

# TEMPORELE ASPECTEN IN GIS

Peter van Oosterom, Sectie GIS-technologie, Onderzoeksinstituut OTB, TU Delft

Temporele aspecten zijn niet nieuw en ook niet specifiek voor Geografische Informatie Systemen (GIS). Ook in andere typen informatiesystemen komen temporele aspecten (steeds vaker) voor. Daar waar men zich in het begin van de automatisering vooral richtte op het vastleggen van de huidige situatie (banksaldo, eigenaar, etc) is er daarna ook steeds meer aandacht gekomen voor het vastleggen van de complete historie.

Zo ook in GIS en Langran (1992) geeft hier een aantal voorbeelden ter motivatie voor:

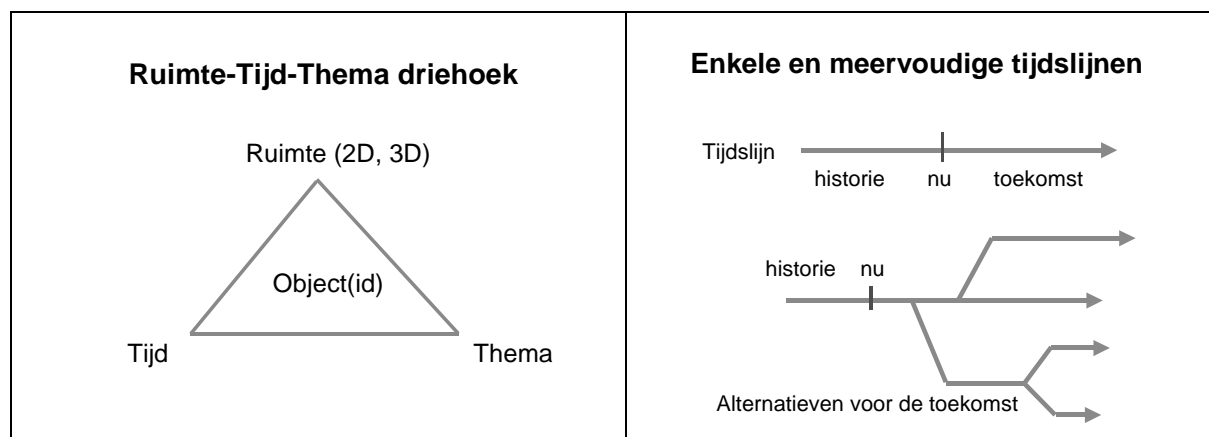
- Waar en wanneer deed zich een bepaalde verandering voor?
- Welke soorten veranderingen hebben plaatsgevonden?
- Wat is het tempo van de verandering (trend)?
- Wat is de periodiciteit van de herhalende verandering (indien aanwezig)?
- Waar was dit object twee jaar geleden?
- Hoe is dit gebied de afgelopen vijf jaar veranderd?
- Welk proces ligt er ten grondslag aan deze verandering?

Gezien het belang van dergelijke vragen (waarvan er vele ook buiten GIS van toepassing zijn) is er de afgelopen twee decennia veel onderzoek naar temporeel GIS gedaan en dit onderzoek is in vele gevallen nog steeds gaande. Zo gaf Langran (1992) het volgende overzicht van functies in een temporeel GIS: inventaris (complete beschrijvingen); analyse (verklaren, onderzoeken, voorspellen); bijhouding (oude versies door nieuwe versies laten opvolgen); kwaliteitscontrole (monitoren en evalueren van nieuwe gegevens en controle op consistentie met oude gegevens); inplannen (vaststellen drempelwaarden, welke bepaalde acties automatisch starten; b.v. aankoop); en visualiseren (kaartweergaven of tabellen van een temporeel proces).

