



UPPSALA
UNIVERSITET

UPTEC W13030

Examensarbete 30 hp
September 2013

Utformning av mjukvarusensorer för avloppsvatten med multivariata analysmetoder

Sandra Abrahamsson

REFERAT

Utformning av mjukvarusensorer för avloppsvatten med multivariata analysmetoder

Sandra Abrahamsson

Varje studie av en verklig process eller ett verkligt system är baserat på mätdata. Förr var den tillgängliga datamängden vid undersökningar ytterst begränsad, men med dagens teknik är mätdata betydligt mer lättillgängligt. Från att tidigare enbart haft få och ofta osammanhängande mätningar för någon enstaka variabel, till att ha många och så gott som kontinuerliga mätningar på ett större antal variabler. Detta förändrar möjligheterna att förstå och beskriva processer avsevärt.

Multivariat analys används ofta när stora datamängder med många variabler utvärderas. I det här projektet har de multivariata analysmetoderna PCA (principalkomponentanalys) och PLS (partial least squares projection to latent structures) använts på data över avloppsvatten insamlat på Hammarby Sjöstadsvärk.

På reningsverken ställs idag allt hårdare krav från samhället för att de ska minska sin miljöpåverkan. Med bland annat bättre processkunskaper kan systemen övervakas och styras så att resursförbrukningen minskas utan att försämra reningsgraden. Vissa variabler är lätta att mäta direkt i vattnet medan andra kräver mer omfattande laboratorieanalyser. Några parametrar i den senare kategorin som är viktiga för reningsgraden är avloppsvattnets innehåll av fosfor och kväve, vilka bland annat kräver resurser i form av kemikalier till fosforfällning och energi till luftning av det biologiska reningssteget. Halterna av dessa ämnen i inkommande vatten varierar under dygnet och är svåra att övervaka.

Syftet med den här studien var att undersöka om det är möjligt att utifrån lättmätbara variabler erhålla information om de mer svårsmätbara variablerna i avloppsvattnet genom att utnyttja multivariata analysmetoder för att skapa modeller över variablerna. Modellerna kallas ofta för mjukvarusensorer (soft sensors) eftersom de inte utgörs av fysiska sensorer.

Mätningar på avloppsvattnet i Linje 1 gjordes under tidsperioden 11 – 15 mars 2013 på flera ställen i processen. Därefter skapades flera multivariata modeller för att försöka förklara de svårsmätbara variablerna. Resultatet visar att det går att erhålla information om variablerna med PLS-modeller som bygger på mer lättillgänglig data. De framtagna modellerna fungerade bäst för att förklara inkommande kväve, men för att verkligen säkerställa modellernas riktighet bör ytterligare validering ske.

Nyckelord: multivariat analys, PCA, principalkomponentanalys, PLS, partial least squares projection to latent structures, mjukvarusensor, avloppsreningsverk, Hammarby Sjöstadsvärk.

*Institutionen för informationsteknologi, Uppsala universitet
Box 337, SE-751 05 Uppsala*

ABSTRACT

Design of soft sensors for wastewater with multivariate analysis

Sandra Abrahamsson

Studies of real processes are based on measured data. In the past, the amount of available data was very limited. However, with modern technology, the information which is possible to obtain from measurements is more available, which considerably alters the possibility to understand and describe processes. Multivariate analysis is often used when large datasets which contains many variables are evaluated.

In this thesis, the multivariate analysis methods PCA (principal component analysis) and PLS (partial least squares projection to latent structures) has been applied to wastewater data collected at Hammarby Sjöstadsverk WWTP (wastewater treatment plant).

Wastewater treatment plants are required to monitor and control their systems in order to reduce their environmental impact. With improved knowledge of the processes involved, the impact can be significantly decreased without affecting the plant efficiency. Several variables are easy to measure directly in the water, while other require extensive laboratory analysis. Some of the parameters from the latter category are the contents of phosphorus and nitrogen in the water, both of which are important for the wastewater treatment results. The concentrations of these substances in the inlet water vary during the day and are difficult to monitor properly.

