inzichten die zijn opgedaan. Voor een aantal van de onderwerpen (zoals piloting en de integratie van resultaten) ligt de moeilijkheid niet in het feit dat men in theorie niet weet hoe dit moet en dat het veel tijd kost, maar in het feit dat in praktijk die tijd (en/of budget) er vaak niet is om zaken zo grondig uit te voeren als eigenlijk zou moeten. Degenen die met evaluaties belast zijn, moeten daar eigenlijk wat vaker op hun strepen gaan staan en de tijd die ze nodig hebben (voor reflectie en discussie) ook echt opeisen.

Belangrijk is ook dat er genoeg aandacht is voor de relevante situationele en verklarende variabelen, omdat die veel invloed kunnen hebben op de effecten die bepaald moeten worden. Dit blijft een lastig punt voor maatregelen in het verkeer – ze zijn lastig te meten, kunnen binnen de meetperiode veel te vaak of juist totaal niet voorkomen.

Evaluaties blijven belangrijk om te bepalen of maatregelen (innovatief of al langer in gebruik) in de praktijk goed uitpakken. Ook op stedelijk niveau is dit belangrijk. In het Europese CIVITAS initiatief (www.civitas.eu) worden al jaren maatregelen geïmplementeerd in verschillende steden in Europa. Evaluaties zijn tot nog toe (te) vaak kwalitatief van aard; kwantitatieve evaluaties worden lastig gevonden, vanwege het ontbreken van expertise op dit gebied en vanwege beperkte middelen (tijd, budget) bij stedelijke wegbeheerders en verkeersmanagers. Gezien het toenemende aantal ITS maatregelen dat in de stad toegepast wordt, is het van belang ook daar een

methodologie als FESTA te omarmen en evaluatie-expertise op te bouwen. Het uiteindelijke doel is om beleidsmakers te kunnen voorzien van goed onderbouwde cijfers over baten en kosten van ITS maatregelen, en zo de beperkte beschikbare budgetten zo effectief en efficiënt mogelijk in te zetten.

Referenties

Amitran (2014-1). <http://www.amitran.eu>.

Amitran (2014-2). <http://www.amitran.eu/knowledge-base>.

Amitran (2014-3). D4.1, Requirements and design of the methodology, 30 april 2014.

Arcadis (2012). Kader Verkeerskundige Evaluatie Regelscenario's, 24 mei 2012.

Connekt (2013). Beter geïnformeerd op weg, Routekaart 2013 – 2023, Hoofddocument. In opdracht van het Ministerie van Infrastructuur en Milieu, oktober 2013.

DRIVE C2X (2014). <http://www.drive-c2x.eu>.

EC (2011). Intelligent Transport Systems in action – Action Plan and Legal Framework for the deployment of Intelligent Transport Systems (ITS) in Europe. European Commission, Directorate-General for Mobility and Transport, 2011, beschikbaar @ http://ec.europa.eu/transport/its/road/action_plan/action_plan_en.htm.

eCoMove (2014), Cost-benefit analysis, integration of evaluation results and eCoMove implementation road map, D650.65 of the eCoMove Project, Project reference FP7-ICT-2009-4, (op termijn) beschikbaar @ <http://ecomove-project.org/publications/deliverables/>.

Energy ITS & ECOSTAND (2013), "Guidelines for assessing the effects of ITS on CO2 emissions", Energy ITS and ECOSTAND projects, with contributions from the University of California, 2013, available @ <http://www.nedo.go.jp/content/100521807.pdf>.

FOT-Net consortium (2014-1). FESTA Handbook version 5, 2014, <http://fot-net.eu/Documents/festa-handbook-version-5-2014/>

FOT-Net consortium (2014-2). FOT-Net D3.2, Working Group Additions to FOT Methodology, http://fot-net.eu/download/Deliverables/fot-net_2_d3_2_addition_to_fot_methodology_v9.pdf.

MuConsult (2011). Evaluatiemethodiek benutten - Deelrapport: Richtlijnen ex-post beleidsevaluaties. Amersfoort, MuConsult, Kenmerk VW82, 22 december 2010, beschikbaar @ <http://traffic-quest.nl/nl/kennis/state-of-the-art/92-leidraden-voor-evaluaties>.

Wilmink, I., K. Malone, E. Jonkers, R. Brouwer, R. de Lange, M. Keuken, A. Eisses, N. Rosmuller & J. Mak (2011), "Leidraad evaluaties benutting (versie 2011)", Delft, 30 mei 2011, beschikbaar @ <http://traffic-quest.nl/nl/kennis/state-of-the-art/92-leidraden-voor-evaluaties>.

Witteveen & Bos (2012). Leidraad evaluatie regelscenario's. 8 Mei 2012, beschikbaar @ http://traffic-quest.nl/images/stories/documents/Evaluatie/leidraad_evaluatie_regelscenarios_versie_2012.pdf