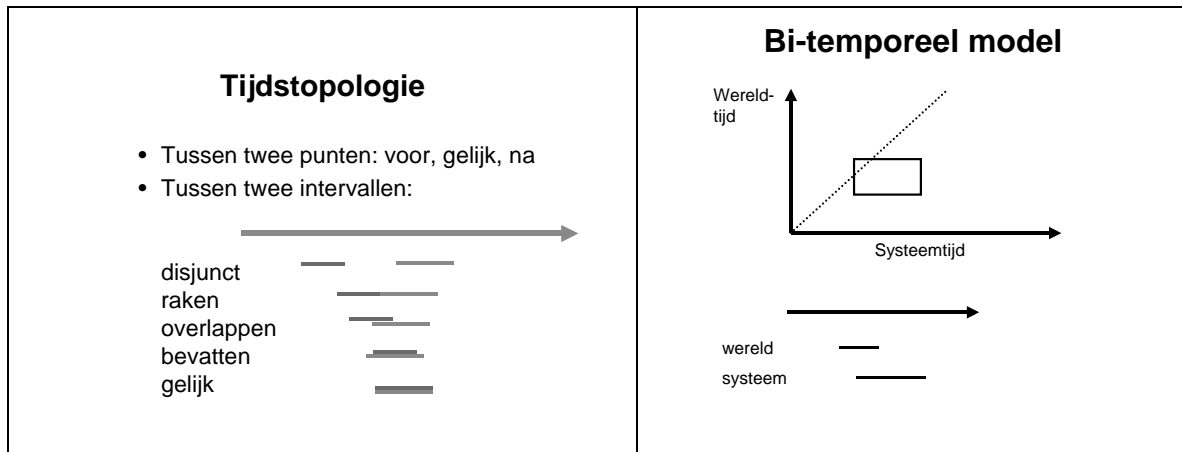
In paragraaf 1 worden enkele temporele basisprincipes besproken. De meer technische systeemaspecten komen aan bod in paragraaf 2. In paragraaf 3 zal een concreet ruimtelijk-temporeel systeem (Kadaster) getoond worden en het artikel sluit af met discussie (3D+tijd of 4D datatypen) en conclusie in paragraaf 4.

## 1. Temporele basisprincipes

Aan de hand van Figuur 1 worden hier enkele temporele principes binnen GIS toegelicht. Elk ruimtelijk-temporeel object heeft een identiteit (object id) en thematische attributen zoals weergegeven in de *ruimte-tijd-thema driehoek* (Figuur 1 links). Een belangrijk concept in temporeel modelleren is de *tijdslijn* (Figuur 1 rechts), waar het duidelijk wordt dat er een richting is waarin de tijd altijd doorloopt. Een *tijdspunt* (of tijdstip of moment) is gedefinieerd door een enkel punt op deze tijdslijn. Er is een heel speciaal punt en dat is 'nu', dat continu opschuift naar rechts op de tijdslijn. Een *tijdsinterval* is een periode van tijd gedefinieerd door een begin- en een eindtijdstip op de tijdslijn. De kleinste tijdseenheid wordt 'chronos' genoemd en is bijvoorbeeld vergelijkbaar met de resolutie of pixelgrootte in een raster GIS. Zich herhalende patronen of gebeurtenissen kunnen worden gemodelleerd, waarbij de *frequentie* aangeeft hoe vaak (snel) de herhaling plaatsvindt per tijdseenheid. Net als in het ruimtelijke domein kunnen er topologische relaties worden gedefinieerd, maar dan nu tussen de temporele primitieven.



Figuur 1: ruimte-tijd-thema driehoek (links) en tijdslijnen (rechts)



Figuur 2: tijdstopologie (links) en het bi-temporele model (rechts)

Figuur 2 links toont een aantal mogelijke topologische relaties tussen twee tijdsintervallen. Normaal gesproken is er één tijdslijn in het verleden, maar kunnen er meerdere tijdslijnen in de toekomst zijn (minder eenvoudige topologie); zie Figuur 1 rechts (onder). In principe zouden er ook meerdere tijdslijnen in het verleden kunnen zijn zoals bijvoorbeeld in het geval dat deze deels onbekend is zoals in de geologie waar meerdere scenario's zijn gemaakt die de ontwikkelingen kunnen verklaren. In dergelijke gevallen kan het verleden ook door meerdere tijdslijnen worden gerepresenteerd.

Er kunnen meerdere soorten tijd worden beschouwd bij het modelleren in een GIS en het is goed deze duidelijk van elkaar te onderscheiden. Hieronder een aantal voorbeelden van verschillende relevante tijden die vastgelegd zouden kunnen worden:

- wanneer iets in de echte wereld plaatsvond;
- wanneer iets is waargenomen (datum van foto);
- wanneer het in de database is opgenomen;
- wanneer het voor het laatst is gecontroleerd;
- wanneer het tijdstip van handtekening / poststempel / registratie;
- wanneer het voor het laatst getoond is aan gebruiker (op scherm of kaart), etc.