The purpose of this study was to investigate whether it is possible, from the more easily measured variables, to obtain information on those which require more extensive analysis. This was done by using multivariate analysis to create models attempting to explain the variation in these variables. The models are commonly referred to as soft sensors, since they don't actually make use of any physical sensors to measure the relevant variable.

Data were collected during the period of March 11 to March 15, 2013 in the wastewater at different stages of the treatment process and a number of multivariate models were created. The result shows that it is possible to obtain information about the variables with PLS models based on easy-to-measure variables. The best created model was the one explaining the concentration of nitrogen in the inlet water.

Keywords: multivariate analysis, PCA, principal component analysis, PLS, partial least squares projection to latent structures, soft sensor, WWTP, Hammarby Sjöstadsverk.

*Department of Information Technology, Uppsala University
Box 337, SE-751 05 Uppsala*

FÖRORD

Den här rapporten omfattar 30 hp och har genomförts som en avslutning på utbildningen till civilingenjör i Miljö- och vattenteknik vid Uppsala universitet. Examensarbetet har utförts i samarbete med IVL Svenska Miljöinstitutet med handledning från Linus Bergfors. Ämnesgranskare var Bengt Carlsson, professor vid Institutionen för informationsteknologi, Uppsala universitet.

Jag vill tacka min handledare Linus Bergfors för stöd och uppmuntran under projektets gång samt ämnesgranskare Bengt Carlsson för goda råd och idéer. Jag vill även tacka Christian Baresel, Lars Bengtsson, Elin Ottosson och övriga medarbetare på Hammarby Sjöstadswerk för all den hjälp jag fått för att kunna genomföra mina mätningar på Linje 1. Slutligen vill jag tacka de som läst och lämnat synpunkter på rapporten.

Sandra Abrahamsson
Stockholm, 2013

POPULÄRVETENSKAPLIG SAMMANFATTNING

Utformning av mjukvarusensorer för avloppsvatten med multivariata analysmetoder

Sandra Abrahamsson

På Henriksdalsberget i Stockholm ligger Hammarby Sjöstadsvverk som är en försöks- och demonstrationsanläggning för framtidens VA-teknik. Rening av avloppsvatten är en viktig miljöfråga, detta eftersom en otillräcklig reduktion av föroreningar kan ge allvarliga konsekvenser hos recipienten. Bland annat genom övergödning av våra sjöar, vattendrag och kustnära områden.

Avloppsvatten släpptes förr helt orenat ut från städerna. Det var i början på 1940-talet som de första kommunala avloppsreningsverken byggdes i Sverige. På 1970-talet skedde en stor utbyggnad av reningsverken och utsläppen till recipient förbättrades avsevärt. Idag är de flesta hushåll i tätorterna anslutna till kommunala avloppsreningsverk och reningsgraden är oftast hög. I reningsverken avlägsnas fasta partiklar, organiskt material och närsalter från vattnet med hjälp av olika reningssteg. Processerna som tillämpas kallas för mekanisk, kemisk och biologisk rening samt filtrering.

I takt med att samhället blir mer miljöfokuserat ställs hårdare krav på avloppsreningsverken för att minska sin miljöpåverkan. Energieffektiviseringar och en högre reningsgrad är utmaningar som många verk står inför. Med bland annat bättre processkunskaper kan systemen övervakas och styras så att resursutnyttjandet minskar. Några parametrar som är viktiga för reningsgraden är innehållet av näringsämnen fosfor och kväve, vilka påverkar resursanvändningen i form av bland annat kemikalier till fosforfällning och energi till luftning av det biologiska reningssteget. Halterna i inkommande vatten av dessa ämnen varierar under dygnet och kan därför vara svåra att mäta direkt (on-line) med fysiska sensorer. Sensorerna som finns tillgängliga är ofta kostsamma och kräver kontinuerligt underhåll. Därför analyseras variablerna istället i laboratorium, för det mesta med dagliga eller veckovisa provtagningar. Eftersom det kan dröja upp till flera dagar från att provet tagits till att det analyserats är det väldigt svårt att ha uppsikt över processen och kunna göra justeringar i tid.

