

Modelit
Elisabethdreef 5
4101 KN Culemborg



info@modelit.nl
www.modelit.nl
+31(345)531717

Melissa: Schermopvolgschema en Datadictionary

Datum 16 juni 2008



Documentatiepagina

Opdrachtgever	RWS - DID		
Titel	Technische Documentatie Validatiesoftware Nat Deelrapport: Melissa: Schermopvolgschema en Datadictionary		
Versie	0.2	20 juni 2008	
Projectteam opdrachtgever	W.T.B. van der Lee, P. Haaring.		
Projectteam Modelit	N.J. van der Zijpp, K.J. Hoogland.		
Projectomschrijving	Vervaardigen aanvullende documentatie Validatiesoftware Nat		
Trefwoorden	Validatiesoftware, Donar, Datadictionary, Melissa		

Inhoud

1	Achtergrond van dit document.....	1
2	Schermpvolgschema.....	2
3	Datamodel	2
3.1	Hoofdstructuur.....	2
3.2	db.dia.....	3
3.3	db.model.....	9
3.4	db.vlg.....	12
3.5	db.cmb.....	13
3.6	db.donarlijst.....	13
3.7	db.cnscriteria.....	14
3.8	db.valcriteria.....	14
3.9	db.stuurfiles.....	15
3.10	db.transaction.....	16
4	Overzicht projecten en documentatie.....	18
4.1	Inleiding.....	18
4.2	Historisch overzicht uitgevoerde projecten met relatie naar Melissa.....	18
4.3	Beschikbaarheid documentatie.....	20

1 Achtergrond van dit document

Dit document is vervaardigd in het kader van het project "Technische Documentatie Validatiesoftware Nat". Naar aanleiding van onder andere een audit uitgevoerd in opdracht van Rijkswaterstaat-RIKZ wordt in dit project aanvullende documentatie vervaardigd voor validatiesoftware die door Modelit voor RIKZ is ontwikkeld.

Dit documentatieproject wordt uitgevoerd conform het projectplan neergelegd in de offerte van datum 12 juli 2007. De offerte maakt onderscheid tussen:

- Documentatie per Applicatie;
- Documentatie per toolbox;
- Documentatie per utility.

Dit document heeft betrekking op het onderdeel "Documentatie per Applicatie", deze bestaat uit:

- Het schermopvolgschema;
- Documentatie van de datadictionary;
- De algemene handleidingen (reeds aanwezig);
- Documentatie per source file (reeds aanwezig);
- Voor niet triviale numerieke modellen: een beschrijving van de werking van deze modellen

Alleen punt 1, 2 en 5 vragen in dit project aandacht. Het schermopvolgschema is opgenomen in hoofdstuk 2 en de documentatie van de datadictionary in hoofdstuk 3. In aanvulling op de vereisten van de offerte is in hoofdstuk 4 een historisch overzicht opgenomen van alle projecten en documenten die samenhangen met de Melissa applicatie. Uit dit overzicht blijkt ook waar de beschrijving van de niet triviale modellen te vinden is.

2 Schermopvolgschema

Het schermopvolgschema is geïntegreerd met de handleiding van Melissa en derhalve niet in dit document opgenomen.

3 Datamodel

De gegevens die tot een Melissa werkgebied behoren worden opgeslagen in een hiërarchische structuur. Deze structuur, ook wel datamodel genoemd, wordt in dit hoofdstuk beschreven. Omdat de structuur vrij uitgebreid is wordt deze niet in één keer gepresenteerd en beschreven maar worden bepaalde vertakkingen en subvertakkingen in aparte secties en subsecties beschreven. In deze gevallen is een verwijzing van de vorm "[+ Zie sectie x voor verdere detaillering]" opgenomen.