Zo onderscheidt het bekende *bi-temporele model* voor elk object een wereltijdinterval (bestaan in werkelijkheid) en een systeemtijdinterval (bestaan in database). Het resultaat is een 2D rechthoek, die rechts onder de diagonaal (de gestippelde lijn in Figuur 2 rechts) ligt. Het is logisch dat de systeemtijd achterloopt bij de wereld tijd (zowel begin- als eindtijdstip) in geval van een systeem dat de werkelijkheid registreert. Hierdoor liggen

linkeronder- en rechterbovenhoek van de rechthoek onder de diagonaal (en daardoor ook het grootste deel van de rechthoek).

## 2. Systeem aspecten

Na de meer conceptuele basisprincipes rondom tijd, komt er bij het bouwen van een systeem nog een aantal praktische en technologische aspecten aan bod. Een van de eerste vragen is hoe fijn of hoe grof de gegevensgranulariteit in verhouding tot de tijd is. Hieronder een viertal *granulariteiten* van grof naar steeds fijner, waaruit in de praktijk een keuze zal moeten worden gemaakt:

- Per (gehele) *gegevensverzameling*: bijv. traditionele topografische kaarten en luchtfoto's die elke 6 jaar ververs worden;
- Per *objectklasse*: bijv. de toekomstige topografische kaarten waarin wegen elke twee jaar ververs zullen worden;
- Per *objectinstantie*: bijv. bij de huidige kadastrale kaart waarbij elk object individueel kan worden ververs; of
- Per *attribuut*: bijv. wanneer delen van een objectinstantie gelijk blijven, maar een attribuut kan apart worden ververs (actuele grondwaterstand op bepaalde meetlocatie).

Er is geen 'optimale' oplossing voor alle situaties, dit hangt af van het soort gegevens (discrete objecten door de mens gemaakt of meer continue natuurlijke fenomenen) en de behoefte aan verschillende tijdsfuncties. Naast de keuze voor een bepaalde granulariteit, moet er ook worden nagedacht over de twee verschillende aanpakken voor het modelleren van dynamische systemen Dit zijn:

- Gebeurtenisgericht modelleren: de transacties of veranderingen worden als aparte entiteit binnen het systeem gemodelleerd, dus met een eigen

identiteit en eigen attributen. Wanneer de begintoestand en alle gebeurtenissen bekend zijn, dan is het mogelijk om elke toestand van het systeem te reconstrueren door het doorlopen van de ketting van gebeurtenissen. Het is ook mogelijk om de huidige toestand te representeren en niet de begintoestand en dan terug te gaan in de tijd door het 'ongedaan' maken van de gebeurtenissen;

- Toestandsgericht modelleren: alleen de toestanden, dus de resultaten van de gebeurtenissen, worden expliciet gemodelleerd. Elk object krijgt tenminste twee tijdsattributen die aangeven op welk interval de objectversie geldig is. Door twee opeenvolgende versies te vergelijken is het mogelijk te reconstrueren wat er bij een specifiek transactie is gebeurd. Het is eenvoudig om de toestand op elk gewenst moment in de tijd op te vragen door de juiste objectversies te selecteren op basis van hun tijdsinterval.

