**MODELLEREN EN SIMULEREN: HOOFDSTUK 5: Modelanalyse**

1. Inleiding

* Model geeft slechts een benaderende beschrijving van de werkelijkheid
  + => Vaak is het niet voldoende om te zeggen of een zeker set van parameters een goede fit oplevert of niet
  + => We moeten ook weten of het model (gecombineerd met de parameterset) robuust is, dwz we moeten een antwoord ku formuleren op de volgende vragen:
    - 1) Als onze metingen of berekeningen een kleine afwijking hebben tov de modelparameters, zullen de voorspellingen vh gebruikte model dan ook een kleine afwijking hebben of zullen ze drastisch verschillen?
    - 2) Hoe zeker kunnen we zijn vd gekozen parameters & in welke mate hangt het gedrag vh model af van deze keuzes?
* Modelanalyse
  + Biedt een antwoord op de vorige vragen
  + Geeft aandacht aan 2 soorten analyses:
    - 1) Gevoeligheidsanalyse = sensity analysis
      * = de studie naar de invloed van variaties in modelparameters/beginvoorwaarden, enz. op modelresultaten
      * Hoe gevoelig is het model als we parameters etc variëren?
    - 2) Onzekerheidsanalyse = uncertainty analysis
      * = de studie naar de onzekere aspecten v/e model & naar hun invloed op de inherente onzekerheid vd modelresultaten
      * Hoe werken deze onzekerheden door in modelresultaten?!
  + OPM: de begrippen gevoeligheidsanalyse & onzekerheidsanalyse w in literatuur vaak door elkaar gebruikt
    - Toch enkele verschillen:
    - In gevoeligheidsanalyse: kijken naar gevoeligheden voor variaties in parameters, beginvoorwaarden & inputs
    - In onzekerheidsanalyse: kansen/onzekerheden waarmee die variaties optreden bekijken => komt niet aanbod in gevoeligheidsanalyse
      * Doel: onzekerheid vd modeloutput te kwantificeren die het gevolg is vd onzekerheden vh modelleerproces

2. Gevoeligheidsanalyse

* Gevoeligheidsanalyse onderscheid 2 technieken:
  + Lokale gevoeligheidsanalyse: de gevoeligheidsanalyse gebeurt lokaal rond 1 bepaald punt vd parameterruimte
    - => de invloed vh veranderen vd parameterwaarde w lokaal bepaald
  + Globale gevoeligheidsanalyse: de invloed v/e parameter over de gehele parameterruimte wordt bepaald
  + => elke techniek voor & nadelen => ku geen uitspraak maken over welke beste is

2.1 Differentiële gevoeligheidsanalyse

* Differentiële gevoeligheidsanalyse
  + = klassieke techniek voor gevoeligheidsanalyse uit te voeren
  + = een vorm van lokale gevoeligheidsanalyse waarbij de invloed v/e parameter in 1 bepaald punt vd parameterruimte wordt bepaald
  + De gevoeligheid
    - = de verandering vd variabele tov de verandering vd parameter
  + 1) De partiële afgeleide vd variabele naar de parameter (1 getal)
    - = de gevoeligheid vd variabele y tov de parameter

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + - * Als verandert zal y ook veranderen:
  + 2) Voor dynamische modellen
    - Variabele y varieert in de tijd => dus de gevoeligheid varieert ook in de tijd
    - Gevolg: de gevoeligheid w gegeven door een **gevoeligheidsfunctie** ipv 1 getal
    - Gevoeligheidsfunctie:
      * Gevoeligheid is functie vd tijd

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + 3) Gevoeligheid vd variabele y tov meerdere parameters (j=1,….p) te bepalen => gebruik maken van meerdere gevoeligheidsfuncties
    - (afzonderlijke partiele afgeleiden)
  + Gevoeligheidsfunctie bepalen:
    - => y(t) wordt numeriek bepaald => hierdoor ook de S(t) numeriek bepalen
    - => gebruik maken van de eindige verschillenmethode die de lineariteit vh model verondersteld

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + - * => partiele afgeleide wordt benaderd!
      * = de perturbatie vd parameter
      * Voorwaartse verschil gebruikt => dus positieve perturbatie
* Voorbeeld: gevoeligheidsfunctie obv de positieve perturbatie ppt p215
  + Links: de modeloutput y(t) ifv de tijd voor de parameterwaarde en +
  + Rechts: de dynamische gevoeligheidsfunctie obv de positieve perturbatie
  + => y plotten voor elke waarde van in fie vd tijd
  + => dan een beetje verhogen => output verandert een beetje
  + => bij grote verschillen tssn en +∆ => grote gevoeligheidsfie en omgekeerd
* Voorbeeld: gevoeligheidsfunctie obv de negatieve perturbatie (achterwaarts verschil)
* Voorbeeld: gevoeligheidsfunctie obv de centrale perturbatie (centrale verschil)
  + Centraal verschil = verschil tssn pos. en neg perturbatie ongeveer => zie formule
  + Links: modeloutput y(t) ifv de tijd voor parameterwaarden + en -
  + Rechts: de dynamische gevoeligheidsfunctie obv de centrale perturbatie

