

# Statistisch benadering van testen

The logo for Refis features the word "Refis" in a large, black, sans-serif font. The letter "e" is stylized with a diagonal slash. The text is set against a light blue background that has a curved, wave-like shape on the right side.

# Refis

system reliability engineering

Merellaan 5  
3722 AK Bilthoven  
tel: +31(0)30 225 36 37  
fax: +31(0)30 225 36 49  
[www.refis.nl](http://www.refis.nl)  
[info@refis.nl](mailto:info@refis.nl)



Ministerie van Verkeer en Waterstaat /  
Bouwdienst Rijkswaterstaat

## Inhoud

1. Principes van softwarebetrouwbaarheid
2. Case Rijkswaterstaat
3. Toegevoegde waarde
4. Samenvatting en conclusie

# Testen is ....

Het opsporen en (laten) verhelpen van fouten in een informatiesysteem

# Maar:

Niet alleen de fouten die gevonden worden zijn interessant, de fouten die *niet* gevonden worden zijn minstens zo belangrijk!

## Uitgangspunt:

Naarmate er fouten worden gevonden en opgelost, neemt de betrouwbaarheid van een systeem toe.

## Definitie betrouwbaarheid (*Reliability*):

“De kans dat het systeem gedurende een bepaalde tijd, onder gelijkblijvende omstandigheden, foutloos functioneert.”

*Betrouwbaarheid:  $R(t)$*

## Faalkans

“De kans dat in het systeem in een bepaalde tijdsperiode een fout optreedt”

*Faalkans:  $F(t)$*

$$R(t) = 1 - F(t)$$

## Faalintensiteit

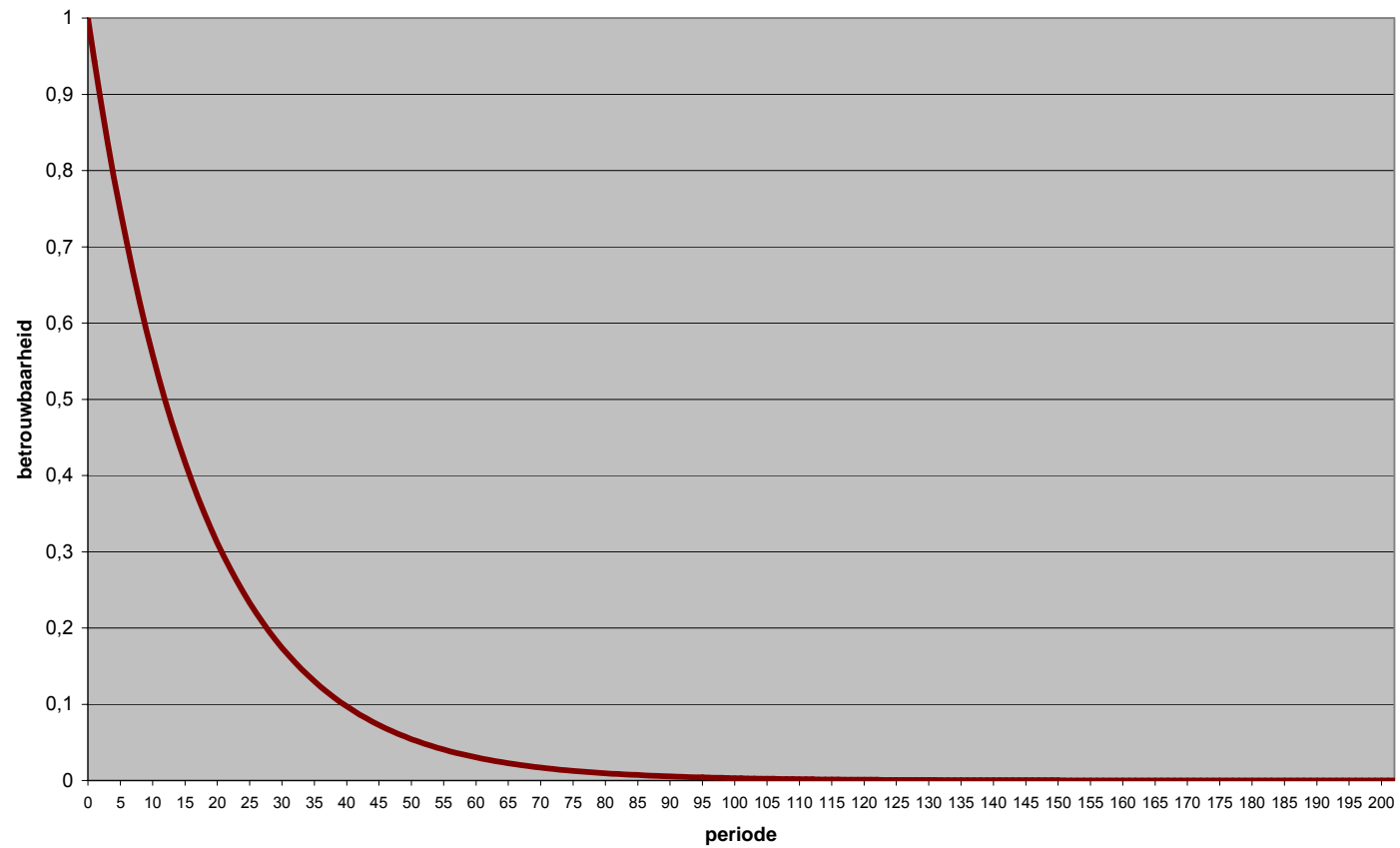
“Het aantal optredende fouten per tijdseenheid”