In de praktijk worden gebeurtenis- en toestandsgericht modelleren ook vaak gecombineerd. In het Core Cadastral Domain Model (Lemmen en Van Oosterom, 2006) worden zowel de gebeurtenissen (binnenkomst brondocumenten zoals akten en metingen) als de resulterende toestanden (aangepaste rechten en percelen) opgeslagen. Een oplossing die een redelijke balans biedt tussen mogelijkheden en complexiteit, is het implementeren van tijd op objectinstantie (record)niveau in een toestandsgerichte aanpak. Ieder object is hiervoor uitgebreid met twee extra attributen:  $t_{min}$  en  $t_{max}$ , die het geldige tijdsinterval van deze objectversie weergeven (Stonebraker en Rowe, 1986). Nieuwe objecten krijgen naast hun gewone attributen ook de huidige tijd ingevuld bij  $t_{mix}$ . Objecten, die momenteel nog steeds geldig zijn, hebben voor  $t_{max}$  de speciale waarde  $TMAX\_VALUE$ . Bij een verandering van een bestaand object wordt er een kopie gemaakt van het gehele record en bij de oude versie de  $t_{max}$  ingevuld met de huidige tijd. Bij de nieuwe versie wordt dezelfde tijdwaarde bij  $t_{min}$  ingevoerd (zodat versies aansluiten) en de nieuwe attribuutwaarden worden in deze versie ingevoerd. Bij het verwijderen van een object wordt alleen de  $t_{max}$  waarde ingevuld met de huidige tijd, ter indicatie dat dit object vanaf nu niet meer geldig is. Over het algemeen zijn de twee meestgestelde vragen aan een ruimtelijk-temporeel model: 1. geef de kaart op een bepaald moment ( $t$ ) in de tijd en 2. geef de veranderingen in de kaart in een bepaald

tijdsinterval ( $t_1-t_2$ ). Beide vragen kunnen met het beschreven model efficiënt beantwoord worden. Gezien het feit dat de temporele behoefte vrij generiek is, is deze ook in een uitbreiding van de standaard database vraagtaal opgenomen, TSQL en inmiddels beschikbaar in verschillende DBMS implementaties (als eerste in de Postgres research database).

Het gevolg van het bewaren van historie is dat het systeem alleen maar groeit. Er zijn wel opties om historische data op te ruimen en/of elders te archiveren.

Speciale aandacht in het ruimtelijk-temporeel modelleren vragen de objectidentificaties (oid). Wordt er normaal van uitgegaan dat deze uniek zijn, in een temporeel systeem kunnen meerdere versie van hetzelfde object (records met dezelfde oid) bestaan. Om elke versie uniek te kunnen aanduiden moet de oid worden uitgebreid met een tijdstempel (of versienummer): het paar  $oid+t_{min}$  (dit is beter dan  $oid+t_{max}$ , omdat  $t_{max}$  één keer kan veranderen). Bij onderlinge object-verwijzingen moet alleen het oid gedeelte expliciet worden gebruikt (tijd is impliciet), dit in verband met het zogenaamde 'olievlekprobleem' indien tijd ook expliciet zou worden opgeslagen als verwijzing. Hierdoor zou niet alleen het object waar naar wordt verwezen veranderen, maar ook het object dat hier naar wijst zou moeten veranderen. Dit alleen maar omdat er een nieuwe versie komt van het gerelateerde object waar naar verwezen wordt en omdat het object dat verwijst dan zelf ook verandert (namelijk de tijdcomponent van de verwijzing), zouden ook de objecten die hier weer naar verwijzen moeten meeveranderen. Enfin, zie hier het olievlekprobleem dat zou ontstaan door een enkele verandering en indien verwijzingen zouden bestaan uit zowel de oid als een expliciete tijdcomponent. Om dit te voorkomen moet dus alleen de oid opgeslagen worden voor de verwijzing.