Varje studie av en verklig process eller ett verkligt system är baserat på mätdata. Förr var den tillgängliga datamängden vid undersökningar ytterst begränsad, men med dagens teknik är mätdata betydligt enklare att få fram. Multivariat analys består av statistiska metoder som behandlar två eller fler variabler som en enda enhet och försöker skapa ett övergripande resultat genom att ta hänsyn till förhållandet mellan variablerna. Multivariata analysmetoder kan användas på avloppsreningsverk för att erhålla information om de variabler som är svåra att mäta i avloppsvattnet genom att utgå från mer lättillgängliga variabler. Genom antagandet att variablerna i avloppsvattnet på något sätt är beroende av varandra kan multivariata modeller över avloppsvattnet tas fram. De svårsmätbara variablerna kan då förklaras från de lättsmätbara variablerna. Modellerna kallas ofta för mjukvarusensorer (soft sensors) eftersom de inte utgörs av fysiska sensorer.

I det här projektet har de multivariata analysmetoderna PCA (principalkomponentanalys) och PLS (partial least squares projection to latent structures) använts på data över avloppsvatten insamlat på Hammarby Sjöstadsverk för att försöka skapa modeller som kan förklara de svårämbara variablerna i avloppsvattnet genom att analysera mer lättillgängliga variabler. Variabler som anses vara lätta att mäta direkt i avloppsvattnet är bland annat temperatur, pH, syrehalt, redoxpotential och konduktivitet. Variabler som anses vara svåra att mäta är totalfosfor, fosfatfosfor, totalkväve, ammoniumkväve, nitratkväve och COD (kemisk syreförbrukning). Insamlingen av data skedde på flera olika ställen i processen under en mätkampanj som pågick under tidsperioden 11 mars - 15 mars 2013.

Resultatet visar att det går att erhålla information om variablerna med PLS-modeller som bygger på mer lättillgänglig data. De framtagna modellerna fungerade bäst för att förklara inkommande kväve, men för att verkligen säkerställa modellernas riktighet bör ytterligare validering ske.

ORDLISTA

Anoxisk: utan tillgång till fritt syre.

Autotrof organism: en organism som använder icke-organiskt material för sin tillväxt.

COD (*chemical oxygen demand*): förkortning för kemisk syreförbrukning som är ett mått på förekomsten av organiskt material i vattnet.

Heterotrof organism: en organism som använder organiskt material för sin tillväxt.

Konduktivitet: beskriver ett materials förmåga att leda elektrisk ström.

Latent variabel: variabel som utan att själv vara observerbar används för att beskriva observerade variabler.

Loadings: parameter som används i multivariata analysmetoder som beskriver hur de nya latent variablerna relateras till de ursprungliga variablerna.

Multivariat analys: samlingsnamn för statistiska metoder för beskrivning och analys av mångdimensionella datamängder.

Omsättningstid: ett mått på den tid ett ämne tillbringar i en viss del av sitt kretslopp.

PCA (*principal component analysis*): förkortningen för en multivariat analysmetod som kallas för principalkomponentanalys.

PLS (*partial least squares projection to latent structures*): förkortningen för en multivariat analysmetod som kallas för partial least squares projection to latent structures.

Recipient: plats dit det renade avloppsvattnet släpps (ex. sjö, havsvik eller vattendrag).

Redoxpotential: potential som utgör ett mått på den intensitet reduktion och oxidation sker vid en redoxprocess, där ett ämne reduceras och ett annat oxideras.

Scores: parameter som används i multivariata analysmetoder som beskriver observationernas koordinater i det nya koordinatsystemet.

SS (*suspended solids*): förkortning för suspenderat material och beskriver andelen fasta partiklar i vattnet.