Faalintensiteit:  $\lambda$

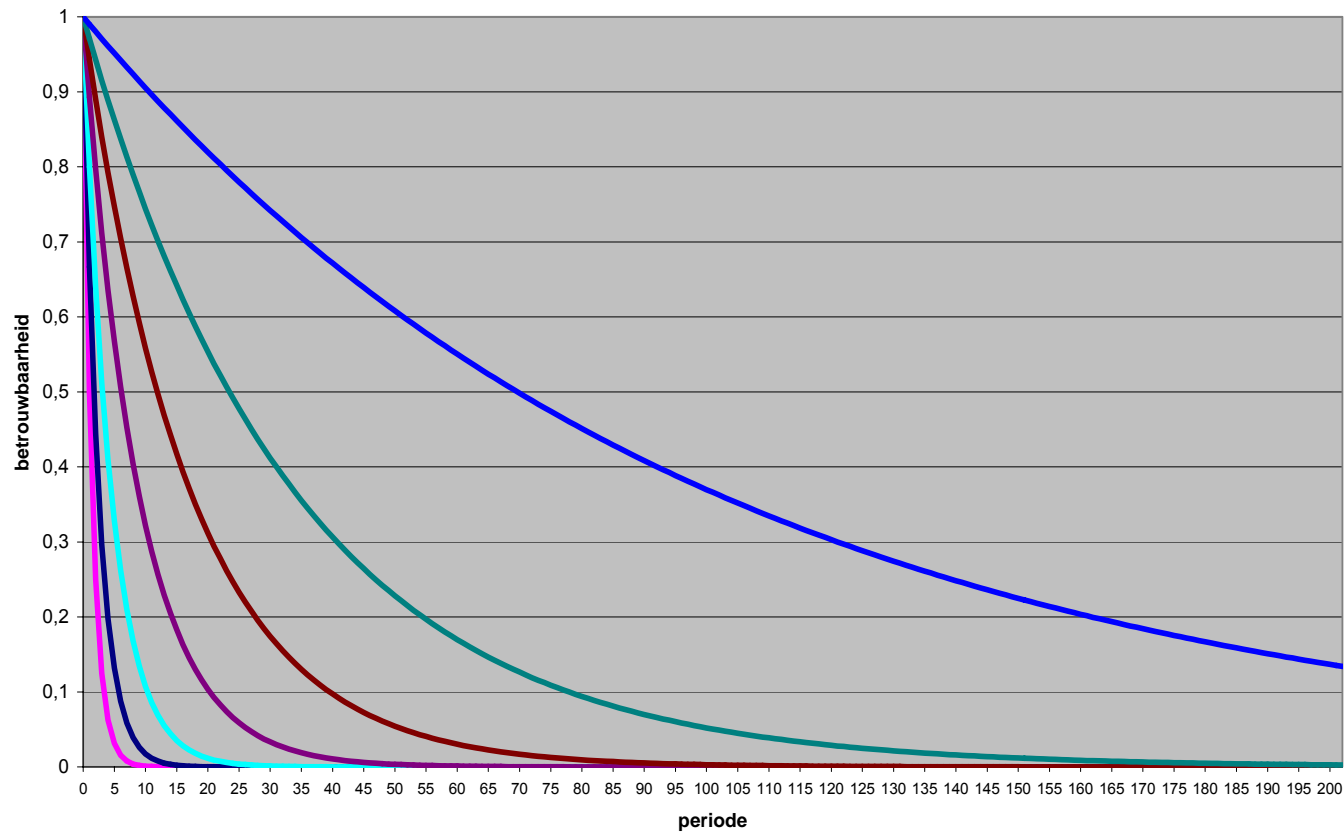
## Betrouwbaarheid

- ◆ Weergegeven als getal tussen 0 en 1:
  - ◆ de betrouwbaarheid van het systeem voor de komende 1 seconde, is bijna 1
  - ◆ de betrouwbaarheid van het systeem voor de komende 100 jaar, is bijna 0

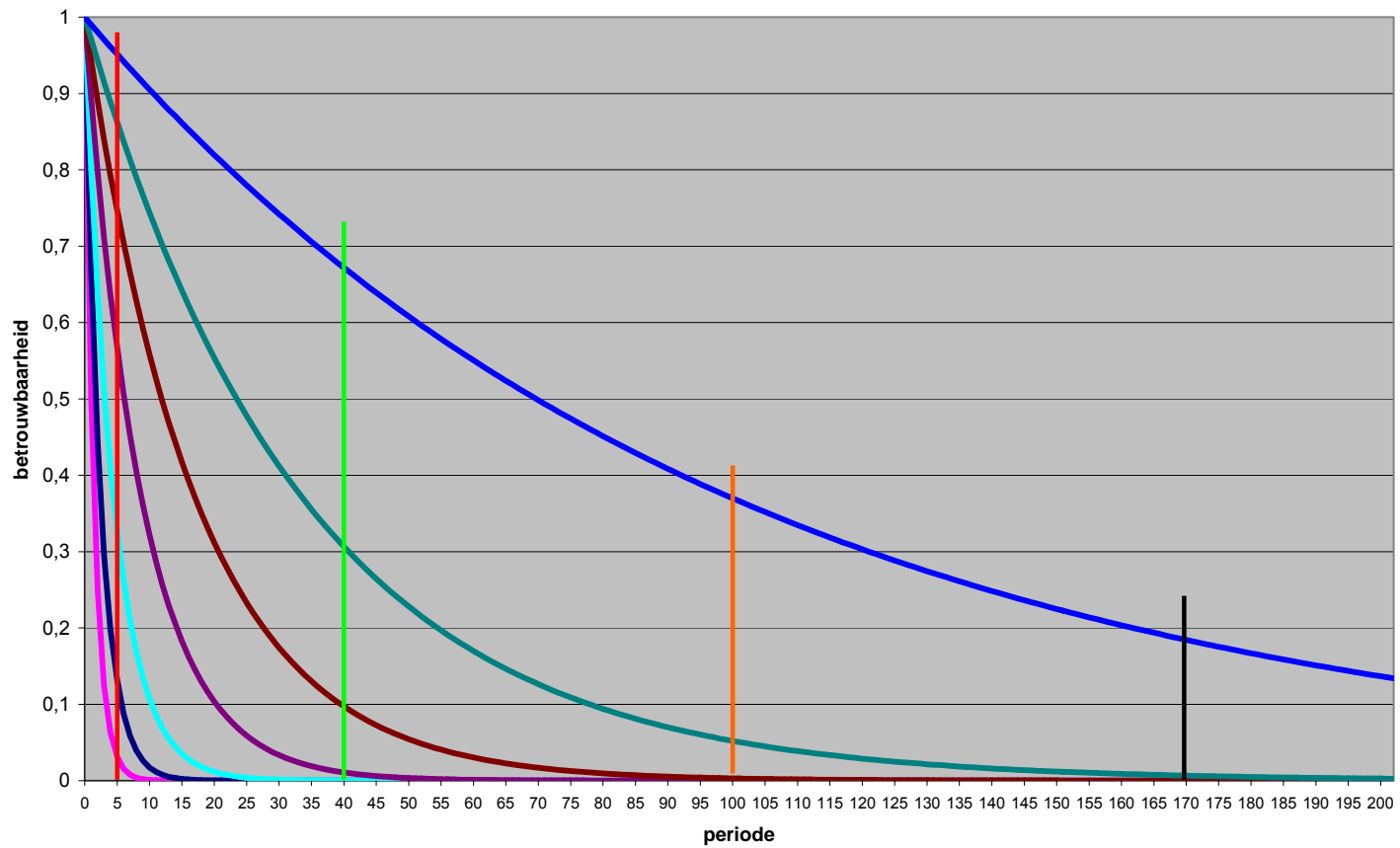
## Initiële betrouwbaarheid:



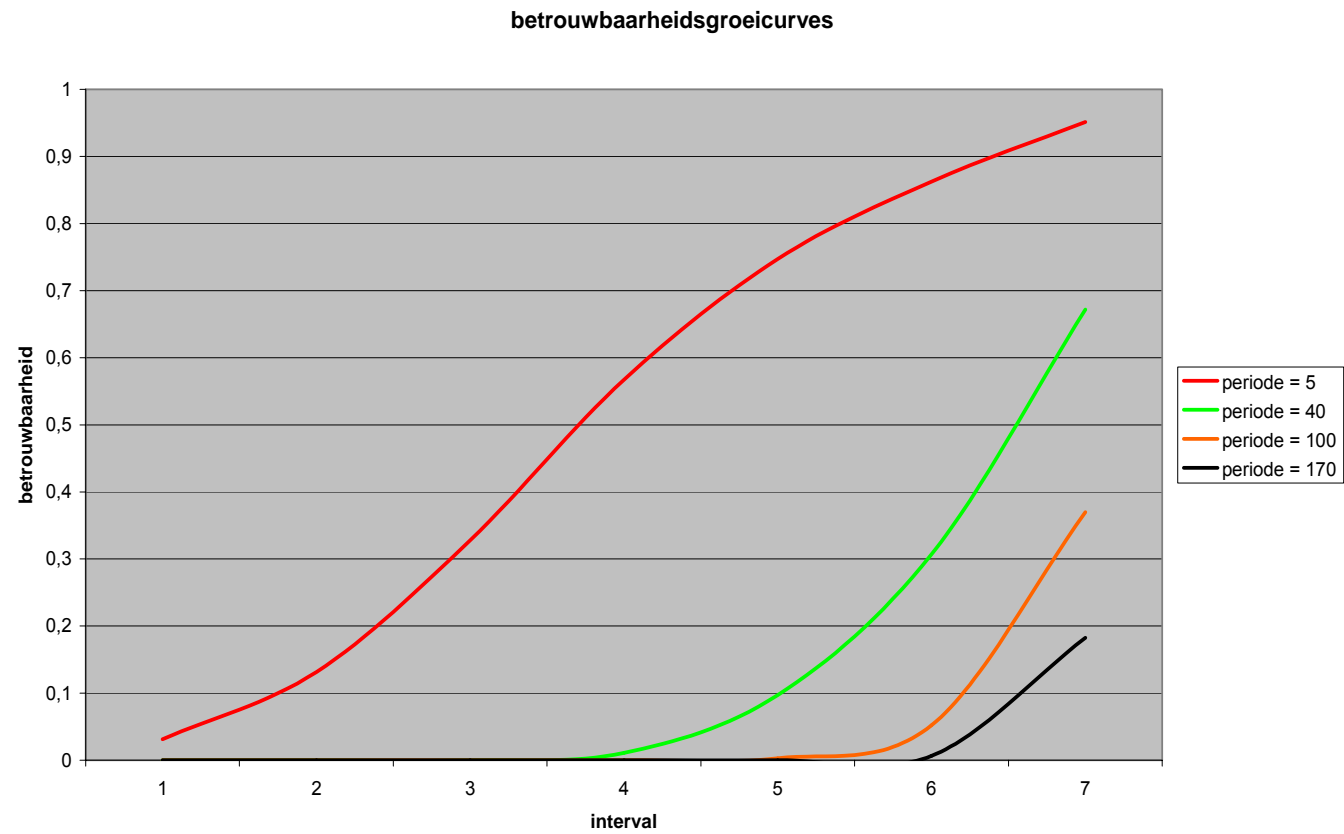
## Betrouwbaarheid na n testintervallen:



## Betrouwbaarheid na n intervallen:



# Betrouwbaarheidsgroei-curves



## Modelleren

Gepoogd wordt de betrouwbaarheidsgroei te beschrijven middels mathematische modellen.

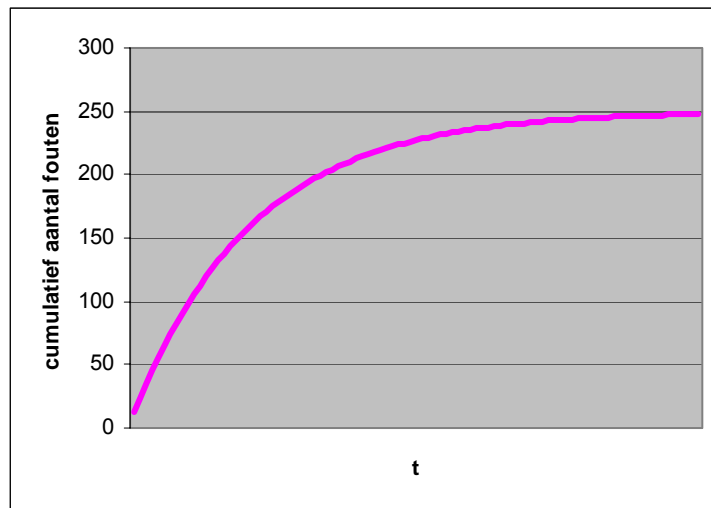
## Voorbeelden

- ◆ Jelinski-Moranda (1972)
- ◆ Geometric (1975)
- ◆ Musa-Okomoto (1984)
- ◆ Niet-homogeen Poisson (NHPP, 1979)
- ◆ Yamada S-shape (1983)
- ◆ Schneidewind (1975)
- ◆ .....

## Verschillende uitgangspunten

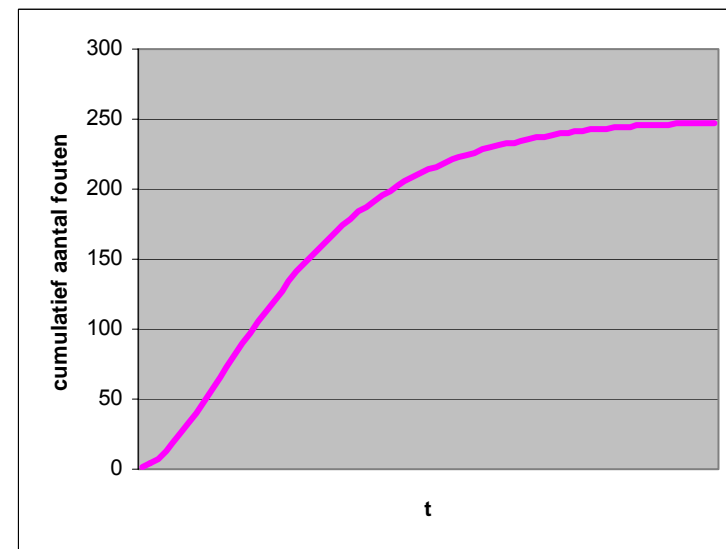
- ◆ al of niet gelijkmatige spreiding van defects
- ◆ al of niet onafhankelijkheid van defects
- ◆ al of niet “perfect debugging”
- ◆ invloed van oplossen defects op faalintensiteit
- ◆ etc. etc.

# Grafieken



cumulatief aantal fouten NHPP

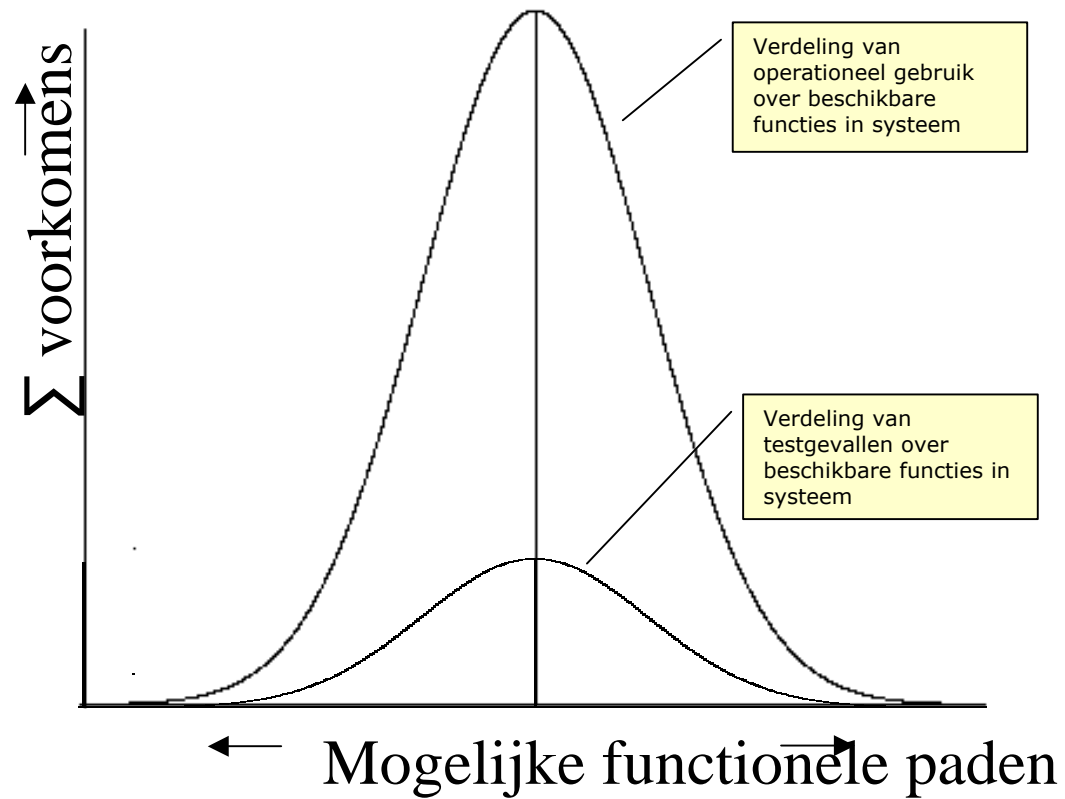
cumulatief aantal fouten Yamada



## Eisen voor toepasbaarheid

- ◆ testen conform gebruik in productie
- ◆ een goede defectregistratie
- ◆ een goede tijdregistratie
- ◆ voldoende fouten in de test