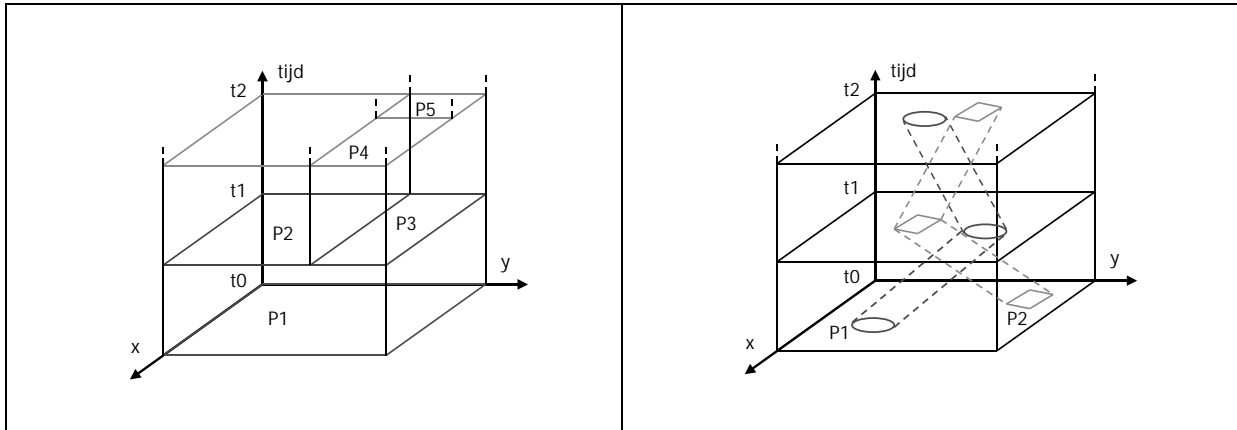
Bij meer ingewikkelde veranderingen zoals een splitsing van een object zouden de ouder-kind relaties tussen de object(versies) opgeslagen kunnen worden. Echter een redelijk deel hiervan is impliciet in de gelijke oid aanwezig (alleen  $t_{min}$  is per versie verschillend) en de overige gevallen kunnen met ruimtelijk-temporeel overlap worden gevonden.

Tot slot verdient het visualiseren van ruimte en tijd in GIS pakketten of bijv. in webomgevingen nog specifieke aandacht.

Hiervoor bestaan verschillende mogelijkheden:

- toon momentopname op een specifiek tijdstip  $t_1$ ;

- toon de veranderingen in een bepaald tijdsinterval t1-t2;
- toon de situatie op twee momenten t1 en t2 naast elkaar;
- toon in een animatie de veranderingen van t1 naar t2 (in relatieve of absolute tijdstappen);
- toon via cartografische symbolen de temporele gebeurtenissen (kleur voor tempo van verandering, symbolen markeren locatie van verandering, etc);
- toon de tijd als 3<sup>de</sup> dimensie (zie Figuur 3 links en rechts).

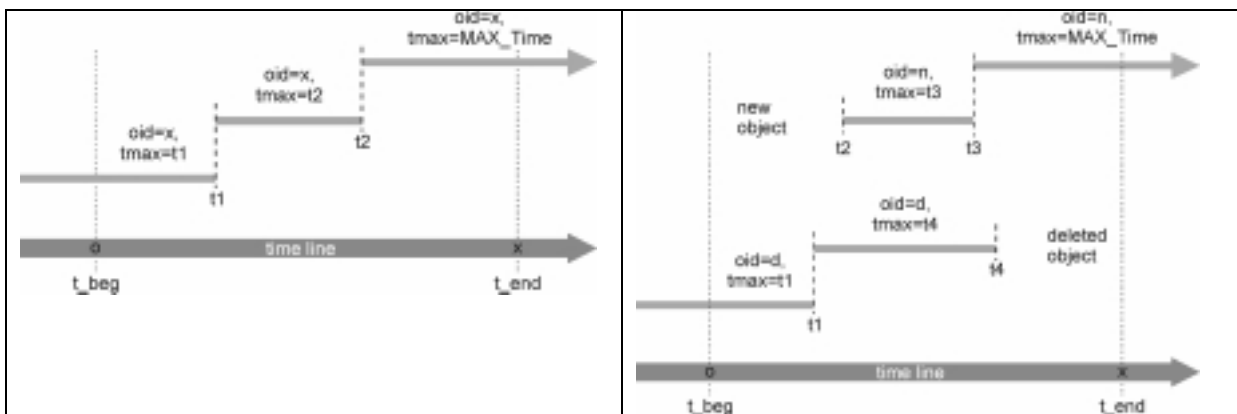


Figuur 3: Visualisatie van tijd als 3<sup>de</sup> dimensie: percelen splitsen (links) verplaatsen van twee object (rechts).

### 3. Kadastertoepassingen

Om de temporele basisprincipes en systeem-aspecten in actie te zien, wordt nu een specifieke toepassing wat verder besproken. Sinds 1997 wordt in Nederland de historie in de kadastrale kaart opgeslagen (Van Oosterom, 1997), in het Landmeetkundig en Kartografisch Informatiesysteem (LKI). Een belangrijke reden

van de invoering van de historie was de productie van mutatiebestanden voor de afnemers van de kadastrale kaart. Het werd zonder historie in LKI steeds moeilijker om deze zogenaamde was-woordt mutatiebestanden te produceren (door geometrische vergelijking van huidige situatie met opgeslagen kopieën van eerder geleverde momenten).