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

2.2 Relatieve gevoeligheid

* Differentiële gevoeligheid
  + De hoger beschreven gevoeligheid S is afh. vd absolute grootte van zowel de variabele y als de parameters (j=1,….p) gebruikt bij berekening vd gevoeligheidsfie
  + Differentiele gevoeligheid = absolute gevoeligheid
  + Nadeel: versch gevoeligheidsfuncties ku niet onderling vergeleken worden
    - Oplossing: relatieve gevoeligheidsfuncties
* Relatieve gevoeligheid
  + Afbeelding met schermafbeelding

    Automatisch gegenereerde beschrijving1) Relatieve gevoeligheid vd output y(t) tov een parameter
    - Voordeel: laat toe om de gevoeligheid van eenzelfde variabele tov verschillende parameters onderling te vergelijken
  + Afbeelding met schermafbeelding

    Automatisch gegenereerde beschrijving2) relatieve gevoeligheid tov variabele
    - Voordeel: laat toe om de gevoeligheid van versch variabelen tov dezelfde parameter onderling te vergelijken
      * => nagaan welke variabele het meest gevoelig is voor bep parameter
  + Afbeelding met schermafbeelding

    Automatisch gegenereerde beschrijving3) totale relatieve gevoeligheid
    - Voordeel: laat toe om alle combinaties van gevoeligheidsfuncties van alle gekozen variabelen & parameters te vglen
    - = combo van de 2 vorige
* Rangschikken parameters in model
  + De relatieve gevoeligheidsfuncties ku worden gebruikt om versch parameters in een model te rangschikken volgens de gevoeligheid vh model
  + Stel generiek model met
    - Uitgangsvariabelen yi (i=1,….v)
    - Parameters (j=1,……p)
    - Tijdstippen van metingen tk (k=1,…..n) (n tijdstippen)
    - Vb: model 3 parameters, welke parameter meest gevoelig? => rangschikken
      * Reden: de gevoeligste parameter moet zo nauwkeurig mogelijk bepaald worden
  + Methode:
    - 1) totale relatieve gevoeligheid vd variabele yi tov parameter op tijdstip tk

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + - 2) om de versch parameters te rangschikken volgens gevoeligheid => moet de gevoeligheid van alle variabelen yi (i=1,…v) tov de parameter in rekening worden gebracht
      * Vb: door kwadratisch gemiddelde te nemen vd totale relatieve gevoeligheden Si,j,k op dat tijdstip voor versch variabelen

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + - 3) bij dynamische modellen (met dynamische gevoeligheidsfuncties) kan de gevoeligheid op versch tijdstippen tk sterk variëren => men moet het gemiddelde van op versch tijdstippen nemen

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + Conclusie: op deze manier w voor alle parameters 1 enkele maat voor de gevoeligheid v/e parameter bekomen & ku deze gebruikt w om de belangrijkheid v/e parameter te kwantificeren

2.3 Lokaal karakter van de gevoeligheidsfuncties

* Lokaal karakter vd gevoeligheidsfuncties
  + Lokale gevoeligheidsfunctie w berekend door een parameter te perturberen rond een bep. waarde
  + De keuze van deze parameterwaarde => kan grote invloed op resultaat vd gevoeligheidsanalyse hebben
* Vb: 2 functies
  + Y=ax+b: lineair in parameters
  + Y=a2x +b: niet lineair in de parameter a
  + Figuur: illustreert lokale karakter vd gevoeligheidsfie voor lineair & niet lineair model
    - Voor beide functies weergave vh verloop van y in functie vd parameter a bij constante waarden van x=1 en b=1
    - De lokale gevoeligheidsfunctie naar parameter a (=) = de rico vd raaklijn aan deze curves in een bepaald punt vd parameterruimte
  + Voor lineaire functie: de rico is overal dezelfde
    - Gevolg: de lokale gevoeligheidsfunctie is dus onafh vd plaats in de parameterruimte
    - De gevoeligheid = de rico
  + Voor niet lineaire functie: de rico (& dus de lokale gevoeligheidsfunctie) zal wijzigen afhankelijk vd waarde van a waarrond men werkt
    - Voor kleine waarden van a = lokale gevoeligheid klein
    - Voor grote waarden van a = lokale gevoeligheid groot !!!!