# Testen conform gebruik in productie



## Defectregistratie

- ◆ registratie van tijdstip optreden defects
- ◆ registratie van aard en ernst

## Tijdregistratie

- ◆ start- en eindtijd testen
- ◆ aantal test uren per testinterval

## Inhoud

1. Principes van softwarebetrouwbaarheid
2. Case Rijkswaterstaat
3. Toegevoegde waarde
4. Samenvatting en conclusie

## Case Rijkswaterstaat: de opdracht

Voer een betrouwbaarheidsanalyse uit op de besturingssoftware van een van de civiel-technische constructies van Rijkswaterstaat.



## Doel

Bepaal de faalkans van de software die het openen en sluiten van de constructie aanstuurt.

### Falen:

- ◆ niet sluiten terwijl dat wel moet
- ◆ sluiten terwijl dit niet moet
- ◆ openen terwijl dit niet moet
- ◆ niet openen terwijl dit wel moet

## Beschikbare gegevens

- ◆ defectgegevens uit 2 opeenvolgende testtrajecten
- ◆ aantal defects per testdag
- ◆ aard en ernst van defects

## Aannames

- ◆ Getest conform verwachte productie situatie
- ◆ Alle testintervallen van gelijke lengte (testdagen van 8 uur)
- ◆ Defects treden onafhankelijk van elkaar op
- ◆ Perfect debugging
- ◆ Testresultaten samen nemen

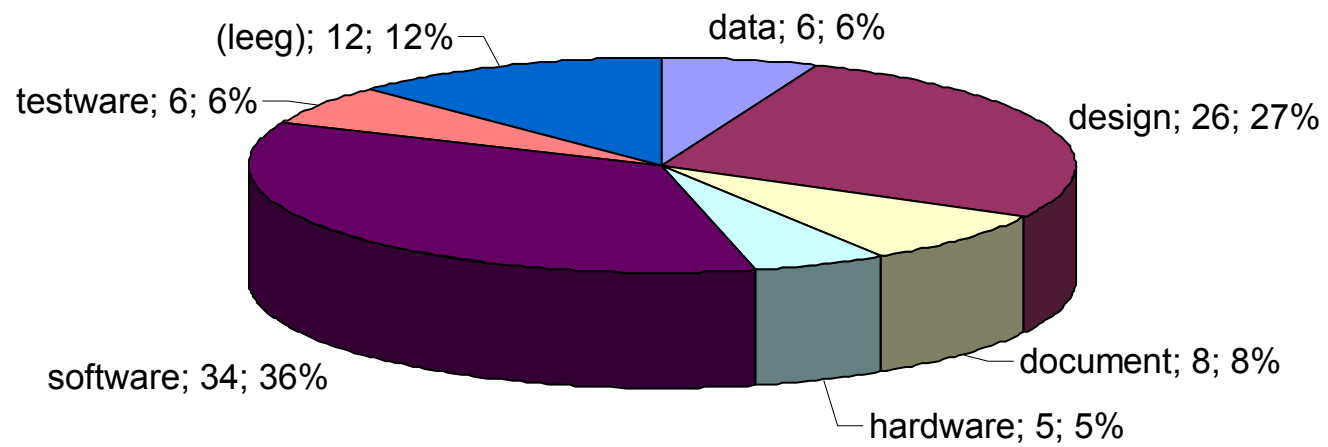
## Uitgevoerde stappen

- ◆ Filteren defectgegevens
- ◆ Vaststellen betrouwbaarheidsgroei
- ◆ Modelselectie
- ◆ Instellen modelparameters
- ◆ Modelberekeningen uitvoeren
- ◆ Interpretatie van resultaten