3.1 Hoofdstructuur

db	
+----dia	(Struct array) Opslag structuur voor tijdreeksen. Elk element bevat één tijdreeks. [+ Zie sectie 3.2 voor verdere detaillering.
+----model	(Struct array) Opslag structuur voor modellen voor het schatten van tijdreeksen. [+ Zie sectie 3.3 voor verdere detaillering.
+----vlg	(Struct array) Opslag structuur voor het bepalen van de groepering van de tijdreeksen. [+ Zie sectie 3.4 voor verdere detaillering.
+----cmb	(Struct array) Opslag structuur met informatie over welke tijdreeksen gelijktijdig getoond moeten worden. [+ Zie sectie 3.5 voor verdere detaillering.
+----donarlijst	(Struct array) Opslag structuur met informatie voor het rechtstreeks ophalen van tijdreeksen uit de Donar database. [+ Zie sectie 3.6 voor verdere detaillering.
+----cnsriteria	(Struct array) Opslag structuur voor de consistentie criteria. [+ Zie sectie 3.7 voor verdere detaillering.
+----valcriteria	(Struct array) Opslag structuur voor de validatie criteria. [+ Zie sectie 3.8 voor verdere detaillering.
+----stuurfiles	(Struct) Opslag structuur voor de inhoud en locatie van ingelezen stuurfiles.

	+ Zie sectie 3.9 voor verdere detaillering.
+----currentfilename	(String)
	Naam van het bestand met het huidige werkgebied.
+----transaction	(Struct array)
	Datastructuur voor het logboek.
	+ Zie sectie 3.10 voor verdere detaillering.

3.2 db.dia

In het veld db.dia worden tijdreeksen opgeslagen die uit .dia bestanden zijn ingelezen. Alle meta-informatie wordt daarbij opgeslagen zodat de reeksen later weer weggeschreven kunnen worden naar .dia bestanden. Ten behoeve van het maken van berekeningen met deze reeksen zijn een aantal velden toegevoegd met daarin de status, standaardafwijking, gemeten en voorspelde waarden en berekeningsresultaten van getijspecials functies.

db.dia

+----ID	(Integer)
	Uniek ID. Dit ID wijzigt na het aanmaken van een tijdreeks in het werkgebied niet meer, ook niet bij het invoegen of verwijderen van tijdreeksen. Het ID wordt gekozen als het tot op dat moment hoogste ID +1.
+----blok	(Struct array)
+----W3H	Deze velden bevatten structures met informatie uit de DIA blokken [W3H], [MUX], [TYP], [RGH], [RKS], [TPS] en [MUX] van de ingelezen tijdreeks. Deze informatie is weer nodig op het moment dat de tijdreeksen worden weggeschreven naar een .dia bestand, zie DONAR handleiding deel 7.
+----MUX	
+----TYP	
+----RGH	
+----RKS	
+----TPS	
+----WRD	
+----groep	(String)
	Groep waartoe deze reeks behoort, kan gebruikt worden om reeksen in bepaalde volgorde weer te geven, zie ook sectie 3.4.
+----combinatie	(Integer array)
	ID's van de tijdreeksen die gelijktijdig met deze tijdreeks getoond moeten worden, zie ook sectie 3.5.
+----taxis	(Double array)
	Deze variabele bevat tijdstippen, in Matlab datum formaat, waarnaar de equidistante tijdreeksen verwijzen.
+----W	(Double array)
	Gemeten waarden.
+----statusW	(UInt8 array)
	De status van de meetwaarden.
	Onderscheiden worden:

	0
	O
	k,
	25
	Int
	er
	po
	lat
	ie,
	98
	N
	on
	ob
	se
	rv
	ed
	,
	99
	Hi
	aa
	t.
+----V	(Double array)
	Voorspelde waarden.
+----statusV	(Uint8 array)
	Statuswaarde V
	Onderscheiden worden:

	0 O k, 24 O nb et ro u w ba ar , 99 Hi aa t.
+----StatusValidat 	(UINT8 array) Code waarmee gevalideerde van niet gevalideerde data kan worden onderscheiden, mogelijke waarden: 0 Ni et ge va lid ee rd , 1 G ev ali de er d.
+----StatusOutlier 	(UINT8 array) Outlierstatuswaarde van V. Onderscheiden worden:

	0
	Ok,
	99
	Hia
	at,
	110
	An
	der
	s,
	115
	Inc
	ons
	iste
	nt,
	120
	Out
	lier.
+-----stdv	(Double array)
	Veld waarin de standaardafwijking
	behorende bij het model waarmee V is
	berekend wordt opgeslagen.
+-----compute	(Logical)
	Variabele waarin wordt opgeslagen of voor
	deze tijdreeks V waarden moeten worden
	berekend. Onderscheiden worden:
	0
	Ni
	et
	be
	re
	ke
	ne
	n,
	1
	B
	er
	ek
	en
	.
+-----use	(Logical)

	Variabele waarin wordt opgeslagen of de gegevens van deze tijdreeks mogen worden gebruikt bij een MLRTI berekening. Onderscheiden worden: 0 Ni et to eg es ta an , 1 T oe ge st aa n.
+----valide	(Double) Code die de DONAR validatie status voor deze tijdreeks aangeeft. Onderscheiden worden:

	0
	O
	ng
	ec
	on
	tr
	ol
	ee
	rd
	,
	1
	G
	ec
	on
	tr
	ol
	ee
	rd
	,
	2
	D
	efi
	nit
	ief
	.

+----getij	
+----GDPS	(Double array)
	De getijparameterdataset voor deze reeks.
+----laaglaag	(Char array)
	Rapportage van de berekening van de laag-
	laagwaterspring voor deze reeks.
+----havengetal	(Char array)
	Rapportage van de berekening van het
	gemiddelde havengetal voor deze reeks.
+----vliegplan	(Char array)
	Rapportage van de berekening van de
	onderschrijding conform vliegplan. voor deze
	reeks.
+----tussenschreiding	(Char array)
	Rapportage van de berekening van de over-
	en onderschrijding voor deze reeks.
+----ongelijkheid	(Char array)
	Rapportage van de berekening van de
	gemiddelde dagelijkse ongelijkheid voor deze
	reeks.

3.3 db.model

Het veld db.model is een struct array. Elk element van deze array bevat de informatie van een model dat gebruikt kan worden voor het bijschatten van een tijdreeks.

db.model	
+----status	(Double)

	0
	N
	i
	e
	t
	g
	e
	k
	a
	l
	b
	r
	e
	e
	r
	d
	,
	1
	G
	e
	k
	a
	l
	b
	r
	e
	e
	r
	d
	,
	2
	G
	e
	k
	a
	l
	b
	r
	e
	e
	r
	d
	z
	o
	n
	d
	e
	r
	d
	a
	t
	a
	.

+----type	(String)
	Soort model, voorlopig kunnen alleen nog
	maar modellen van het type 'regressie' worden
	gebruikt.
+----naam	(String)
	Unieke naam van het model.
+----aantekeningen	(Char array)
	Ruimte voor het maken van aantekeningen en
	opmerkingen voor dit model. Deze tekst kan
	uit meerdere regels bestaan.
+----activeren	(Logical)
	Code die aangeeft of het model gebruikt mag
	worden (1) of niet (0) voor het bijstellen van
	tijdreeksen.
+----berekeningscode	(Double)
	code die momenteel nog niet in gebruik is,
	maar in de toekomst kan worden gebruikt om
	aan te geven onder welke omstandigheden het
	model van toepassing is.
+----model	
+----aantal_tijdstap	(Integer)
	Het aantal tijdstappen (van 'eenheid_tijdstap')
	tussen iedere periode. Dit aantal wordt bij het
	berekenen en gebruiken van de modellen
	gebruikt.
+----eenheid_tijdstap	(String)
	De eenheid van de gehanteerde tijdstap. Deze
	eenheid wordt bij het berekenen en gebruiken
	van regressie coëfficiënten gebruikt. Mogelijke
	waarden: 'min', 'uur' en 'culminatie'.
+----invoer	(Struct array)
	Definieert welke reeksen (locatie-parameter
	combinaties) met welke tijdsverschuivingen
	gebruikt worden in het regressiemodel. In het
	veld coef worden de geschatte coëfficiënten
	opgeslagen.
+----uitvoer	(Struct)
	Definieert welke reeks (locatie-parameter
	combinatie) voorspeld moet worden. De
	velden tShift en coef worden niet gebruikt.
+----intercept	(Double)
	Het geschatte intercept van het lineaire
	regressiemodel.
+----resultaat	
+----dNumber	(Integer)
	Het aantal datapunten op basis waarvan het
	model is gekalibreerd.
+----dbegin	(Double)