2.4 Globale gevoeligheidsanalyse

* Aangezien de gevoeligheid w bepaald door de plaats in de parameterruimte waar men de analyse uitvoert => kan de gevoeligheidsanalyse versch resultaten voor & na parameterschatting opleveren
  + Na parameterschatting komt men immers elders in de parameterruimte terecht => waardoor de uitgevoerd gevoeligheidsanalyse waardeloos is, want ze bevat geen info over het model in het nieuwe punt vd parameterruimte
  + Dus de lokale gevoeligheidsanalyse probleem = positie bepaald de gevoeligheid
* Globale gevoeligheidsanalyse
  + = om de gevoeligheid in de gehele parameterruimte te onderzoeken
  + => vaak gebaseerd op Monte-Carlo simulatietechnieken

2.5 Monte-Carlo simulatietechniek

* Monte-Carlo simulatietechniek
  + = simulatietechniek waarbij een fysiek proces vele malen w gestimuleerd, telkens met andere startcondities of parameterwaarden
  + = dus gebaseerd op groot aantal simulaties door herhaaldelijk random te samplen uit een vooropgestelde verzameling => om zo tot resultaat te komen op bep vraag
  + Verdelingsfunctie
    - = het resultaat van deze verzameling simulaties, die het hele gebied van mogelijke uitkomsten weergeeft
  + Wordt toegepast wanneer:
    - 1) Er een grote variatie of onzekerheid bestaat met betrekking tot de startcondities
      * Resultaat van 1 enkele simulatie is dan niet voldoende representatief
      * Dus veel simulaties met versch startcondities
    - 2) Als de variatie of onzekerheid van die startcondities bekend is of met voldoende betrouwbaarheid ingeschat & gekwantificeerd kan worden
      * Dit doen bij onzekerheidsanalyse
  + Kenmerken
    - Monte-carlo simulatie koppelt maw een onzekeheid/ variatie van inputs & parameters met een deterministisch model
    - Monte-Carlotechnieken w vaak gebruikt voor gevoeligheidsanalyses en onzekerheidsanalyses
  + Bestaat uit 3 fasen/ stappen

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + - Fase 1 onderste alinea: parameters kiezen met grote kans van voorkomen
  + Besluit
    - Monte-Carlosimulatie techniek kan dus beschouwd worden als een groot aantal (analytische of numerieke) oplossingen v/e deterministisch model
      * => om zo stochastische eigenschap vh model te onderzoeken
    - Monte-Carlosimulatie technieken w vooral gebruikt voor het uitvoeren van gevoeligheidsanalyses & onzekerheidsanalyses
  + Overzicht ppt p227
    - Stap 1) 2 inputs, 3 parameters => gekenmerkt door een kansverdeling (dus niet 1 getal) => 1 waarde komt veel voor = de nominale waarde
    - Stap 2) algoritme gaat bep waarde selecteren in gebied inputs (gebied met grootste kans) en onderzoeken => 1 waarde voor elke input en elke parameter bestuderen => 1000en simulaties => krijg je groot aantal outputs die ook op een bep manier verdeeld zijn => min, max,.. bepalen

2.6 De methode vd gestandaardiseerde regressiecoëfficiënten

* Methode vd gestandaardiseerde regressiecoëfficiënten
  + = globale gevoeligheidsanalyse gebaseerd op (combinatie)
    - 1) Monte-Carlosimulaties
    - 2) Regressieanalyse
* Stel: model met p stochastische parameters (parameterruimte heeft dimensie p), waarmee n simulaties werden uitgevoerd
  + Fig: de dynamische, tijdsafhankelijke modeloutput y(t) van deze n simulaties
  + Op elk tijdstip kan lineaire regressie w uitgevoerd vd n simulaties:

Afbeelding met schermafbeelding, computer

Automatisch gegenereerde beschrijving

Afbeelding met schermafbeelding, computer

Automatisch gegenereerde beschrijving

* + - Met b= de vector met regressiecoefficienten & e= foutenvector
  + Door stelsel op te lossen naar b => vinden we de waarden voor de regressiecoeff
    - De gevonden waarden bi geven respectievelijk de bijdrage vd parameters voor het verklaren vd variantie van y
  + Analoog aan de relatieve gevoeligheden => dient men de waarden te corrigeren voor de spreiding op zowel de parameter als de variabele
    - Door de waarden te vermenigvuldigen met de resp. standaardafwijkingen in de parameter en de variabele
    - De gestandaardiseerde regressie-coëfficiënten
      * = deze gecorrigeerde waarden
      * = maat voor de globale gevoeligheid vd variabele tov een parameter

