

Modelonzekerheid in GeoTOP

TNO – Geologische Dienst Nederland

Versiehistorie

Documentversie	GeoTOP versie	Toelichting
24 juni 2014	GeoTOP v1.2	De in dit document beschreven modelonzekerheid is opgenomen in modelgebied Westelijke Wadden dat met GeoTOP v1.2 is geïntroduceerd.

Definities

Het voxelmodel GeoTOP kent twee maatstaven van modelonzekerheid. Dit zijn:

- Modelonzekerheid van **geologische eenheid** – de mate waarin het model in staat is om een eenduidige schatting te geven van de geologische eenheid waartoe de voxel behoort.
- Modelonzekerheid van **lithoklasse** – de mate waarin het model in staat is om een eenduidige schatting te geven van de voor de voxel representatieve lithoklasse.

In beide gevallen wordt de modelonzekerheid uitgedrukt als een getal van 0 tot 1 waarbij 0 staat voor een zeer lage modelonzekerheid (het model is goed in staat een eenduidige schatting te geven) en 1 voor een zeer hoge (het model is niet in staat om een eenduidige schatting te geven).

In GeoTOP v1.2 zijn beide modelonzekerheden opgenomen als attribuut in het voxelmodel van modelgebied Westelijke Wadden. In een volgende versie zullen de modelonzekerheden ook in de overige modelgebieden worden opgenomen.

Waarom onzekerheid?

In GeoTOP is de ondergrond onderverdeeld in een regelmatig grid van aaneengesloten voxels (volumecellen) van 100 bij 100 m in de horizontale richtingen en 50 cm in de verticaal. Voor elke voxel wordt geschat tot welke geologische eenheid de voxel behoort en wat de voor de voxel representatieve lithoklasse is.

Het belangrijkste uitgangspunt voor GeoTOP zijn de boormonsterprofielen in de DINO-databank. Elk van deze boormonsterprofielen geeft gedetailleerde informatie over de opbouw van de ondergrond op één specifieke locatie. Aan het aardoppervlak is iets minder dan 10% van de voxels doorboord (we hebben de beschikking over circa 425.000 boormonsterprofielen op een oppervlak van 41.000 km²). Met de diepte neemt dit percentage snel af. Dit betekent dat we voor verreweg de meeste voxels in het model een schatting moeten doen op basis van de in de omgeving van de voxel aanwezige boormonsterprofielen. Hoe goed het model hiertoe in staat is, is onder andere afhankelijk van:

- de geologische complexiteit (de lithoklasse van een homogeen samengestelde eenheid is beter te schatten dan die van een heterogeen samengestelde eenheid);
- de hoeveelheid en de kwaliteit van de boormonsterprofielen in de omgeving van de voxel;
- de aan het model opgelegde randvoorwaarden zoals verbredingsgrenzen;
- het gebruikte algoritme met de bijbehorende parameters zoals de gehanteerde ruimtelijke correlatiefunctie.

In GeoTOP worden de schattingen met behulp van stochastische interpolatietechnieken gemaakt. Deze technieken komen er in essentie op neer dat het model 100 keer wordt doorgerekend met telkens een andere, maar statistisch gezien even waarschijnlijke, uitkomst. Voor de lithoklasse van een voxel wordt dan bijvoorbeeld 80 keer klei geschat, 10 keer veen en 10 keer kleiig zand.

