

POPULAIR-WETENSCHAPPELIJKE SAMENVATTING (SUMMARY IN DUTCH)

DR. E.A. GROEN

1. ACHTERGROND

De productie van voedsel kan nadelige gevolgen voor het milieu hebben, zoals een bijdrage aan klimaatverandering door de uitstoot van broeikasgassen. De belangrijkste broeikasgassen in de landbouw zijn CO_2 (koolstofdioxide), N_2O (lachgas) en CH_4 (methaan). De uitstoot van broeikasgassen tijdens voedselproductie vindt bijvoorbeeld plaats bij bemesting van landbouwgrond (in de vorm van CO_2 en N_2O), verbranding van fossiele brandstoffen bij transport van veevoer ingrediënten (in de vorm van CO_2) en tijdens de mestopslag (in de vorm van CH_4). Om deze vervuiling terug te dringen, moeten eerst de desbetreffende productieprocessen in kaart gebracht worden.

Echter, productieprocessen in de landbouw zijn onderhevig aan *natuurlijke variatie*: externe invloeden, zoals van het weer en klimaat. Daarnaast zijn alle cijfers ook nog onderhevig aan *epistemische onzekerheid*: ze kunnen meetfouten bevatten, of men weet niet zeker of de cijfers die worden gebruikt ook wel exact hetzelfde productieproces weerspiegelen. De cijfers die nodig zijn als invoer van een model om de milieu-impact van een product te berekenen, zijn dus onderhevig aan onzekerheden. Dit leidt tot onzekerheid rondom de uitkomst van een milieu-impact *model*. Daarmee lijkt het meenemen van onzekerheid in eerste instantie alleen maar te leiden tot nog meer vragen: *hoe beïnvloedt natuurlijke variatie de modeluitkomst? Zijn de verschillen die ik vind in de modeluitkomst nog wel significant als ik onzekerheden meeneem van de input variabelen? Welke variabelen zijn verantwoordelijk voor de meeste spreiding rondom de uitkomst?* Over deze vragen, en nog vele andere, gaat dit proefschrift.

Een uitleg van de *schuingedrukte* terminologie en methodes, die gebruikt worden in dit proefschrift, is terug te vinden in de kaders.

Onzekerheid: in de context van dit proefschrift is het een paraplueterm voor zowel *natuurlijke variatie* (variatie door verschillen in klimaat, weersinvloeden, menselijk handelen) als *epistemische onzekerheid* (onzekerheid door gebrek aan kennis maar ook bijvoorbeeld meetfouten, of ontbrekende data). **Spreiding:** de gekwantificeerde onzekerheid rondom een datapunt van een *input variabele* of een *output variabele*. **Standaardafwijking:** spreiding van een dataset, bijvoorbeeld van de normale verdeling. **Distributiefunctie:** de kansverdeling waarmee een variabele beschreven wordt, bijvoorbeeld bij een standaard normale verdeling vallen 99.7% van de waarden van de kansverdeling binnen $\pm 3x$ de *standaardafwijking* van de gemiddelde waarde van de verdeling. **Correlatie:** (lineaire) samenhang tussen twee grootheden.

DOEL VAN DIT PROEFSCHRIFT

Het doel van mijn proefschrift is de betrouwbaarheid verbeteren van modellen die milieu-impacts kwantificeren, gebruik makend van methodes voor *onzekerheidsanalyse* en *gevoeligheidsanalyse*. Hiervoor heb ik verschillende methodes met elkaar vergeleken en expliciet gekeken naar het effect van *correlatie* op de *spreiding* rondom de uitkomst van milieu-impact modellen. Alle methodes gebruikt in dit proefschrift zijn toegepast op voedselproductiesystemen.

Gevoeligheid: hoe sterk een variabele doorwerkt op de modeluitkomst. Er bestaan drie soorten *gevoeligheidsanalyses*: (1) *lokale gevoeligheidsanalyse*: kwantificeert wat er gebeurt met de modeluitkomst als de input variabelen één voor één een heel klein beetje worden veranderd; (2) *screening analyse*: kwantificeert wat er gebeurt met de modeluitkomst als de input variabelen één voor één worden veranderd tussen hun daadwerkelijke minimum en maximum waarden; (3) *globale gevoeligheidsanalyse*: kwantificeert hoeveel van de spreiding rondom de output variabele kan worden verklaard door de input variabelen, door middel van de *gevoeligheidsindex*.

ENKELE RESULTATEN UITGELICHT

In **Hoofdstuk 2** worden twee methodes voor gevoeligheidsanalyse gecombineerd: de “multiplier methode” (een *lokale gevoeligheidsanalyse*) en de “methode van elementaire effecten” (een *screening methode*). De methodes worden toegepast op een productie systeem in de varkenshouderij, inclusief de keten voorafgaand aan de houderij, zoals de productie van de veevoer ingrediënten. De resultaten van beide methodes worden met elkaar gecombineerd om tot suggesties tot reductie van milieu-impacts te komen en om betrouwbaarheid van de modeluitkomst te verhogen.

Variabele: een grootte in het model, bijvoorbeeld brandstof (in liter), elektriciteit (in kWh) of broeikasgas (in kg CO₂). De variabelen die het model ingaan worden **input variabelen** genoemd, de modeluitkomst wordt ook wel de **output variabele** genoemd. Bijvoorbeeld, om de totale broeikasgassen van melk te bepalen is 10 liter diesel nodig. In dit geval is 10 de variabele (ook al varieert deze niet daadwerkelijk). Het kan ook zijn dat er gemiddeld 10 liter diesel nodig is, maar dat er een spreiding om dit gemiddelde ligt met een standaardafwijking van 0.5. In dat geval is de variabele gelijk aan de distributiefunctie, met een gemiddelde van 10 en een standaardafwijking van 0.5. **Model:** in de context van dit proefschrift gaat het om modellen die de milieu-impact van een productiesysteem bepalen. Deze modellen bevatten meerdere input variabelen (meestal in de orde van enkele honderden) en meestal één output variabele, bijvoorbeeld het broeikasgas CO₂.