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

3. Onzekerheidsanalyse

* Eerder vermeld: zie ppt
* Onzekerheidsanalyse = uncertainty analysis
  + Tijdens de versch stappen v/e modelleerproces kunnen er onzekerheden in het model sluipen => die zorgen er uiteindelijk voor dat de output vh model een zekere inherente onzekerheid met zich meedraagt
  + Stelt de vraag: Hoe zeker kunnen we zijn vd gekozen parameters & in welke mate hangt het gedrag vh model af van deze keuzes?
  + = de studie naar de onzekere aspecten v/e model & naar hun invloed op de inherente onzekerheid van de modelresultaten

3.1 Toepassingen

* Toepassingen onzekerheidsanalyse
  + 1) Validatie van modellen
    - Modellen w gevalideerd adhv onafhankelijke dataset
      * De modeloutput zal echter nooit perfect de dataset beschrijven
    - Een betrouwbaarheidsinterval (rond de output) geeft de onzekerheid aan waarmee het model zou moeten overeenkomen met de validatieset
  + 2) Besluitvorming
    - Vaak de modeloutput vergelijken met een waarde vb: wettelijke norm
    - Als betrouwbaarheidsmodel rond modeloutput is => kan besluit nemen met bepaalde zekerheid
    - Vb: als een 95% betrouwbaarheidsinterval volledig onder de wettelijke norm zit => kan men met zekerheid stellen dat in 95% vd gevallen de modeloutput kleiner zal zijn dan deze waarde & dus onder norm zal liggen
  + 3) Voorspellingen
    - Modellen gebruikt om voorspellingen te doen
    - Vb: weervoorspellingen, voorspellingen van klimaatveranderingen etc
    - Vb: de tweewekelijkse temperatuursvoorspelling (ppt p235)
      * = voorspelling met zekere onzekerheid
      * => betrouwbaarheidsinterval meegeven => onzekerheid groter naarmate verder in de toekomst
    - Vb: voorspellingen vd globale T stijging of CO2 concentraties

3.2 Bronnen van onzekerheid

* Bron 1: Onzekerheid in de modelstructuur
  + Elk model is een benaderende beschrijving vd werkelijkheid
    - Gevolg: bij vastleggen vd modelstructuur maken we sws een fout
  + Modellen onderscheiden zich op vlak van:
    - 1) Het aantal processen en het aantal toestandsvariabelen dat wordt meegenomen in de systeembeschrijving
      * Als een cruciaal proces niet in rekening w gebracht => grote gevolgen op de modeloutput
    - 2) De complexiteit en mate van detail waarmee de processen beschreven w
      * Hoe complexer het model => hoe kleiner de onzekerheid
      * Echter: meer detail kan ook problematisch zijn
        + Vb: gedetailleerde modellen vereisen veel parameters & te veel parameters levert probz op voor parameterschatting
    - 3) Aantal onafhankelijke variabelen
      * Vaak 2 onafhankelijke variabelen: tijd én ruimte!
      * In combinatie met een dynamisch model => levert dit partiele DV op
        + Dit kan men omzeilen door de ruimtelijke component te discretiseren waardoor de partiele DV w omgezet in een set van gewone DV
      * Discretietechnieken op zich zijn ook een bron van onzekerheid
        + = vb eindige verschillen methode = benadering
  + Vaak moeilijk kwantificeerbaar
  + Typische indicaties v/e onvolledige modelstructuur of gebrek aan detail:
    - 1) optreden van duidelijke patronen in de residuen
    - 2) parameterschattingen die variëren naargelang het tijdsinterval dat gekozen w voor de schatting
  + Oplossingen om modelonzekerheid te identificeren
    - 1) opstellen van versch kandidaatmodellen voor hetzelfde systeem
    - 2) variatie vd modeloutput vd versch modellen geeft dan een idee vd onzekerheid
* Bron 2: Onzekerheden in de programmering
  + = gerelateerd aan de implementatie vh model in een modelleer of programmeertaal
  + 1) Kunnen fouten optreden bij schrijven van code of bugs in software
  + 2) Numerieke benaderingen & afrondingsfouten ku van software tot software verschillen
  + Oplossing: kwaliteitscontrole
    - Dit kan bestaan uit ringtesten => hierbij laat men het model door versch programmeurs in versch softwarepaketten opstellen => vervolgens w de versch modeloutputs vergeleken
* Bron 3: Meetfouten
  + Met ELKE meting is een meetfout geassocieerd
  + Meetfouten w uitgedrukt adhv de standaardafwijking
  + Meetfouten zijn eigen aan sensoren en meetmethodes
    - Maar ook menselijke fouten treden op
  + Meetfouten geven op 2 manieren aanleiding tot modelonzekerheid
    - 1) parameterschatting: de metingen die gebruikt w voor de parameterschatting (om model te kalibreren) zijn onderhevig aan meetfouten & vormen de basis voor het bepalen vd geschatte parameters
    - 2) Modelinput: Modelinput is de bepalende factor voor de oplossing => dus fouten hierin geven aanleiding tot een afwijkende oplossing
* Bron 4: Onzekerheid in de modelinput
  + = onzekerheid in *toekomstige* input => daardoor niet te verwarren met de eerder vermelde onzekerheid bij de parameterschatting/ modelkalibratie & modelvalidatie
  + Vb: klimaatmodellen
    - maken gebruik van toekomstige input die vaak niet gekend is
    - oplossing: werken met meerdere scenario’s
* Bron 5: Onzekerheid in de modelparameters
  + (Geschatte) parameters worden gekenmerkt door betrouwbaarheidsinterval
    - => hiermee moet rekening worden gehouden
  + Vb: constante vs variabele parameters (vb: temperatuursafhankelijkheid)