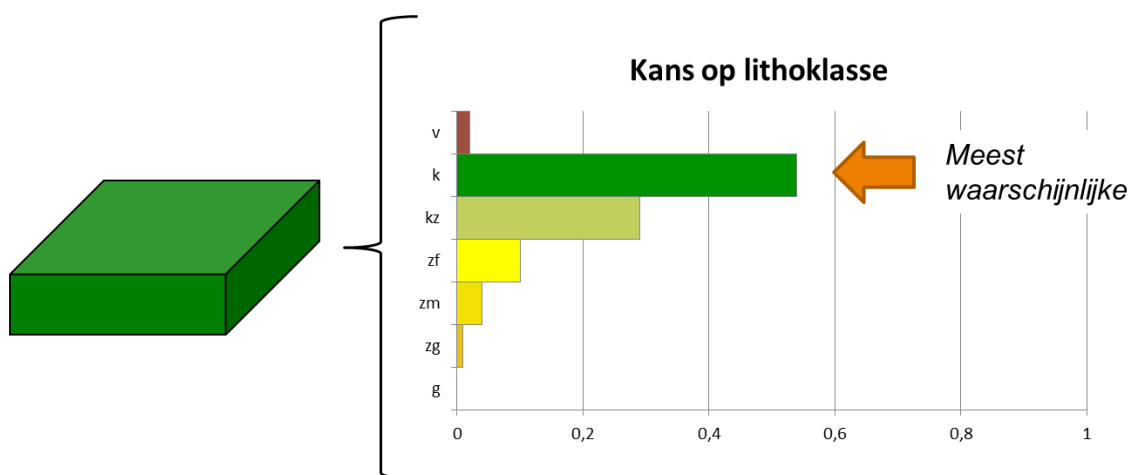
De verschillende uitkomsten van de modelberekeningen geven aan hoe goed het model in staat is om een eenduidige schatting te geven: in het beste geval leidt elke modelberekening tot dezelfde uitkomst, in het slechtste geval komen alle mogelijke uitkomsten even vaak voor.

Kansen

Een eerste stap om modelonzekerheid te bepalen is het berekenen van kansen. Uit de verschillende uitkomsten van de modelberekeningen wordt voor elke lithoklasse de kans op voorkomen berekend door het aantal keren dat de lithoklasse is geschat te delen door het aantal modelberekeningen (100). In het eerder beschreven voorbeeld is de kans op klei dan 0.8, de kans op veen 0.1 en de kans op kleiig zand eveneens 0.1.

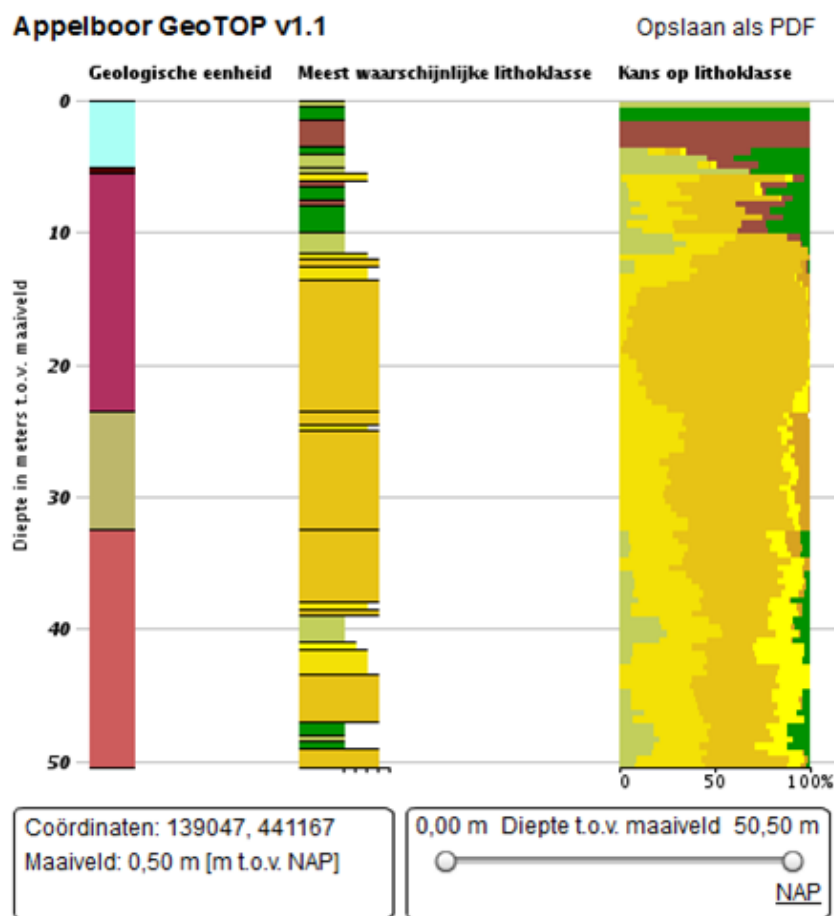
Op een vergelijkbare manier kan ook de kans worden berekend dat een voxel tot een bepaalde geologische eenheid behoort. De voxel hoort bijvoorbeeld in 75 modelberekeningen tot het Hollandveen Laagpakket en in 25 gevallen tot het Laagpakket van Walcheren. De bijbehorende kansen zijn dan 0.75 en 0.25.