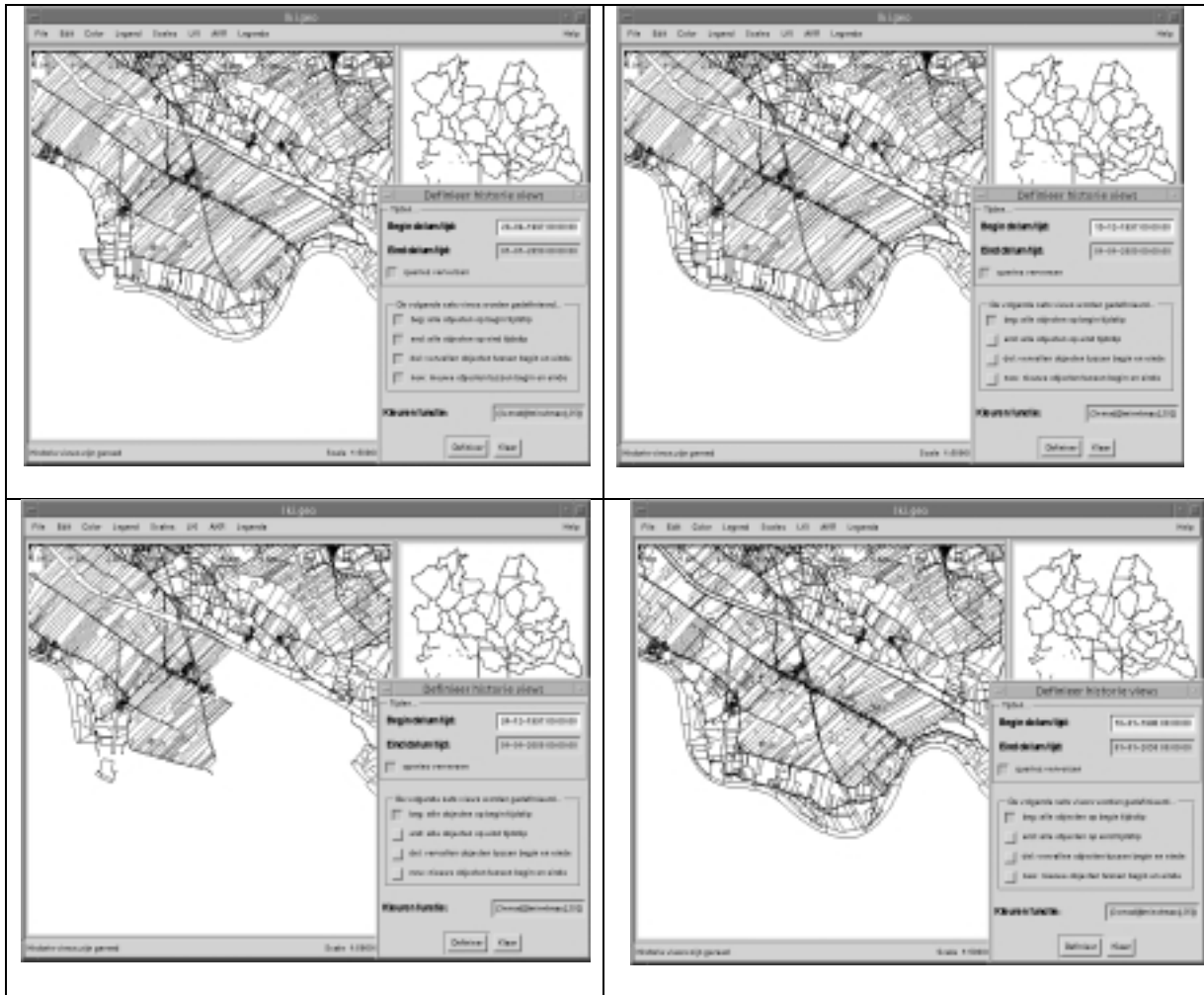
Figuur 4: De tijdslijn en een aantal versies van een veranderend object x (links), van een nieuw object n (rechts boven) en van een verwijderd object d (rechts onder)