	Het aantal datapunten op basis waarvan het model is gekalibreerd.
+----dend 	(Double) Het aantal datapunten op basis waarvan het model is gekalibreerd.
+----jaar 	(Double) jaartal van genereren.
+----over1 	(Double) De fractie van het aantal residuen waarvoor geldt: overschrijding >1*standaardafwijking.
+----over2 	(Double) De fractie van het aantal residuen waarvoor geldt: overschrijding >2*standaardafwijking.
+----over3 	(Double) De fractie van het aantal residuen waarvoor geldt: overschrijding >3* standaardafwijking.
+----wminvmax 	(Double) Maximum van de residuele absolute afwijking.
+----wminvmin 	(Double) Minimum van de residuele absolute afwijking.
+----wminvstdafw 	(Double) Residuele standaardafwijking.
+----result 	(Double array) De door het model voorspelde waarden.
+----target	(Double array) De door het model te voorspellen waarden.

3.4 db.vlg

De tabel in het hoofdscherm met tijdreeksen kan gesorteerd worden. Met behulp van een extra kolom met een groepsnaam kan deze lijst in elke volgorde weergegeven worden die de gebruiker wenst. De groepsnaam kan rechtstreeks in de tabel of via een stuurfile met extensie .vlg ingesteld worden. De instellingen, een groepsnaam voor een parameter-locatie combinatie, worden opgeslagen als elementen van de struct array in het veld db.vlg.

db.vlg	
+----sLoccod 	(String) Locatie code.
+----sParcod 	(String) Parameter code.
+----groep	(String)

Groepsnaam.

3.5 db.cmb

Wanneer in het hoofdscherm een tijdreeks wordt geselecteerd verschijnen de gegevens van deze reeks in de onderste helft van het scherm. Er bestaat de mogelijkheid om in te stellen dat naast de gegevens van de geselecteerde tijdreeks, ook gegevens van andere tijdreeksen worden getoond. De instellingen voor het tonen van combinaties van tijdreeksen kunnen opgegeven worden via een stuurfile met extensie .cmb of rechtstreeks via de tabel met tijdreeksen. De instellingen, die gelijktijdig te tonen locatie-parameter combinaties voor een bepaalde parameter-locatie combinatie, worden opgeslagen als elementen van de struct array in het veld db.cmb.

db.cmb

+-----sLoccod	(String)
	Locatie code.
+-----sParcod	(String)
	Parameter code.
+-----reeks	(Struct array)
+-----sLoccod	Locatie en parameter van de tijdreeksen die
+-----sParcod	gelijktijdig getoond moeten worden met de
	huidige tijdreeks.

3.6 db.donarlijst

Melissa dient voor een belangrijk deel om tijdreeksen, afkomstig van het MSW en opgeslagen in DONAR, te controleren en te valideren. Na deze bewerking moeten de gegevens weer kunnen worden opgeslagen in DONAR. In het databasebeheer scherm bestaat de mogelijkheid voor de gebruiker om tijdreeksen rechtstreeks uit DONAR op te halen door de op te halen parameter, de locatie en de periode te specificeren. In het veld db.donarlijst wordt deze selectie opgeslagen.

db.donarlijst

+-----sLoccod	(String)
	Locatiecode zoals gebruikt in Donar.
+-----sLocoms	(String)
	Locatieomschrijving zoals gebruikt in Donar.
+-----sCrdtyp	(String)
	Coördinatenstelsel zoals gebruikt in Donar.
+-----lXcrdgs	(Double)
	X-coördinaat van de locatie.
+-----lYcrdgs	(Double)
	Y-coördinaat van de locatie.
+-----datanaam	(String)
	Parameter.