Voor individuele voxels kan de kansverdeling worden weergegeven in een histogram, waarmee een visualisatie van de modelonzekerheid in de betreffende voxel wordt verkregen (Figuur 1).



Figuur 1 – Visualisatie van modelonzekerheid van een individuele voxel door het weergegeven van de kans op lithoklasse in een histogram. In dit voorbeeld is de meest waarschijnlijke lithoklasse klei, met een kans van ruim 50%. Er is ook een vrij grote kans op kleiig zand (~30%), de kans dat de voxel zand of veen bevat is echter klein (~20%).

Ook in een verticale voxel-stack (“appelboor”) zijn de verschillende kansen gezamenlijk te visualiseren (Figuur 2).



Figuur 2 – Visualisatie van modelonzekerheid van een verticale voxel-stack door het weergeven van de kans op lithoklasse in een gestapeld histogram. In diepte-bereiken waar één lithoklasse overheerst (bijvoorbeeld tussen 14 en 20 m) is de modelonzekerheid klein, in diepte-bereiken waar meerdere lithoklassen met vergelijkbare kansen voorkomen (bijvoorbeeld tussen 5 en 10 m) is de modelonzekerheid groot.

In een horizontale doorsnede (rasterkaart) of een verticale doorsnede (profiel) is het echter niet langer mogelijk om de kansen gezamenlijk te visualiseren: je ziet altijd naar keuze één van de kansen.

Ook voor kans op geologische eenheid geldt dat het niet mogelijk is om op basis van de kansen op enkele tientallen geologische eenheden een beeld van modelonzekerheid te verkrijgen.

Modelonzekerheid

TNO – GDN heeft een maat van modelonzekerheid vastgesteld die in één getalswaarde wordt uitgedrukt in plaats van met een afzonderlijke kans voor elke mogelijke lithoklasse of geologische eenheid. Deze maat is afgeleid van het concept van *informatie-entropie* (Shannon, 1948), die voor het eerst door Wellmann & Regenauer-Lieb (2012) is toegepast op 3D geologische modellen. In plaats van de term informatie-entropie gebruiken we modelonzekerheid.

De modelonzekerheid heeft onder andere de volgende, voor ons doel goed bruikbare eigenschappen:

1. Modelonzekerheid is 0 als elke modelberekening tot dezelfde geschatte lithoklasse leidt, ofwel er is één lithoklasse met kans 1, en alle andere lithoklassen hebben kans 0.
2. Modelonzekerheid is 1 (maximale waarde) als alle mogelijke lithoklassen met dezelfde kans voorkomen. Het model kan dan geen eenduidige schatting geven van de lithoklasse van de voxel.
3. Hoe meer mogelijke lithoklassen met een kans groter dan 0, hoe groter de modelonzekerheid.
4. Hoe groter de verschillen tussen de kansen, hoe kleiner de modelonzekerheid.

Voorbeelduitwerking

In onderstaande tabel is de modelonzekerheid (H) uitgewerkt voor een model met drie mogelijke lithoklassen (bijvoorbeeld zand, klei, veen, met kansen p_1, p_2, p_3).

P_1	P_2	P_3	H
1	0	0	0
1/3	1/3	1/3	1
0.5	0.5	0	0.63
0.49	0.49	0.02	0.71

In de eerste situatie is de kans op de eerste lithoklasse 1, en hebben de beide andere lithoklassen een kans 0. Hieruit volgt dat het model zeer goed in staat is om een schatting te geven en de modelonzekerheid is daarom 0.