INNEHÅLLSFÖRTECKNING

1	Inledning	1
1.1	Syfte.....	3
2	Bakgrund.....	4
2.1	Hammarby Sjöstadsverk.....	4
2.1.1	Linje 1.....	4
2.2	Multivariat analys.....	6
2.2.1	PCA.....	6
2.2.2	PLS.....	7
2.2.3	Programvara för multivariat analys.....	8
3	Material och Metod.....	15
3.1	Mätkampanj	15
3.1.1	Förändringar i flöden.....	15
3.1.2	Datainsamling och provtagning	15
3.1.3	Analyser.....	17
3.2	Tidsförskjutning	18
3.3	Modellering i SIMCA.....	18
4	Resultat och observationer	20
4.1	Mätkampanj	20
4.2	Tidsförskjutning	22
4.3	Modellering i SIMCA.....	23
4.3.1	Inkommande fosfor	23
4.3.2	Inkommande kväve	27
4.3.3	Inkommande COD	31
4.3.4	Nitrat i bireaktorerna	33
5	Diskussion	34
5.1	Mätkampanj	34
5.2	Tidsförskjutning	35
5.3	Modellering i SIMCA.....	35

5.3.1	Inkommande fosfor	35
5.3.2	Inkommande kväve	36
5.3.3	Inkommande COD	36
5.3.4	Nitrat i bioreaktorerna	37
5.3.5	Allmänt	37
6	Slutsatser	39
7	Referenser.....	40
Bilaga A	42
Bilaga B	43
Bilaga C	48

1 INLEDNING

I takt med att samhället blir mer miljöfokuserat ställs hårdare krav på avloppsreningsverken för att minska sin miljöpåverkan. Energieffektiviseringar och en högre reningsgrad är utmaningar som många verk står inför. Med bland annat bättre processkunskaper kan systemen övervakas och styras så att resursutnyttjandet minskar. Några parametrar som är viktiga för reningsgraden är innehållet av fosfor och kväve, vilka påverkar resursanvändningen i form av bland annat kemikalier till fosforfällning och energi till luftning av det biologiska reningssteget. Halterna i inkommande vatten av dessa ämnen varierar under dygnet och kan därför vara svåra att mäta direkt (on-line) med fysiska sensorer. Sensorerna som finns tillgängliga är ofta kostsamma och kräver kontinuerligt underhåll. Därför analyseras variablerna istället i laboratorium, för det mesta med dagliga eller veckovisa provtagningar. Eftersom det kan dröja upp till flera dagar från att provet tagits till att det analyserats är det väldigt svårt att ha uppsikt över processen och kunna göra justeringar i tid.

Multivariata analysmetoder kan användas för att erhålla information om de variabler som är svåra att mäta genom att utgå från mer lättillgängliga variabler. Genom antagandet att variablerna i avloppsvattnet på något sätt är beroende av varandra kan multivariata modeller över avloppsvattnet tas fram. De svårsmätbara variablerna kan då förklaras från de lättsmätbara variablerna. Modellerna kallas ofta för mjukvarusensorer (soft sensors) eftersom de inte utgörs av fysiska sensorer. Multivariat analys är ett samlingsbegrepp för flera olika statistiska metoder som tillämpas inom många områden när stora datamängder med många variabler utvärderas.

Det finns en mängd tidigare studier där mjukvarusensorer har tillämpats. Aguado m.fl. (2006) använde i Valencia, Spanien olika prediktionsmodeller för att uppskatta näringsinnehållet hos avloppsvattnet i en SBR (Sequencing Batch Reactor). Flera olika modelltyper undersöktes och utvärderades. En nackdel med den framtagna modellen var att den var tvungen att vänta tills slutet av varje cykel (batch) innan prediktionerna kunde beräknas. Detta kunde lösas genom att saknade värden ersattes med medelvärden från testdata.

I Alabama, USA har Galcia m.fl. (2011) använt mjukvarusensorer på ett pappersbruk för att prediktera processvariabler för behandlingen av träflis. Modellen lyckades fånga dynamiken i processen tillräckligt väl för att senare kunna implementeras och användas som en enkel och praktisk lösning i industrin. Dock krävdes vidareutvecklingar för att modellen skulle kunna användas under varierande förhållanden, till exempel om koktemperaturen eller träslaget ändrades.