Door introductie van de historie in LKI werden dit eenvoudige SQL queries. Voor een mutatiebestand zijn twee opties denkbaar: 1. alleen de verschillen tussen het begin- en eindtijdstip worden geleverd, of 2. alle mutaties worden geleverd (ook de tijdelijke). Ter illustratie

voor optie 2 resulteert dit voor respectievelijk de 'was' en de 'wordt' bestanden in:

```
select * from parcel p where
  t_beg < p.tmax and p.tmax <= t_end;
/* was */
```

```
en: select * from parcel p where
  t_beg < p.tmin and p.tmin <= t_end;
/* wordt */.
```



Figuur 5: Veranderingen door de tijd, 20 september '97 (linksboven), 15 december '97 (rechtsboven), 24 december '97 (linksonder), 10 januari '98 (rechtsonder)



Figuur 6: Veranderingen door de tijd: tijdstip van verwijdering van een perceel weergegeven via kleurcoderingen (links) of verwijderde gebouwen in een bepaalde periode (rechts)

Een tweede toepassing van de historie is het tellen van de mutaties in de Grootschalige Basiskaart Nederland (GKBN), die net als de kadastrale kaart ook in LKI is opgeslagen (Van Oosterom, Maessen en Quak, 2002). In elke regio bestaat een lokale instantie in de vorm van een aantal samenwerkende partners die gezamenlijk verantwoordelijk zijn voor de bijhouding van de GBKN. Per regio kan verschillen welke partner de feitelijke inwinning van de gegevens uitvoert (en hier ook voor gecompenseerd wil worden). Door de mutaties in LKI te tellen is sprake van een eerlijke basis voor deze verrekening, waarbij nog verschillende mutatiesoorten worden onderscheiden: verwijdering van een element, semantische verandering, het toevoegen van nieuwe elementen (in verschillende categorieën, harde/zachte topografie, wel/niet geconcentreerd).

Naast de beschrijving van deze twee toepassingen is het ook aardig om te kijken naar verschillende visualisaties van de historie in LKI. Figuur 5 toont een viertal stappen in het uitwerken van een landinrichtingsproject. Figuur 6 (links) toont de eindtijd van percelen via een kleurcodering en Figuur 6 (rechts) toont welke gebouwen in een bepaalde periode verwijderd zijn.

#### 4. Discussie en conclusie

In dit artikel zijn oplossingen beschreven voor het opslaan van de ruimtelijke dimensie en de temporele dimensie van een object als aparte attributen van een object. Indien de ruimtelijke dimensie 3D is, wordt ook wel van 3D+tijd of 4D systeem gesproken. Tijdens de AGGN-bijeenkomst ontstond een discussie over of een dergelijk 4D systeem ook baat zou hebben bij het gebruik van het 4D ruimtelijk-temporeel datatype (in plaats van een los temporeel attribuut en een los 3D ruimtelijk attribuut). Tijdens de AGGN-middag bleef deze vraag nog open, hieronder een poging alsnog een antwoord te geven.

De voordelen van een geïntegreerd 4D datatype zijn:

1. Optimale efficiënt 4D zoeken (opgeven van zowel ruimte- als tijdvoorwaarden in de zoekopdracht). Dit kan alleen indien er gebruik wordt gemaakt van een 4D datatype (en indexering/clustering). Anders moet de DBMS ('query plan') eerst op ruimte en vervolgens op tijd selecteren (of omgekeerd), wat minder efficiënt is. Merk op dat waarschijnlijk zelfs een 2D index/cluster aanpak al