In de tweede situatie zijn de kansen op de drie lithoklassen aan elkaar gelijk. Het model is niet in staat om een eenduidige schatting te geven en de modelonzekerheid is daarom 1.

In de derde situatie zijn er twee lithoklassen met gelijke kansen. Het model kan geen eenduidige schatting geven van de eerste twee lithoklassen, maar lithoklasse 3 komt zeker niet voor.

In de laatste situatie wordt een kleine kans op lithoklasse 3 ($p_3 = 0.02$ of 2%) geïntroduceerd waardoor de modelonzekerheid relatief sterk toeneemt.

Berekeningswijze

De modelonzekerheid H van een individuele voxel kan eenvoudig worden berekend uit de som van de producten van elke kans op lithoklasse en de logaritme van de kans op lithoklasse:

$$H = - \sum_i^N p_i \log p_i$$

Hierin staat H voor de modelonzekerheid, N voor het aantal mogelijke uitkomsten (dat wil zeggen het aantal mogelijke lithoklassen) en p_i voor de kans op een bepaalde lithoklasse (bijvoorbeeld kans op veen).

Voor het bijzondere geval dat er een $p_i = 1$ bestaat, is de modelonzekerheid 0. Voor de lithoklasse met $p_i = 1$ geldt dan namelijk $\log 1 = 0$, en voor de overige lithoklassen geldt $\lim_{x \rightarrow 0} (x \log x) = 0$. De totale som is dan eveneens 0.

In het geval dat alle lithoklassen dezelfde kans hebben ($p_1 = p_2 = p_3 = \dots = p_N$) is de modelonzekerheid maximaal. De maximale waarde is gelijk aan 1 als het grondtal van de logaritme overeenkomt met het aantal mogelijke uitkomsten N .

In applicaties waarin alleen met grondtal 10 of e ($^{10}\log$ of \ln) gerekend kan worden kan de formule herschreven worden als:

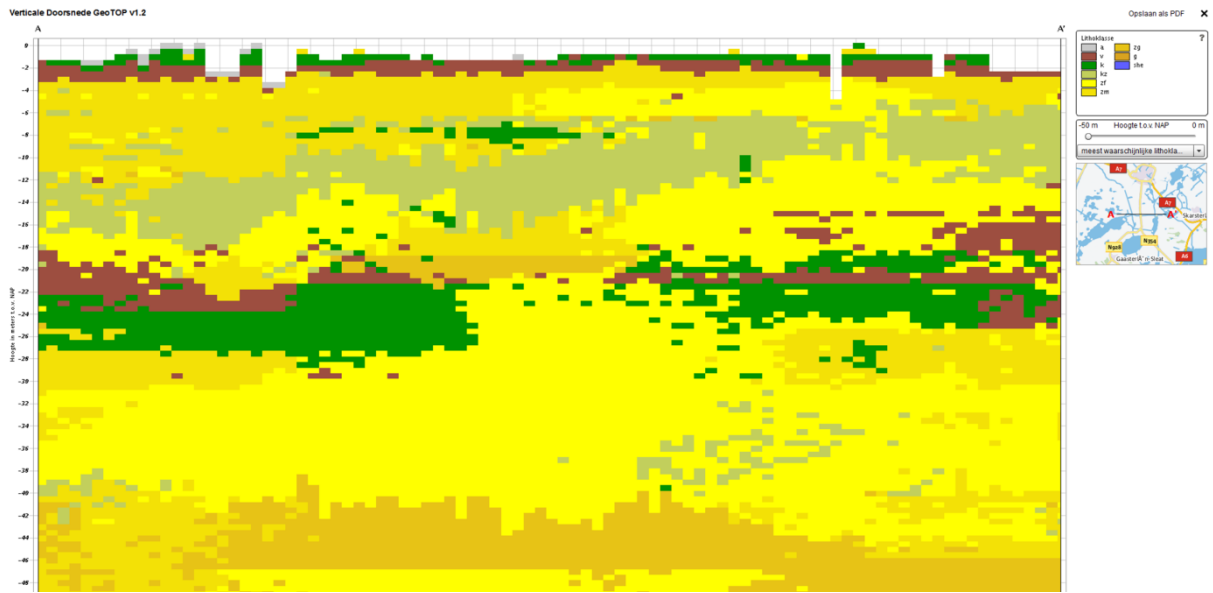
$$H = - \sum_i^N p_i \frac{\ln p_i}{\ln N}$$