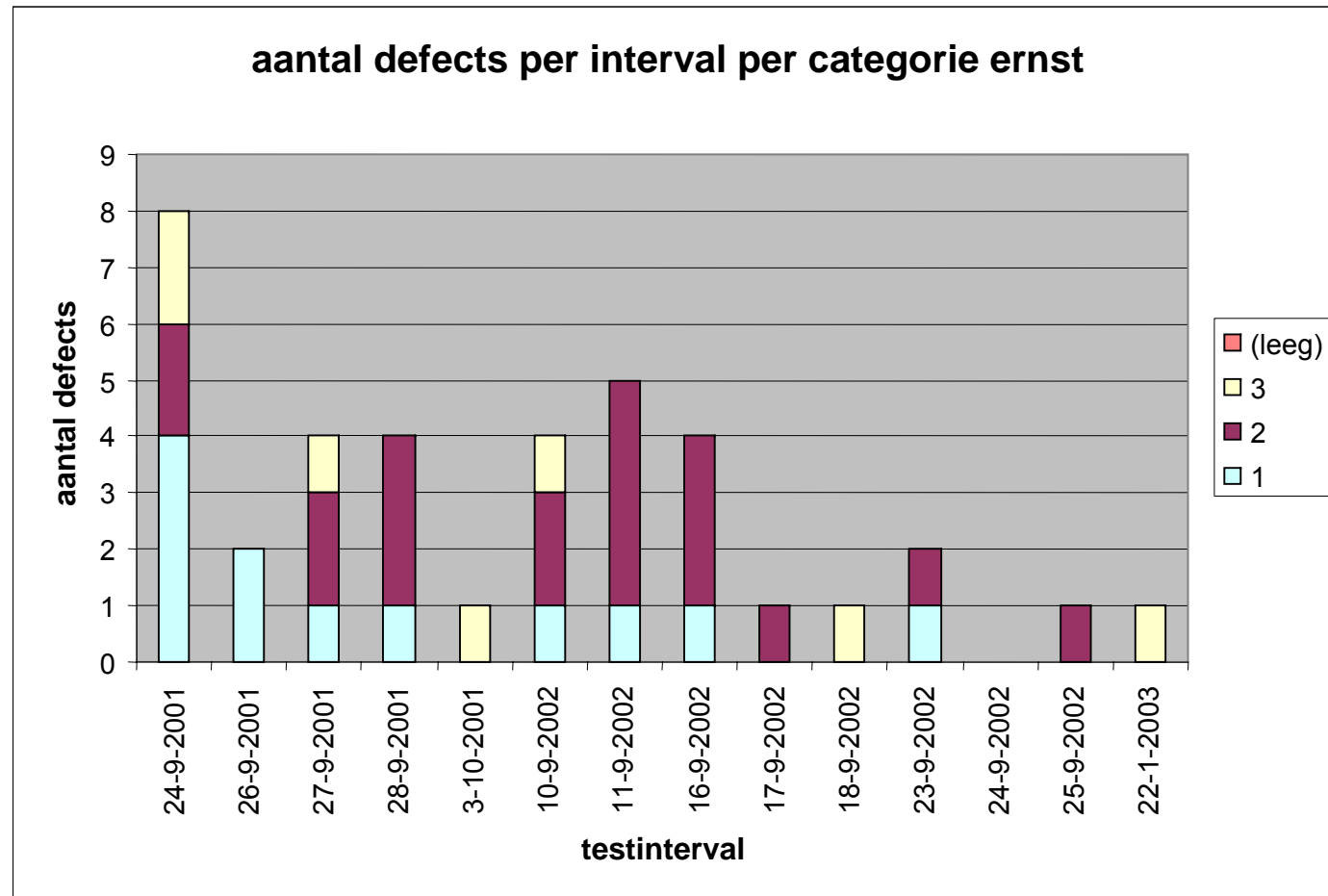
## Filteren defectgegevens

- ◆ dubbele meldingen verwijderen
- ◆ niet relevante defects verwijderen

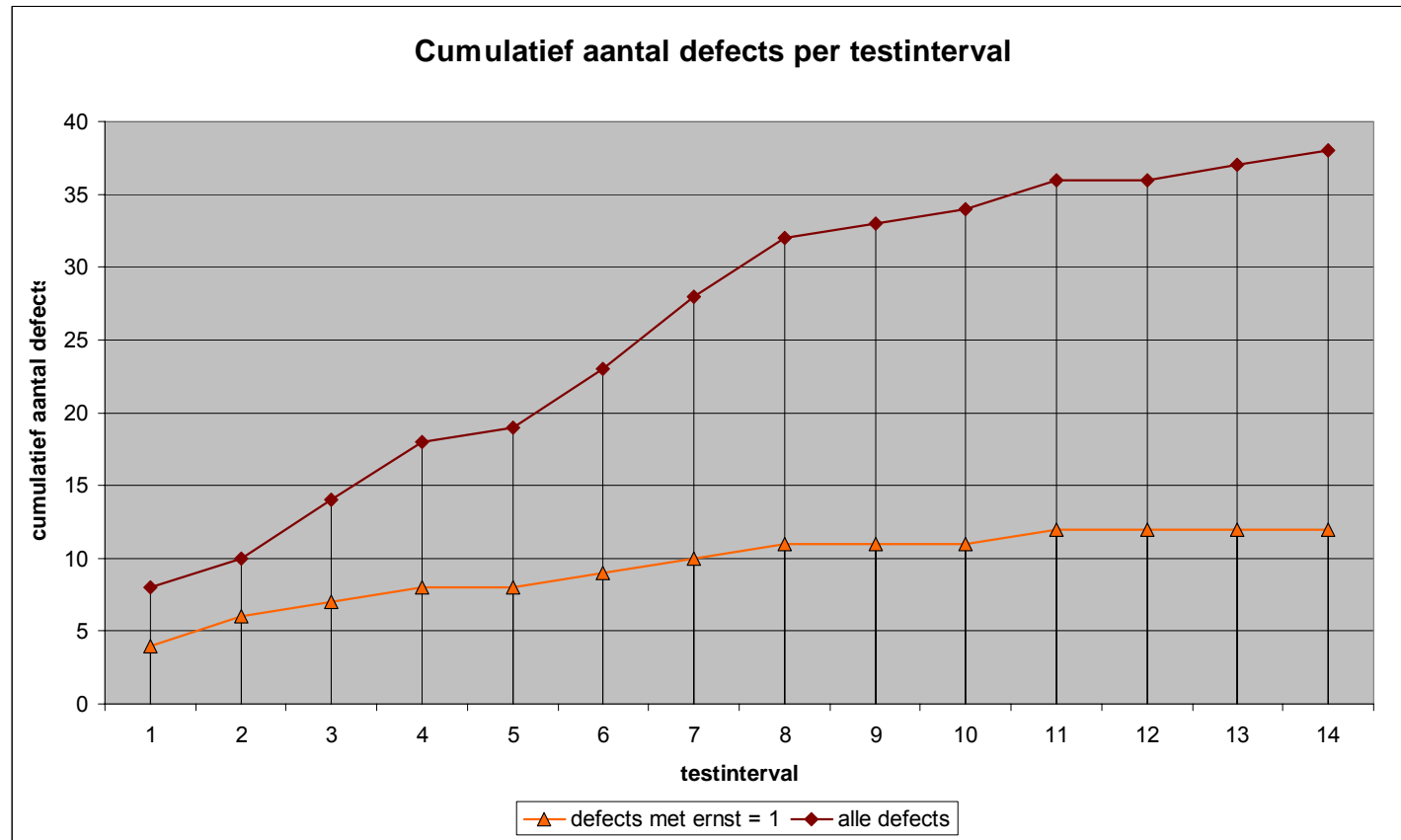
## Soorten defects naar “herkomst”



## Invoer voor betrouwbaarheidsmodellen



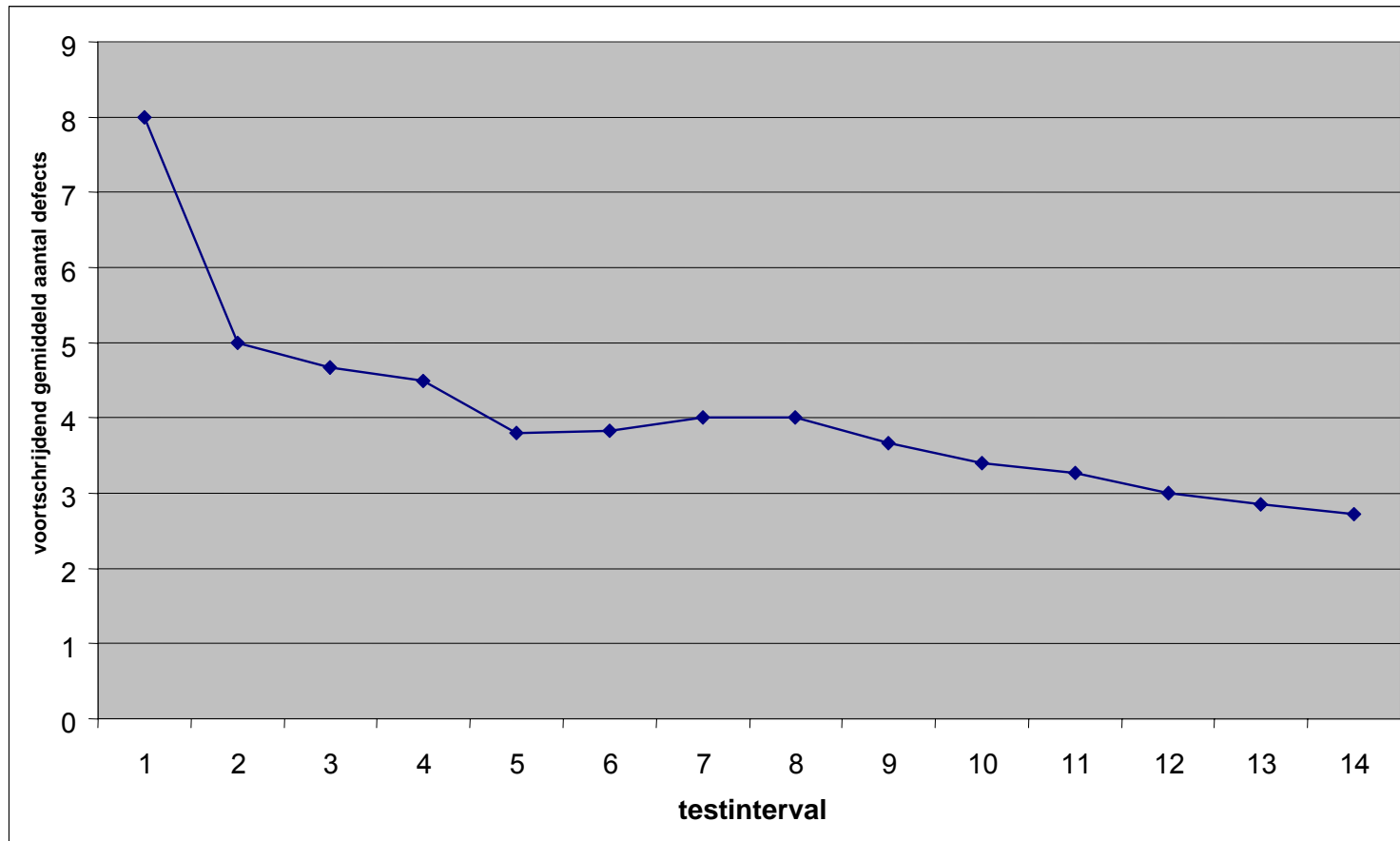
## “Ruwe data”