Lumley (2002) har i Göteborg, Sverige undersökt hur mjukvarusensorer kan användas och implementeras på Rya reningsverk för att erhålla en bättre kontroll av processen. Bland annat har utgående totalfosforhalt studerats. Som en kontroll jämfördes mjukvarusensorn för totalfosfor med en fysisk on-linesensor och ett alarm utlöstes om de avvek för mycket från varandra. Dock krävdes det att en faktor i modellen som påverkade turbiditeten uppdaterades allteftersom processen förändrades. Vanligtvis beräknades den veckovis men för en bättre modell krävdes tätare uppdateringar.

Andersson m.fl. (2003) har i Frövi, Sverige med hjälp av multivariat analys tagit fram processinformation hos en långtidsluftad aktivslamanläggning. En långtidsluftad aktivslamanläggning är vanlig inom skogsindustrin och syftet var att skapa ett system för övervakning och styrning av processen. Resultaten från den multivariata analysen visade att under provtagningen har anläggningen haft tre olika lägen, då utsläppen av bland annat totalfosfor varierat. Mjukvarusensorer för olika variabler konstruerades och lyckades prediktera variablerna väl och implementerades därefter i anläggningen och övervakar nu systemet i realtid.

Nilsson m.fl. (2007) har under åren 2002-2006 undersökt möjligheterna att ta fram resurseffektivare avloppsvattenreningstekniker i Sverige. Under projektet användes bland annat multivariat analys för att genomföra processidentifieringar för att erhålla en bättre övervakning av avloppsreningsverk i Västerås och Örebro. Resultaten visade att de multivariata modellerna hade en bra förklaringsgrad mellan processens olika variabler. Exempelvis påvisades att avloppsvattnets temperatur ger en högre elförbrukning för luftning av det biologiska reningssteget då vattentemperaturen är hög jämfört med om den är låg. I ett annat delprojekt har en undersökning av mjukvarusensorer på Henriksdals reningsverk i Stockholm studerats. Flera mätkampanjer genomfördes för att samla in data. Ett antal mjukvarusensorer togs fram på nyckelvariabler i det inkommande avloppsvattnet med goda resultat. Mjukvarusensorerna implementerades sedan för övervakning på reningsverket. Men med vidareutvecklingar skulle modellerna även kunna tillämpas för att reglera mängden doserad fällningskemikalie i förfällningen.

1.1 SYFTE

Under examensarbetet samlades data in från on-linesensorer och vattenprover tagna på avloppsvattnet från Linje 1 på Hammarby Sjöstadsverk. Därefter används multivariata analysmetoder till att skapa modeller för att förstå och förutspå svårsmätbara variabler i avloppsvattnet genom att analysera mer lättillgängliga variabler.

Variabler som anses vara lätta att mäta direkt i avloppsvattnet är bland annat temperatur, pH, syrehalt, redoxpotential och konduktivitet; dessa värden erhålls från on-linesensorer som finns installerade i reningsverket. Variabler som anses vara svåra att mäta är totalfosfor, fosfatfosfor, totalkväve, ammoniumkväve, nitratkväve och COD (kemisk syreförbrukning), värden för dessa tas fram genom analys av insamlade vattenprover i laboratorium.

Den här rapporten är en del av IVL Svenska Miljöinstitutets projekt om mjukvarusensorer som startade i juni 2012. Hittills har utrusning köpts in och trimmats, reningseffektiviteten studerats samt en första mätkampanj genomförts. Till den här studien har placeringen av vissa on-linesensorer ändrats för att optimera processövervakningen ytterligare. Dessutom har inflödet varierats mot att tidigare ha hållits konstant för att maximera variansen i insamlad data.

Målet med det här arbetet är att:

- Undersöka om det med multivariata analysmetoder är möjligt att utifrån lättmätbara variabler, erhålla information om mer svårsmätbara variabler på avloppsvattnet i Hammarby Sjöstadsverk.
- Ta fram och utvärdera modeller för insamlad data.