Voorbeelden modelonzekerheid van lithoklasse

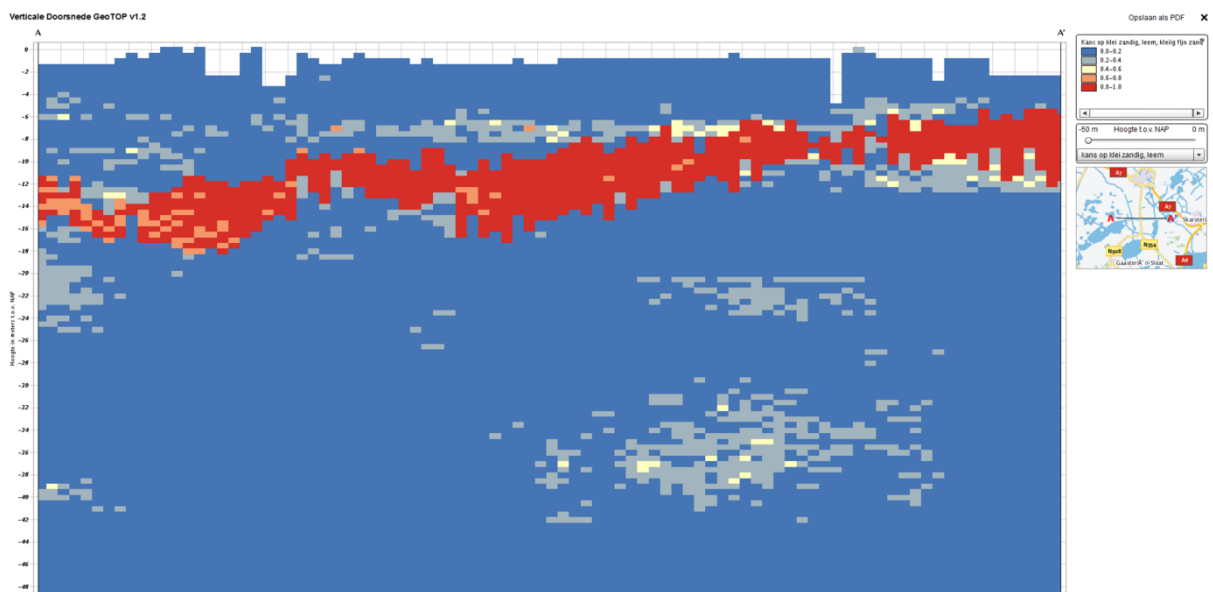
Figuur 3 toont een doorsnede door het lithoklassemodel van GeoTOP met de *meest waarschijnlijke lithoklasse* als attribuut. Dit is het attribuut waar de modelonzekerheid betrekking op heeft. Figuur 4 en 5 geven de kans op lithoklasse weer voor resp. *kleig zand en zandige klei* en *matig grof zand*. De afzonderlijke kansen-plaatjes geven een goed inzicht in de waarschijnlijkheid dat een bepaalde lithoklasse optreedt maar tonen niet de modelonzekerheid van lithoklasse als geheel.

Figuur 6 geeft de *modelonzekerheid* van lithoklasse weer. Deze is volgens bovenstaande formule berekend op basis van zes of zeven *kans op lithoklassen*: zes voor de eenheden waarin grind meegenomen is in grof zand, en zeven voor de eenheden waarin grind als aparte lithoklasse is gemodelleerd.