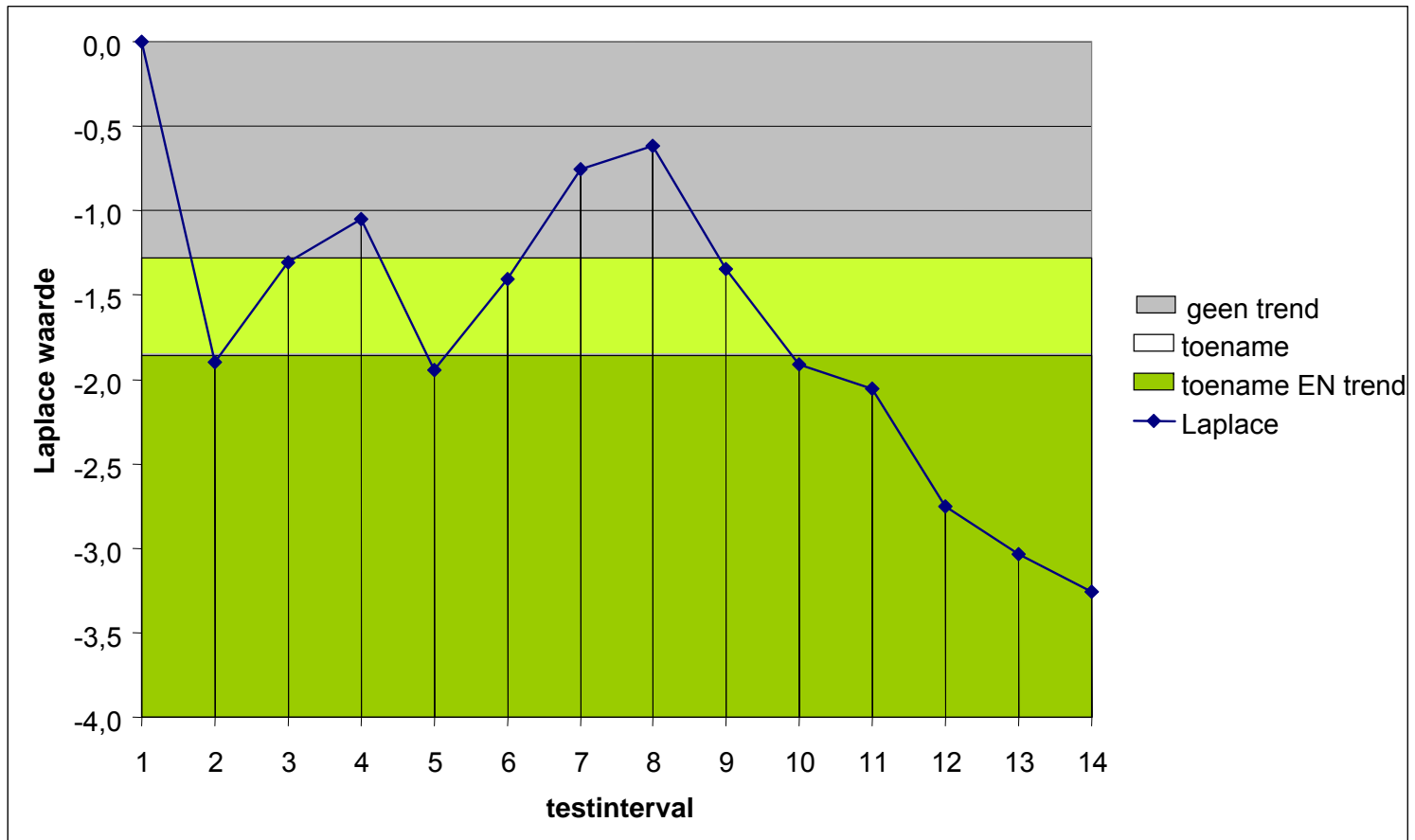
## Vaststellen betrouwbaarheidsgroei

- ◆ Voortschrijdend gemiddelde
- ◆ Laplace-test

## Voortschrijdend gemiddelde



# Laplace test



# Modelselectie: weging en selectie

assumption	atomic model relative importance	atomic model									
		Geometric	Jeinski-Moranda	Littlewood-Verrall (linear and quadratic)	Musa basic	Musa-Okamoto	Nonhomogeneous Poisson (TBF)	Generalized Poisson (ici. Schick-Wolverton)	Nonhomogeneous Poisson (FC)	Schneidewind (all 3 variants)	Yamada S-shaped
1 time-between-failure	2										
2 failure-count per testinterval	2							x	x	x	x
3 all testing intervals are of the same length	3									x	
4 software operated in similar manner as anticipated in operational use	3	x	x	x	x	x	x	x	x	x	x
5 all failures do not have same chance of detection	2										
6 all failures are equally likely to occur	2		x						x	x	
7 detections of faults are independent of each other	2	x			x	x			x	x	x
8 each failure is of the same severity as any other failure	1										
9 failure detection rate forms a geometric progression and is constant between failure occurrences	3	o									
10 failure detection is proportional to current fault content	2		o					o		o	
11 failure rate remains constant over the interval between failure occurrences	3		o								
12 the expected number of failures is a logarithmic (or proportional) function of time	3					o		o			
13 the failure intensity decreases exponentially with the expected number of failures found	3					o					
14 the cumulative number of failures detected at any time follows a Poisson distribution	3						o		o		
15 the hazard rate is proportional to the number of failures remaining in the program	1				o						
16 successive TBF's are independent random variables with exponential distribution	2			o	o						
17 program may get less reliable if more failures are inserted than are removed during correction	3										
18 failures are corrected instantaneously (at end of interval) without introducing new failures	3		x		x		x	x	x	x	x
19 the failure correction rate is proportional to the failure occurrence rate	1				x						
20 the total number of failures expected to be seen has an upper bound	2		o		o					o	o
21 there is no upper bound to the total number of failures	2	o		o		o					
<b>total score</b>		<b>36%</b>	<b>44%</b>	<b>25%</b>	<b>56%</b>	<b>33%</b>	<b>46%</b>	<b>57%</b>	<b>80%</b>	<b>79%</b>	<b>83%</b>
<b>total score incl. optional</b>		<b>71%</b>	<b>83%</b>	<b>58%</b>	<b>88%</b>	<b>87%</b>	<b>69%</b>	<b>93%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>