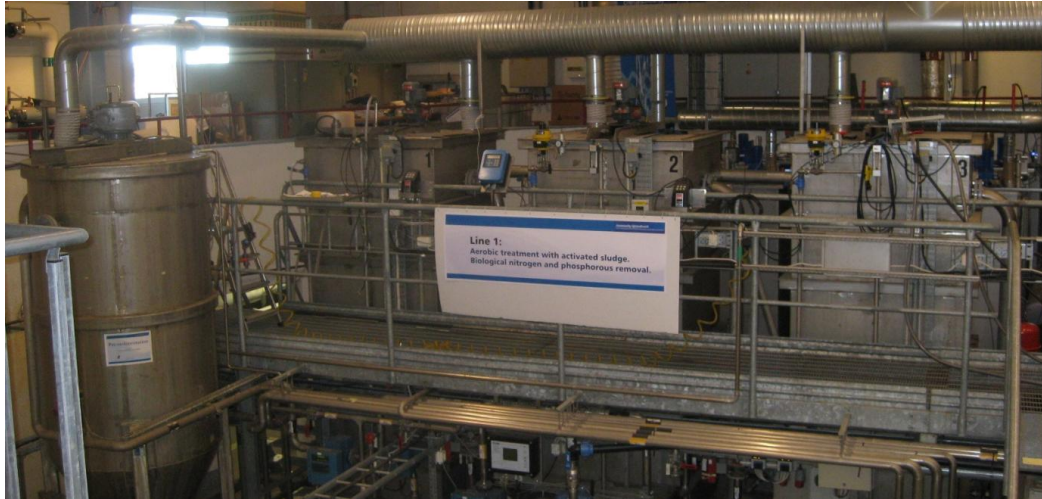
2 BAKGRUND

2.1 HAMMARBY SJÖSTADSVERK

Hammarby Sjöstadsvverk ligger på Henriksdalsberget i Stockholm. Anläggningen används för forskning inom VA-sektorn och ger möjlighet till samarbete mellan forskare, kommunala avloppsreningsverk och företag. Hammarby Sjöstadsvverk invigdes i början av 2000-talet och ligger i nära anslutning till Henriksdals reningsverk (Baresel m.fl., 2011). På reningsverket pågår långa forskningsprojekt men även kortare tester och utvecklingsarbeten. Hammarby Sjöstadsvverk fungerar utöver detta också som en demonstrations- och utbildningsanläggning för bland annat ny utrustning. Från och med 2008 har driften av reningsverket, som tidigare sköts av Stockholm Vatten, överlåtits till en FoU-grupp (Forskning och Utveckling) ledd av IVL Svenska Miljöinstitutet i samarbete med Kungliga Tekniska Högskolan (KTH) (Baresel m.fl., 2011). I dagsläget finns det på anläggningen sex försökslinjer installerade för att kunna applicera och utvärdera olika typer av processteg för behandling av avloppsvattnet. I detta projekt har endast Linje 1 använts, därför kommer övriga försökslinjer ej att behandlas.

2.1.1 Linje 1

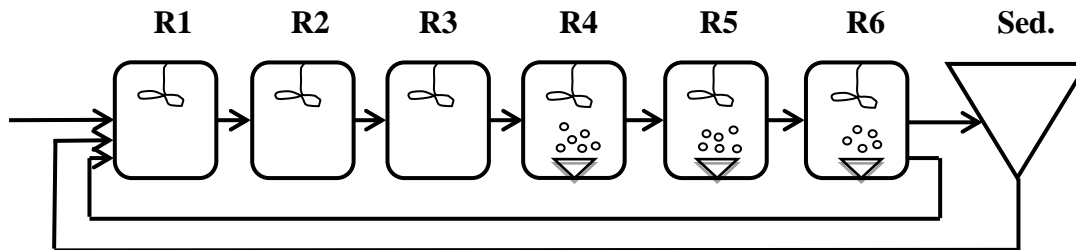
Linje 1 (Figur 1), är utformad som ett traditionellt svenskt reningsverk. Den är en mindre kopia av Henriksdals reningsverk och har en kapacitet på 150 personekvivalenter (Baresel, pers. medd.).