3.3 Onzekerheidspropagatie

* Onzekerheidspropagatie /foutenpropagatie
  + Een onzekerheidsanalyse behandelt de voortgang of voortplanting vd onzekerheid doorheen het model
    - Dergelijke analyse = onzekerheidspropagatie /foutenpropagatie
  + = het analyseren vd effect vd onzekerheden van variabelen (of fouten, meer in het bijzonder willekeurige fouten) op de onzekerheid v/e functie die daarop is gebaseerd Vb: de output v/e model
  + Doel: de grootte vd onzekerheid op de output bepalen obv de grootte & de distributie vd onzekerheid aanwezig in: het model, de inputs, de parameters
  + Verschillende technieken om een onzekerheidspropagatie uit te voeren
    - Klassieke foutenpropagatie, lineaire onzekerheidspropagatie & Monte-carlo onzekerheidspropagatie
* Fouten en afwijkingen: begrippen
  + Afbeelding met schermafbeelding

    Automatisch gegenereerde beschrijvingVgl 1: yi= gemeten waarde, ygem = werkelijke waarde
* Klassieke foutenpropagatie: algemene rekenregels
  + Aantal algemene rekenregels om foutenpropagatie te berekenen

Afbeelding met schermafbeelding

Automatisch gegenereerde beschrijving

* + Vgl 1: AF op een som of een verschil = de som vd absolute fouten op de termen
  + Vgl 3: RF op product of een quotient is de som vd relatieve fouten op de termen
* Monte-Carlo onzekerheidspropagatie
  + Monte-Carlosimulaties gebruikt voor het bestuderen van onzekerheidspropagaties
    - Monte-Carlo simulatie zie eerder
  + Voordeel tov andere technieken: zowel voor lineaire én niet lineaire modellen
  + Methode
    - De inputs en/of parameters zijn door hun inherente onzekerheid eerder stochastisch dan deterministisch => ku worden voorgesteld door stochastische vglen
      * Hierdoor kan men het model voorstellen door oneindige set van (deterministische) vglen
    - Na uitvoeren van een groot aantal simulaties, waarbij de inputs en/of parameters w voorgesteld door een stochastische verdeling, kan uitspraak worden gedaan over de verdeling vd output (& dus over de onzekerheid op de output)
    - Resultaat v/e Monte-Carlo simulatie v/e **dynamisch model of systeem** = n tijdreeksen
      * Hieruit kan men statistische grootheden (gem, mediaan, variantie & percentielen) berekenen
    - Resultaat v/e Monte-Carlo simulatie v/e **statisch model of systeem** = n (verschillende) outputs
      * Deze worden vaak voorgesteld door:
        + 1) histogrammen
        + 2) probability density functions (PDF)
        + 3) cumulatieve distributiefuncties (CDF)
      * => ook hier kan idee w gevormd vd verdeling en onzekerheid van de modeloutput
      * Vb: normaal verdeelde output p251
      * Vb: in geval van biosystemen: geen normaalverdeling p252
  + Nadeel Monte-Carlo onzekerheidspropagatie: het aandeel van elk afzonderlijke onzekerheidsbron in de onzekerheid vd output kan niet worden bepaald
    - Alle onzekerheden w gebundeld in de verdeling vd output
    - Maw nadeel is dat je geen info eruit kan halen over waar de onzekerheid vandaan komt, want de onzekerheden worden gebundeld