## Modelselectie: resultaat

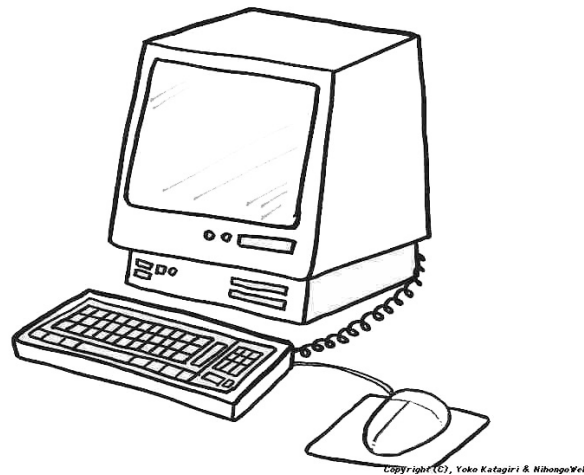
- ◆ Yamada
- ◆ Schneidewind
- ◆ NHPP
- ◆ Gegeneraliseerd Poisson

## Instellen modelparameters

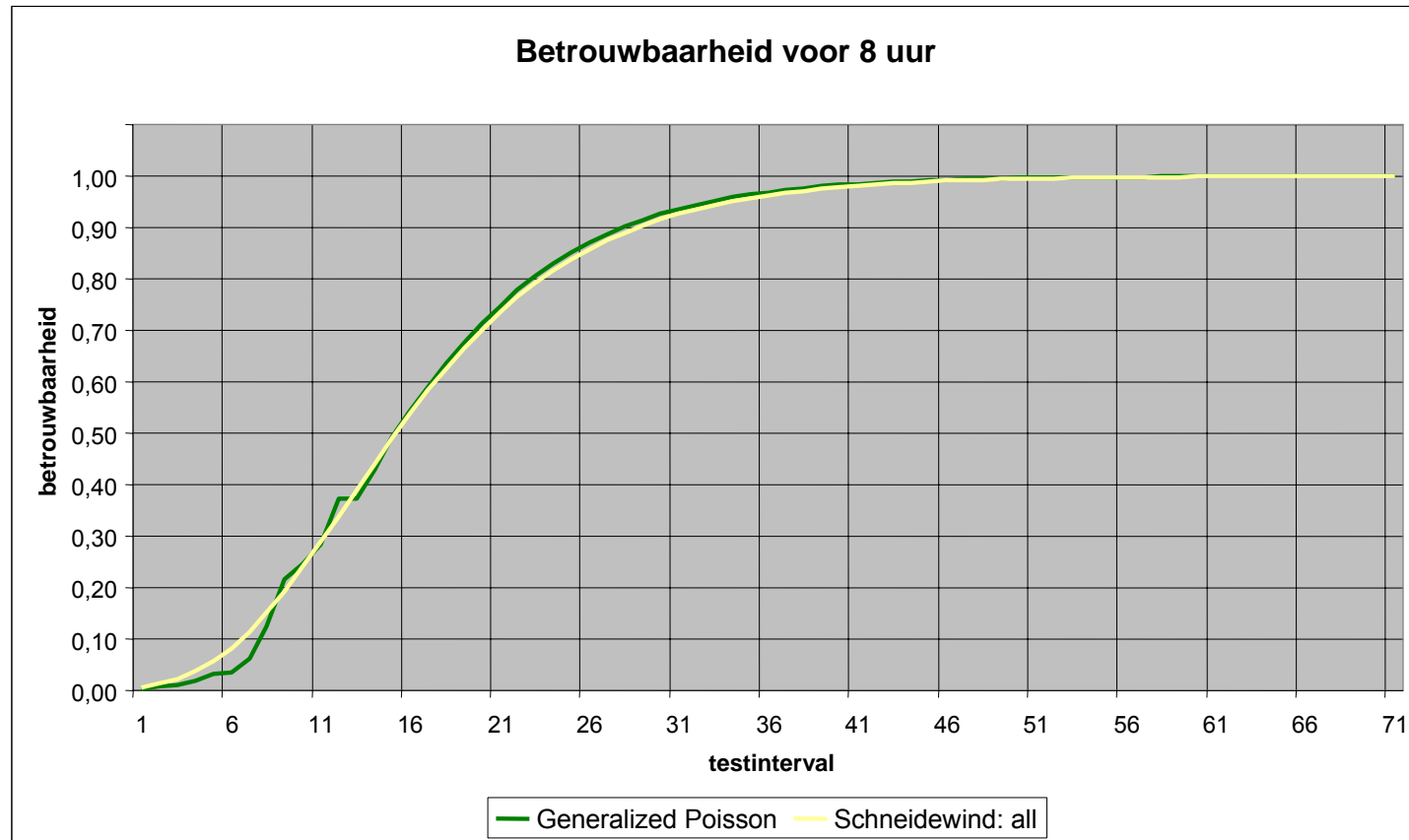
- ◆ welke defectgegevens ?
  - ◆ alle
- ◆ significantie t.b.v. Goodness Of Fit
  - ◆ 5%
- ◆ methode van parameterschatting
  - ◆ Maximum likelihood
- ◆ Voorspellingsperiode
  - ◆ 100