Figur 1. Hammarby Sjöstadsvverk, Linje 1, med försedimenteringsbassängen samt de tre första bioreaktorerna.

Avloppsvattnet in till Linje 1 kommer från Henriksdals reningsverk, där det enbart genomgått mekanisk grovrensning för att avskilja det allra grövsta materialet från vattnet såsom trasor, grus, sand och dylikt. Behandlingen av avloppsvattnet på Linje 1 inleds med en försedimenteringsbassäng på 2,53 m³ vars syfte är avlägsna organiska partiklar för att minska belastningen på efterföljande reningssteg (Svenskt Vatten, 2010). I samband med försedimenteringen sker även kemisk förfällning med järn(III)klorid för att en god avskiljning av fosfor ska uppnås. Därefter följer sex bioreaktorer för biologisk kväverening

(R1-R6, Figur 2). Samtliga bioreaktorer har en volym på 5,04 m³ och en omrörare installerad. De tre sista bioreaktorerna (R4-R6) är luftade. R4 och R5 har börvärden för syrehalten satta på 2,0 mg/l och R6 på 1,0 mg/l. Syrehalten är lägre i R6 eftersom risken för återföring av syrerikt vatten till R1 vid recirkulationen minskas.



Figur 2. Schematisk bild över den biologiska kvävereningen på Linje 1. Reningen består av sex bioreaktorer (R1-R6) där de tre sista är luftade, samt en sedimenteringsbassäng. I reningssteget finns två processflöden som återför vatten och aktivt slam till början av reningssteget: ett så kallat recirkulationsflöde från den sista bioreaktorn samt ett returslamflöde från sedimenteringsbassängen.

Kvävereduktionen är uppbyggd enligt så kallad fördenitrifikation; de första reaktorerna i reningssteget är oluftade. Kvävet avlägsnas från avloppsvattnet med hjälp av mikroorganismer som förekommer som brunaktiga slamflockar; aktivt slam. Framför allt är det bakterier som genom sin respiration omvandlar kvävet mellan olika former för att slutligen erhålla kvävgas som avgår till atmosfären. In till reningsverket kommer kväve framför allt som ammonium. I den luftade zonen (R4-R6) där tillgången på syre är god finns autotrofa bakterier som genom nitrifikation oxiderar ammonium till nitrit och därefter till nitrat. Autotrofa bakterier utnyttjar koldioxid som kolkälla (Svenskt Vatten, 2010).

Från R6 leds nitratiskt vatten via ett recirkulationsflöde tillbaka till R1. I den oluftade zonen (R1-R3), där tillgång till fritt syre saknas (anoxisk miljö), finns heterotrofa bakterier som genom denitrifikation reducerar kväve i form av nitrat till nitrit och slutligen till kvävgas. Heterotrofa bakterier utnyttjar organiska kolkällor som de får genom att bryta ner det organiska materialet i avloppsvattnet (Svenskt Vatten, 2010).

Efter de sex bioreaktorerna följer en sedimenteringsbassäng vars uppgift är att avskilja partiklar från det renade vattnet så att slammet i bassängen koncentreras (Svenskt Vatten, 2010). En del av det koncentrerade aktiva slammet pumpas sedan från sedimenteringsbassängen tillbaka till början av det biologiska kvävereningssteget, så kallat returslamflöde, för att upprätthålla en viss slammängd i systemet. Det slam som inte återförs tas ut som överskottsslam. När vattnet passerat genom sedimenteringsbassängen följer ett sandfilter för en slutlig partikelavskiljning. Hammarby Sjöstadsverk släpper inte ut något vatten till recipient, utan vattnet återförs till Henriksdals reningsverk för kompletterande rening när det passerat genom processen (Baresel, pers. medd.). Linje 1 har även möjlighet att använda kemisk efterfällning av fosfor. Efterfällningen var avstängd under projektet för att inte påverka mätningarna då backspolvatten brukar ledas tillbaka till försedimenteringen.