- efficiënt genoeg zal zijn, omdat de 3<sup>de</sup> dimensie en de temporele dimensie niet zo selectief zijn. Dus dit is niet een erg sterk argument voor een 4D type;
2. Ouder-kind relaties tussen objecten uit een ruimtelijke partitie (zoals bij percelen) worden nu topologische buurvragen (voor 4D 'buren' waarbij de tijd anders is); zie Figuur 3 (links): Perceel P3 heeft ouderperceel P1 en kindpercelen P4 en P5. Topologische buurvragen zijn efficiënter dan het via ruimtelijke temporele overlap bepalen van ouder-kindrelaties (zie paragraaf 3);
3. 4D analyse: hebben twee bewegende objecten op enig moment in de tijd een overlap (zie Figuur 3 rechts)? Indien opgeslagen als 4D datatype dan is dit een eenvoudige query, er van uitgaande dat het 4D datatype de overlap operatie ondersteunt. Indien als aparte attributen opgeslagen dan is het antwoord minder eenvoudig te geven;
4. Wellicht de meest belangrijke vraag is of we een 4D partitie (van 3D ruimte en tijd) als basis voor ons model willen: dus geen 4D overlap en geen 4D gaten. In dat geval is het hebben van 4D geometrie en topologie de meest solide basis.

Er zijn echter ook argumenten, die in het voordeel van aparte opslag van de ruimte- en tijdattributen spreken:

1. Aparte ruimte- en tijdattributen zijn voldoende om alle mogelijke ruimtelijk-temporele situaties vast te leggen;
2. Huidige technologie kan gebruikt worden voor de implementatie (en dit is niet het geval bij 4D datatypen, waar eerst meer R&D moet worden uitgevoerd);
3. Temporeel modelleren vereist de nodige flexibiliteit: het kan meer zijn dan alleen het simpelweg toevoegen van een enkele tijdsdimensie. Er kunnen meerdere soorten tijd van belang zijn (en hierbij is het mogelijk dat een 4D datatype alleen niet voldoende is voor ruimte en alle 'smaken' tijd).

Gezien deze argumenten lijkt het erop dat op de middellange tot lange termijn het 4D datatype de gewenste solide basis kan vormen, vooral als een 4D partitie het uitgangspunt is. Op de kortere termijn lijkt het verstandig om voorlopig met aparte ruimte- en tijdattributen te werken en de 4D ontwikkelingen af te wachten.

Bij het modelleren van het aspect tijd in GIS moeten er vele vragen beantwoord worden: welke tijden zijn van belang (wereld of systeem), welke granulariteit moet er worden vastgelegd (dataset, objectklasse, objectinstantie, attribuut), betreft het discrete of continue veranderingen, ligt de nadruk op vastleggen toestanden of transacties (veranderingen), welke temporele functies moeten worden ondersteund (geef

toestand op tijdstip  $t_1$ , geef de veranderingen in periode  $t_1$ - $t_2$ ), etc. Er is geen oplossing die in alle situaties optimaal zal zijn, maar één ding is duidelijk: de temporele aspecten worden steeds belangrijker binnen GIS. De resultaten hiervan zijn zichtbaar in moderne standaarden (zoals GML3 van OGC/ISO) en datasets (zoals de kadastrale kaart en sinds kort ook de Top10NL).

### Dankwoord

Dank aan mijn (oud) collega's Chrit Lemmen, Bart Maessen en Wilko Quak voor de samenwerking op temporeel gebied, waaruit veel van de ideeën zijn voortgekomen die in dit artikel zijn beschreven.

Verder dank aan de AGGN (en met name Jeroen Baltussen, Frans Lips) en Frans van de Wel (KNMI) voor het organiseren van de AGGN themamiddag 'GIS en Tijd' op 6 juni 2006, waar de fundamentele vraag 'Heeft combinatie van ruimte en tijd in een 4D datatype zin?' aan de orde kwam.

### Literatuur

- Langran, G. (1992) *Time in Geographic Information Systems* (London: Taylor & Francis).
- Lemmen, C.H.J. en P.J.M. van Oosterom (2006) 'Version 1.0 of the Core Cadastral Domain Model', XXIII FIG Congress, 'Shaping the Change', Munich, Germany, October 8-13, 2006
- Oosterom, P.J.M. van (1997) Maintaining Consistent Topology including Historical Data in a Large Spatial Database. *In: Proceedings Auto-Carto 13*, Seattle WA, 8-10 April 1997, pp. 327-336.
- Oosterom, P.J.M. van, B. Maessen en W. Quak (2002) Generic query tool for spatio-temporal data, *In: International Journal of Geographical Information Science*, Volume 16, 8, pp. 713-748.
- M. Stonebraker, M. and L.A. Rowe (1986) The design of Postgres. *ACM SIGMOD*, 15(2), pp. 340-355.