# Berekening

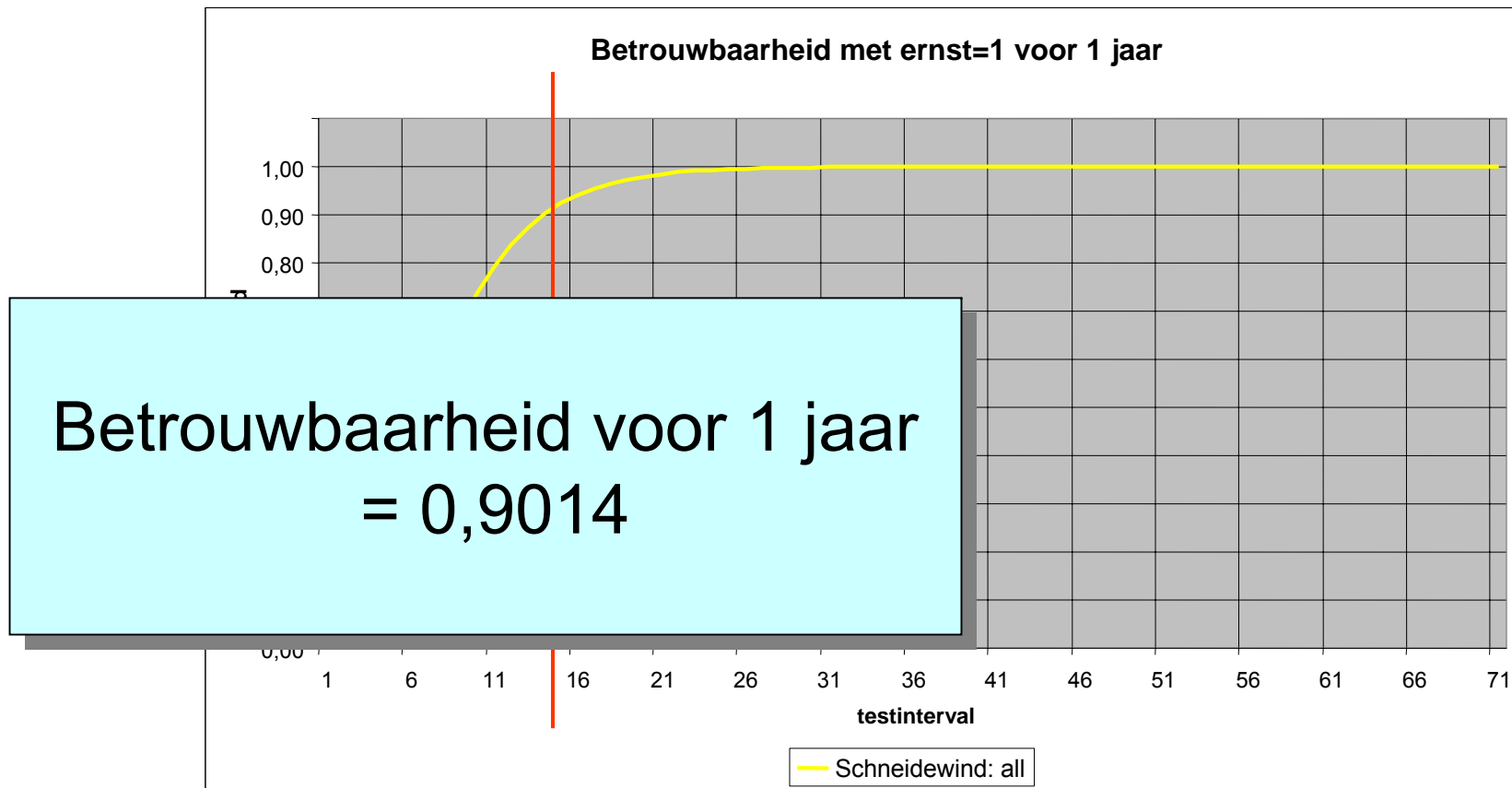
Modelberekening (laten) uitvoeren



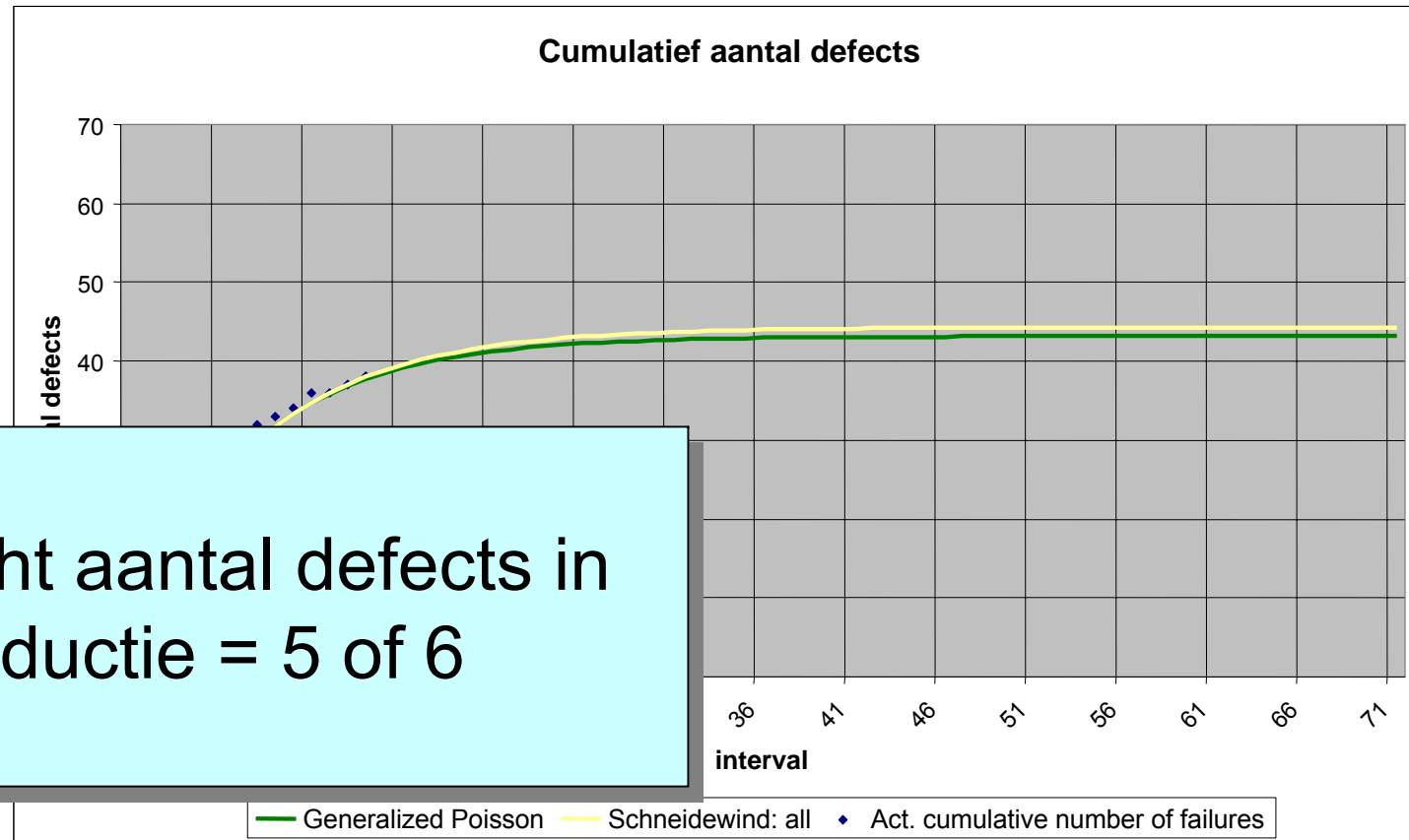
## Uitkomsten: betrouwbaarheid voor 8 uur



## Uitkomsten: betrouwbaarheid voor 1 jaar



## Uitkomsten: cumulatief aantal defects



## Conclusie

- ◆ Betrouwbaarheid voor het eerste jaar = 0,9014

## Ofwel

- ◆ De faalkans voor het eerste jaar = 0,0986  
(1 - 0,9014)

## Met

- ◆ Totaal aantal te verwachten defects van max. 6

## Betrouwbaarheid uitkomsten

- ◆ gemeten waarden versus modelwaarden:  
toeval of niet?
- ◆ Vaststellen m.b.v.  $\chi^2$  - test:  
 $\chi^2 \approx 7,1$   
P-waarde  $\approx 20 \%$

## Inhoud

1. Principes van softwarebetrouwbaarheid
2. Case Rijkswaterstaat
3. Toegevoegde waarde
4. Samenvatting en conclusie

## Toegevoegde waarde

Met behulp van testresultaten inzicht verkrijgen in het systeemgedrag in termen van:

- ◆ aantal nog aanwezige defects
- ◆ moment van optreden
- ◆ betrouwbaarheid / faalkans
- ◆ faalintensiteit

## Toepassingen

- ◆ Aantonen dat gewenste betrouwbaarheid is bereikt:  
kunnen we stoppen met testen?
- ◆ Aantonen dat een afgesproken service niveau wordt gehaald
- ◆ Inschatten van benodigde resources voor correctief onderhoud

## Inhoud

1. Principes van softwarebetrouwbaarheid
2. Case Rijkswaterstaat
3. Toegevoegde waarde
4. Samenvatting en conclusie

## Conclusie

Het gebruik van statistische modellen maakt een adequate voorspelling mogelijk over het gedrag van een informatiesysteem in de toekomst.

## Voetangels en klemmen

- ◆ Onbekendheid met eigenschappen van systeem en proces
- ◆ Onvolledige documentatie
- ◆ Hoe betrouwbaar is de betrouwbaarheidsanalyse?

The logo for Refis features the word "Refis" in a large, black, sans-serif font. The letter "e" is stylized with a diagonal slash. The text is set against a light blue background that has a curved, wave-like shape on the left side, creating a sense of motion or a stylized 'R'.

# Refis

system reliability engineering

Merellaan 5  
3722 AK Bilthoven  
tel: +31(0)30 225 36 37  
fax: +31(0)30 225 36 49  
[www.refis.nl](http://www.refis.nl)  
[info@refis.nl](mailto:info@refis.nl)