+----datadesc	(Struct)
+----MUX	de databeschrijving regels zijn de DONAR
+----WNS	regels die een datatype beschrijven, zoals
+----OGI	WNS of MUX, en eventuele aanvullende
	voorwaarden, zoals bemonsterende instantie
	(BMI).
+----van	(Double)
	Begin van de op te halen periode, in Matlab
	datum formaat.
+----tot	(Double)
	Eind van de op te halen periode, in Matlab
	datum formaat.

3.7 db.cnscriteria

Melissa biedt de mogelijkheid om een aantal relaties te specificeren tussen de gemeten waterhoogte op de verschillende meetstations. Inconsistenties die op grond van deze relaties optreden worden weergegeven als outliers. Deze relaties kunnen gespecificeerd worden in een stuurfile met extensie .cns. In het veld db.cnscriteria worden deze relaties opgeslagen in een struct array waarin elk element een relatie tussen twee stations voorstelt en waarin het maximale toegestane verschil tussen deze twee stations wordt gedefinieerd.

db.cnscriteria	
+----up	(String)
	Station met hoogste waterstand.
+----down	(String)
	Station met laagste waterstand.
+----downplus	(Double)
	Toegestaan verschil in waterstand.

3.8 db.valcriteria

Binnen Melissa kan naast een controle op consistentie ook een controle op de validatiestatus uitgevoerd worden. De instellingen voor de controle op de validatiestatus worden gelezen uit een file met extensie .val. Deze file bevat een aantal regels waarmee de validatie criteria worden ingevoerd. Elke regel begint met een stationsnaam en parameter naam, gevolgd door een factor, gevolgd door een maximale absolute standaardafwijking. Deze instellingen worden opgeslagen in het veld db.valcriteria.

db.valcriteria	
+----station	(String)
	Locatienaam.
+----parameter	(String)
	Parameternaam.

+----stdafwfact	(Double) Wordt gebruikt in een berekening waarin de absolute waarde van W-V wordt vergeleken met deze factor maal de standaardafwijking behorende bij het model waarmee de V waarde is berekend. Een waarde 0 wordt altijd genegeerd.
+----stdafwconst	(Double) De maximale absolute waarde wordt gebruikt indien de absolute waarde van W-V wordt beoordeeld. Ook hier wordt de waarde 0 genegeerd.

3.9 db.stuurfiles

Voor een aantal functies en instellingen binnen Melissa wordt gebruik gemaakt van zogenaamde stuurfiles. Dit zijn bestanden waarin een aantal instellingen zijn vastgelegd, zodat ze steeds opnieuw kunnen worden gebruikt. Ieder type stuurfile is te herkennen aan een aparte. Het veld db.stuurfiles bevat de inhoud van deze stuurfiles in de vorm van een tablefile object. Dit object maakt het mogelijk voor de gebruiker om deze stuurfiles binnen Melissa aan te passen en eventueel weer op te slaan.