In **Hoofdstuk 3 en 4** worden verschillende methodes voor onzekerheids- en *globale gevoeligheidsanalyse* met elkaar vergeleken op verschillende aspecten: (1) type *onzekerheidspropagatie* (wel of geen gebruik makend van *sampling*), (2) hoe goed ze zijn in het bepalen van de spreiding rondom de uitkomst, en (3) hoe goed ze zijn in het verklaren van spreiding rondom de uitkomst. Methodes die het meest geschikt zijn (binnen milieu-impact modellen) zijn *Monte Carlo simulatie* voor onzekerheidspropagatie en op regressie gebaseerde methoden voor globale gevoeligheidsanalyse, de ‘kwadratisch gestandaardiseerde regressie coefficient’ fungeert dan als de *gevoeligheidsindex*.

Onzekerheidspropagatie: het voortplanten van onzekerheden rondom input variabelen door een model, wat resulteert in onzekerheid in de output variabele van dat model. **Onzekerheidsanalyse:** analyseren van de onzekerheid van de uitkomst, bijvoorbeeld het bepalen van de spreiding, of het vergelijken van twee distributiefuncties. **Sampling:** het trekken van random getallen uit een distributiefunctie. Bijvoorbeeld: uit een uniforme verdeling tussen 0 en 1, kunnen de eerste drie samples zijn: 0.1; 0.8; 0.2. **Monte Carlo simulatie:** door middel van sampling de spreiding rondom de output variabele van een model schatten.

In sommige gevallen is er sprake van *correlatie* tussen de spreiding van twee (of meer) variabelen binnen een model. Binnen de landbouw kan men bijvoorbeeld denken aan een correlatie tussen voerinnname en melkproductie van een koe. Als de voerinnname en melkproductie beiden variëren zullen deze twee variabelen gelijk optrekken, maar niet helemaal: de temperatuur in de stal, de samenstelling van het voer, de gezondheid van de koe, kunnen ook allemaal invloed hebben op deze onderlinge afhankelijkheid. Het is echter vaak lastig om de waarde voor de correlatiecoëfficiënt boven tafel te krijgen, in sommige gevallen is deze waarde simpelweg niet aanwezig.

In **Hoofdstuk 5** wordt, zonder dat vooraf de waarde van de correlatiecoëfficiënt bekend is, bepaald in welke gevallen correlatie tussen variabelen genegeerd kan worden. Het blijkt dat, in sommige gevallen, het meenemen van correlatie geen tot weinig invloed heeft op de spreiding rondom de uitkomst, noch op de gevoeligheidsanalyse. In deze gevallen is het dus mogelijk om de correlatiecoëfficiënt te negeren, wat mogelijk veel tijd bespaart die beter besteed kan worden aan het verbeteren van de datakwaliteit van de andere variabelen. In sommige gevallen is heeft de correlatie wel een groot effect op de spreiding en de gevoeligheidsanalyse, dan is het dus wel noodzakelijk om de echte waarde van de correlatiecoëfficiënt te achterhalen. In **Hoofdstuk 6** wordt de in Hoofdstuk 5 beschreven procedure toegepast op een productiesysteem binnen de melkveehouderij. Hier blijkt dat de correlatiecoëfficiënt tussen voerinnname en melkproductie wel degelijk van belang is, en dus niet zonder meer genegeerd kan worden.

In **Hoofdstuk 7** worden de milieu-impacts van verschillende boerenbedrijven met elkaar vergeleken op het gebied van efficiënt nutriënten gebruik. Vooraf waren de distributiefuncties van alle variabelen bekend. Vervolgens werden verschillende methodes voor onzekerheids- en gevoeligheidsanalyse toegepast. De onzekerheidsanalyse liet zien dat het meenemen van epistemische onzekerheden, er in sommige gevallen toe leidde dat de bedrijven geen significant verschillend nutrient gebruik hadden. Vervolgens wordt een globale gevoeligheidsanalyse toegepast, waardoor de variabelen konden worden bepaald die het meest bijdroegen aan de spreiding van de uitkomst. Vervolgens worden de standaardafwijkingen van de belangrijkste variabelen in het model verkleind, wat leidt tot meer onderling verschillende bedrijven.

In **Hoofdstuk 8** worden enkele aanbevelingen gedaan, die zich erop richten de betrouwbaarheid van toekomstige studies met behulp van milieu-impact modellen te vergroten. Ten eerste, het standaardiseren van methoden voor onzekerheids- en gevoeligheidsanalyse in wetenschappelijke literatuur leidt er toe dat de onderlinge vergelijkbaarheid tussen studies verbetert. Ten tweede, het in tandem uitvoeren van lokale en globale gevoeligheidsanalyse geeft beter inzicht in de meest essentiële variabelen dan het toepassen van slechts één methode, met name voor variabelen waarvoor het niet mogelijk is om een spreiding te vinden. Ten derde, door vooraf de grootte van het effect van correlaties tussen input variabelen op de modeluitkomst te bepalen (zoals beschreven in Hoofdstuk 5), kan het risico van het negeren op de spreiding rondom de modeluitkomst worden bepaald.

Deze drie punten dragen bij aan het vergroten van de betrouwbaarheid van milieu-impact modellen, en dragen daarmee bij aan het formuleren van strategieën om toekomstige milieu-impact van voedselproductie verder te reduceren.