Modelonzekerheid in GeoTOP

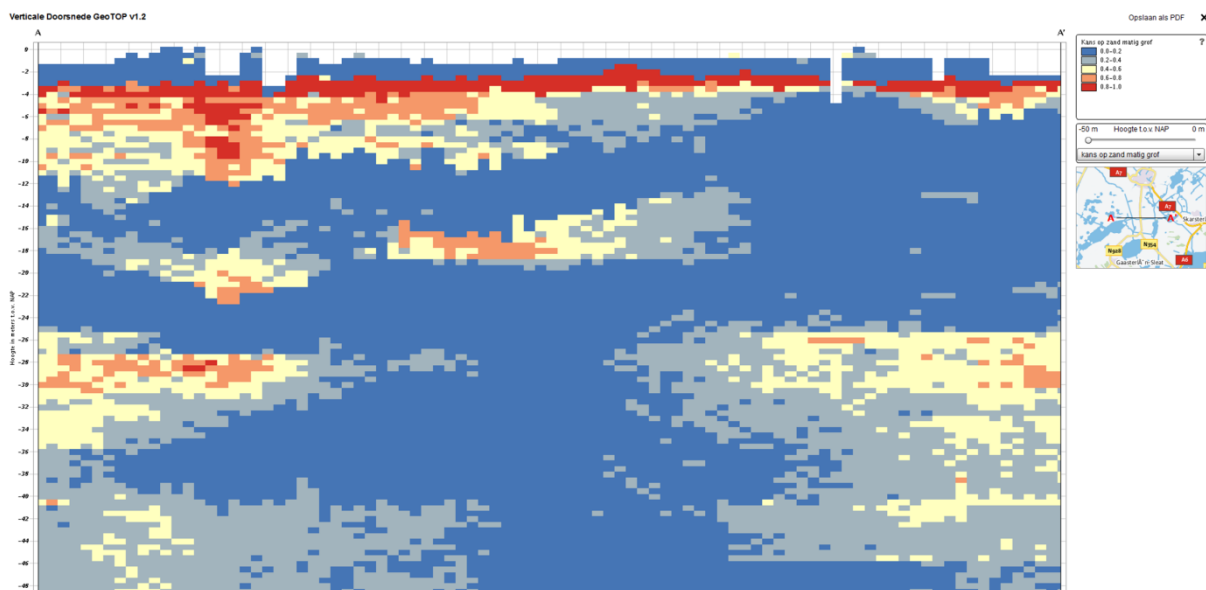


Figuur 3 – Doorsnede door GeoTOP met de meest waarschijnlijke lithoklasse.