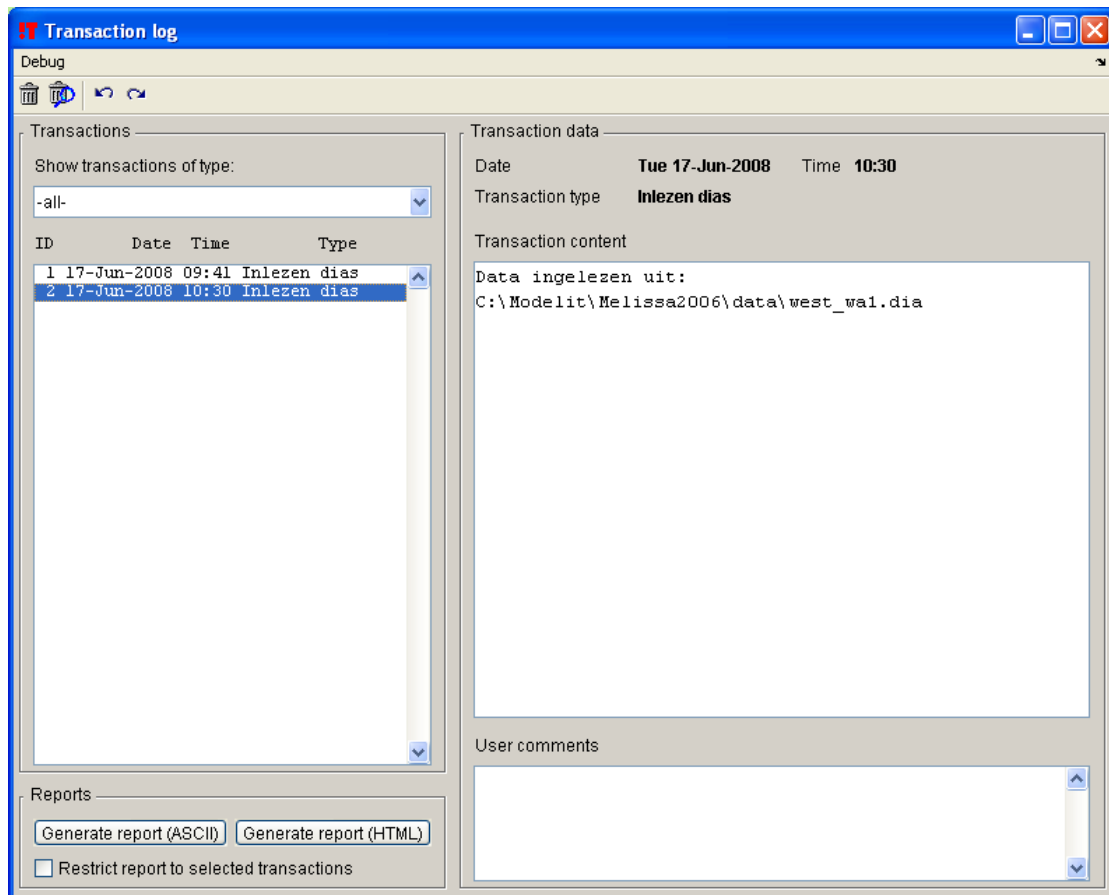
+----cns	(String) Naam van de gebruikt stuurfile met consistentiecriteria, zie ook sectie 3.7.
+----val	(Tablefile object) Stuurfile voor het instellen van de validatiecriteria. Zie ook sectie 3.8.
+----mrc	(Tablefile object) Stuurfile voor het uitvoeren van een meetreekscontrole.
+----hooglaag	(Tablefile object) Stuurfile voor het bepalen van hoog- en laagwater.
+----agger	(Tablefile object) Stuurfile voor het bepalen van aggers, hoog- en laagwater.
+----gdps	(Tablefile object) Stuurfile voor het aanmaken van getijparameterdatasets.
+----zuiveren	(Tablefile object) Stuurfile voor het zuiveren van aggers.
+----vliegplan	(Tablefile object) Stuurfile voor het bepalen van de onderschrijding conform vliegplan.
+----tussenschr	(Tablefile object)

	Stuurfile voor het bepalen van over- en onderschrijding.
+----havengetal	(Tablefile object) Stuurfile voor het bepalen van het gemiddelde havengetal.

3.10 db.transaction

Melissa is voorzien van een logboek waarin rapportages van uitgevoerde bewerkingen worden opgeslagen. Deze functionaliteit wordt geïmplementeerd door een standaard functie die tevens bepalend is voor de gebruikte opslagstructuur. Het veld `db.transaction` bevat een struct array met de opgeslagen logboek aantekeningen.

<code>db.transaction</code>	
+----date	(Double) Tijdstip, in Matlab datenum formaat, waarop de aantekening is gemaakt.
+----content	(Char array) Volledige tekst van het opgeslagen rapport. Deze tekst kan uit meerdere regels bestaan.
+----type	(String) Soort aantekening. Dit is een aanduiding die het mogelijk maakt bepaalde logboek aantekeningen snel op te sporen.
+----comment	(Char array) Tekst die door de gebruiker aan het logboek is toegevoegd. Deze tekst kan uit meerdere regels bestaan.



Figuur 1: Logboek weergave.

4 Overzicht projecten en documentatie

4.1 Inleiding

Melissa is een applicatie voor het off-line valideren van waterstands metingen in rivieren en voor de Hollandse kust. Tevens worden met het pakket de parameters berekend die voor de on-line validatie nodig zijn. Bovenop het datamodel van Melissa zijn naderhand nog een aantal losse modules voor diverse berekeningen gebouwd, de zogenaamde getijspecials. Alhoewel Melissa ontworpen is als applicatie voor het verwerken van tijdsreeksen van waterstandsmetingen, blijkt het pakket in de praktijk ook te voldoen en gebruikt te worden voor het inspecteren en valideren van andere plaatsgebonden tijdreeksen, zoals temperatuurmetingen.

De aanleidingen voor de implementatie van Melissa waren destijds

- Het niet milleniumproof zijn van bestaande software
- De wens om een verzameling losse Fortran programma's te integreren;
- De wens de bediening te vereenvoudigen na de pensionering van de gebruiker van de "oude" software.

In 1999 is een FO voor Melissa opgesteld, gevolgd door een realisatie traject. De realisatie is tot stand gekomen via een serie van steeds in meer detail uitgewerkte prototypes. Na de introductie en nazorg periode is heeft het onderhoud eerste een aantal jaren bij RIKZ-intern plaatsgevonden.

In 2006 heeft Modelit Melissa in onderhoud genomen, na een re-engineering te hebben toegepast op het door RIKZ doorontwikkelde pakket. Bij deze re-engineering is Melissa gebaseerd op de Modelit toolboxes en geschikt gemaakt voor compilatie tot standalone applicatie. Vanaf deze datum is de revisiehistorie bijgehouden.

Dit document geeft een overzicht van de beschikbare documentatie. Deze bestaat enerzijds uit Projectplannen, Functionele Ontwerpen en Handleidingen en anderzijds uit broncode documentatie. Een groot deel van deze documentatie is on-line te vinden vanuit het helpcenter van de applicatie.

4.2 Historisch overzicht uitgevoerde projecten met relatie naar Melissa

Offertedatum	Titel	Document
1999	Functioneel Ontwerp Melissa	Offerte door XI-advies
1999	Implementatie Melissa	projectplan met tijdraming: PJ Dekker en NJ van der Zijpp (1999) Melissa, Functioneel Ontwerp, inclusief aanvullingen, juli 1999, Xi visualisatie [MLRTI.doc, 61 blz]
19-07-2006	Groot Onderhoud Melissa	Offerte voor de re-engineering van Melissa [offerteMelissa20060719.doc]
26-06-2006	Documentatie project	Offerte voor het documentatie project waarvan het huidige document 1 van de deliverables is. [offerte20070626.doc, 11 blz]

4.3 Beschikbaarheid documentatie

Datum laatste versie	Type	Herkomst	Titel	Omschrijving	Online
01-09-2001	Functioneel ontwerp	Project "FO"	Melissa, Functioneel Ontwerp, inclusief aanvullingen	projectplan met tijdraming: PJ Dekker en NJ van der Zijpp (1999) Melissa, Functioneel Ontwerp, inclusief aanvullingen, juli 1999, Xi visualisatie [MLRTI.doc, 61 blz]	Nee
24-06-2008	Technische documentatie	Project "Implementatie"	Melissa, Technische documentatie	Beschrijving (oorspronkelijke) schermopvolgschema, rekenmodules, onderlinge functieafhankelijkheden en tags. [TechdocRevised.doc, 23 blz]	Nee
24-06-2008	Handleiding	Project "Implementatie"	Melissa, Handleiding	Uitgebreide handleiding, inclusief beschrijving achtergronden en fileformaten [MelissaManual.doc, 76 blz]	Ja
actueel	Revisie historie	Beheer en onderhoud	Revisie historie	Revisie historie, bijgehouden vanaf 3 Nov 2006. [revisionsMelissa.txt]	Nee
27-10-2000	Technische Specificatie	Project "Getijspecials"	Getijspecial, Technische Specificatie	In feite FO, TO userguide en handling in één [gtyTechSpecc, 57 blz]	Nee
27-10-2000	Broncode Documentatie	Project "Getijspecials"	Getijspecial, Documentatie	Toelichtingen, Beschrijving afhankelijkheden en Manual pages [gtyDocumentatie.doc, 32 blz]	Nee