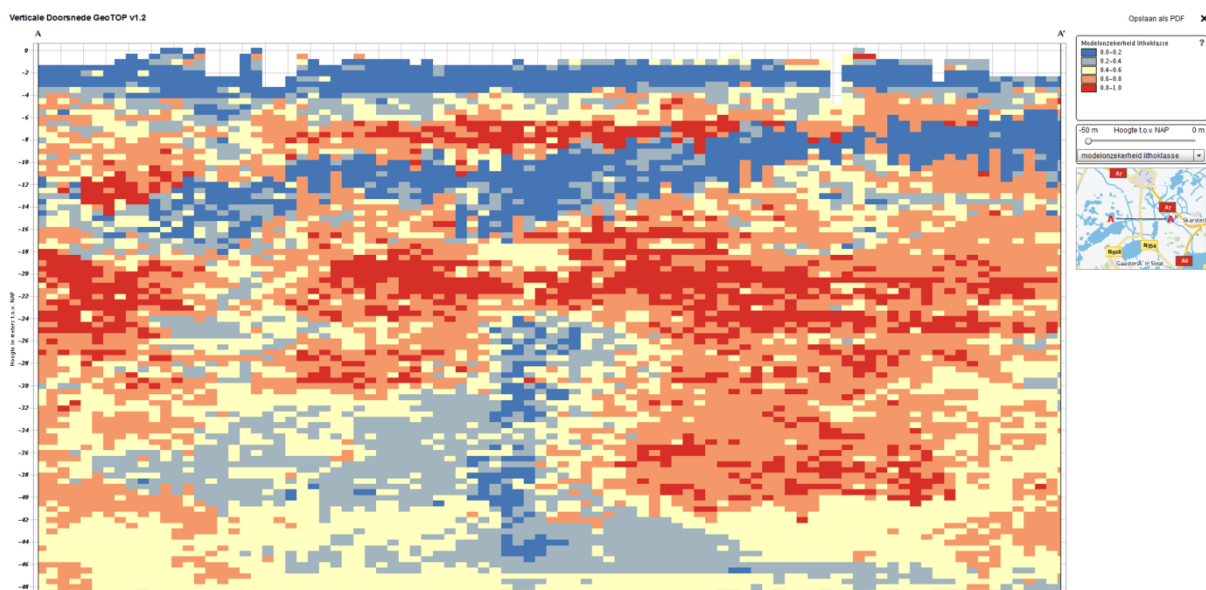


Figuur 4 – Kans op kleiig zand en zandige klei. Vooral in één geologische eenheid (Formatie van Drente, Laagpakket van Gieten) zijn hoge kansen waarneembaar.

Modelonzekerheid in GeoTOP



Figuur 5 – Kans op matig grof zand, met hoge kansen in de eenheid Formatie van Boxtel, Laagpakket van Wierden en daarnaast relatief hoge kansen verspreid in de diepere Pleistocene ondergrond.

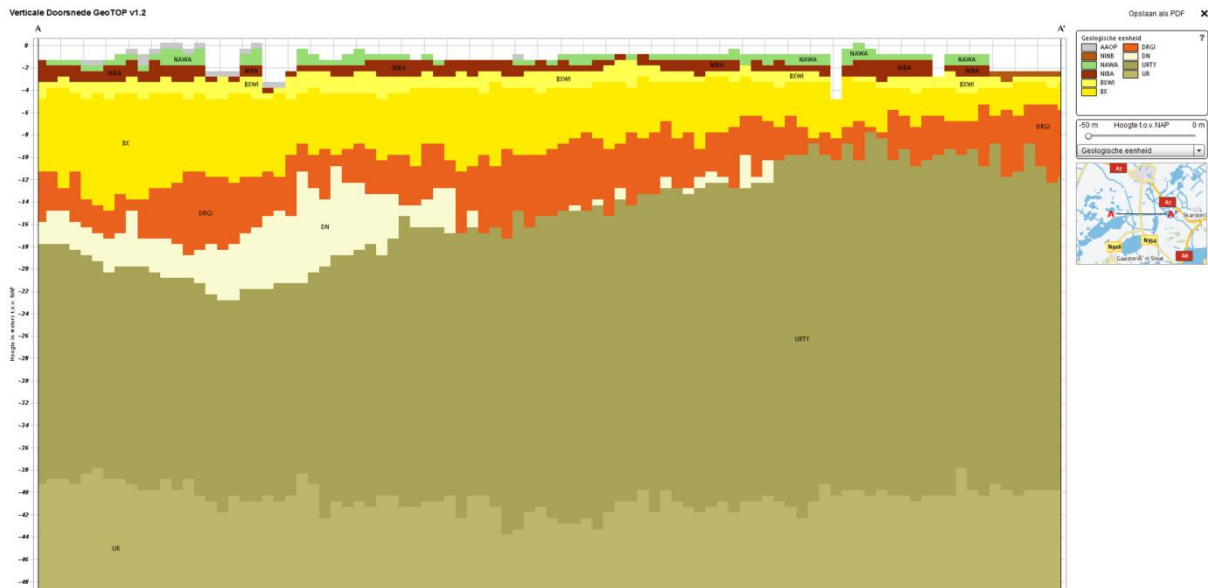


Figuur 6 – Modelonzekerheid van lithoklassen, met o.a. een lage onzekerheid in de eenheden waar een bepaalde lithoklasse overheerst (vergelijk met Figuur 4 en 5).

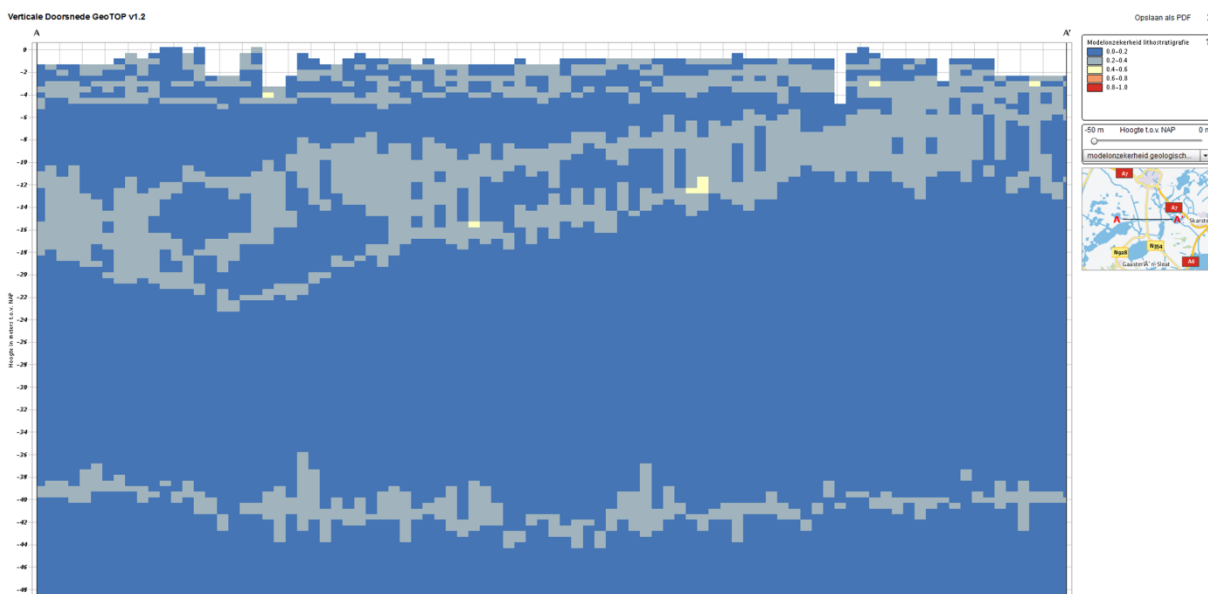
Voorbeelden modelonzekerheid van geologische eenheid

Figuur 7 toont het voxelmodel met de *geologische eenheid* als attribuut. Dit is het attribuut waar de modelonzekerheid betrekking op heeft. Figuur 8 geeft de *modelonzekerheid van geologische eenheid* weer. Deze is volgens de eerder gegeven formule berekend op basis van 21 afzonderlijk kansen. In het figuur zijn duidelijk zones van hoge modelonzekerheid te zien rondom de grenzen tussen de geologische eenheden.

Modelonzekerheid in GeoTOP



Figuur 7 – Doorsnede door GeoTOP met de geologische eenheid.



Figuur 8 – Modelonzekerheid van geologische eenheid.

Beperkingen

- De modelonzekerheid is gebaseerd op de in GeoTOP gebruikte stochastische interpolatietechnieken. Deze technieken houden niet expliciet rekening met onzekerheidsmarges in de gebruikte brongegevens (voornamelijk boormonsterprofielen).
- Buiten de verbreiding van een geologische eenheid is de kans op die eenheid altijd 0. Zie bijvoorbeeld de rechterkant van Figuur 8 waar buiten de verbreiding van de Formatie van Naaldwijk de kans op deze eenheid altijd 0 is. Dit betekent dat de kansen geen onzekerheidsinformatie over de laterale verbreiding van de geologische eenheid bevatten.

Referenties

Shannon, E.C., 1948. A mathematical theory of communication. Bell System Technical Journal 27, p. 379–423.

Stafleu, J., Maljers, D., Busschers, F.S., Gunnink, J.L., Schokker, J., Dambrink, R.M., Hummelman, H.J., Schijf, M.L., 2012. GeoTOP modellering (in Dutch). TNO Report 2012 R10991, 216 p.

Wellmann, J.F. & Regenauer-Lieb, K., 2012. Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics 526–529, p. 207–216.