2.2 MULTIVARIAT ANALYS

Varje studie av en verklig process eller ett verkligt system är baserat på mätdata. Förr var den tillgängliga datamängden vid undersökningar ytterst begränsad, men med dagens teknik är mätdata betydligt enklare att få fram. Multivariat analys består av statistiska metoder som behandlar två eller fler variabler som en enda enhet och försöker skapa ett övergripande resultat genom att ta hänsyn till förhållandet mellan variablerna (Jackson, 1991). Då kan information om relationer mellan variabler och observationer eller grupper av observationer analyseras och avvikande observationer kan upptäckas. Multivariata analysmetoder tillämpas inom många områden, från neurovetenskap till datorgrafik (Shlens, 2005).

I detta avsnitt tas grunderna till de multivariata analysmetoderna principalkomponentanalys (PCA) och partial least squares projection to latent structures (PLS) upp. PCA är en metod som ger en bra överblick över data och är lämplig att utföra som en första analys, även om mer avancerade tekniker ska tillämpas senare. PLS är en metod för att finna samband mellan uppsättningar av variabler som senare används till exempelvis prediktion av önskvärda variabler.

2.2.1 PCA

Principalkomponentanalys (PCA) har ibland kallats för det mest värdefulla resultatet från tillämpad linjär algebra, då det är en enkel metod för att ta fram information från till synes oöverskådliga datamängder (Shlens, 2005). Med PCA möjliggörs identifiering av grupper av variabler som är inbördes relaterade via fenomen som inte kan observeras direkt.

Syftet med PCA är att minska dimensionen hos data som innehåller ett stort antal variabler, men samtidigt försöka bevara så mycket som möjligt av informationen som finns i data (Jolliffe, 2002). Metoden går ut på att de ursprungliga variabler som är delvis korrelerade med varandra överförs till ett mindre antal nya, okorrelerade variabler som kallas för principalkomponenter (PC). Varje PC är en linjär kombination av de ursprungliga variablerna (Yu m.fl., 2010). Principalkomponenter används till att upptäcka och tolka beroenden som existerar hos de ursprungliga variablerna. De kan även användas för att undersöka relationer mellan dem så att mycket av den ursprungliga informationen kan förklaras och presenteras.

Innan principalkomponenterna tas fram centreras och normeras data, så kallad autoskalning genom att dra bort medelvärdet och dividera med standardavvikelsen. PC kan exempelvis beräknas ur kovariansmatrisen, korrelationsmatrisen eller genom singularvärdesuppdelning (SVD). Dessa tekniker tas inte upp i detalj i denna rapport utan hänvisas till annan litteratur, till exempel Jackson (1991) eller Jolliffe (2002). Resultatet för beräkningarna av principalkomponenterna blir dock desamma, oberoende av vilken teknik som används för att ta fram dem.

Den första PC väljs som en linjärkombination av de ursprungliga variabler som har störst varians (Yu m.fl., 2010), det vill säga i den riktning i rummet där variansen är som störst. Eftersom den första PC ofta inte är tillräcklig för att ge en tillfredsställande beskrivning av data kan fler principalkomponenter tas fram. Den andra PC är ortogonal mot den första och

representerar linjärkombinationen av variablerna i riktningen med den näst största variansen. Ibland behövs ytterligare komponenter för att representera data tillräckligt bra. Det finns inget fastställt antal principalkomponenter som behöver tas fram, men många regler och indikatorer som kan användas som vägledning finns, se bland annat Yu (2010).

I modellen för PCA representeras data av en matris X , som har dimensionerna $n \times m$ där n är antalet rader, vilket motsvarar antalet observationer och m är antalet kolonner vilket motsvarar antalet variabler i X . Med PCA-modellen beskrivs enligt Björk (2007) matrisen X genom

$$X = TP^T + E \quad (1)$$

där T kallas för scores och är en matris som beskriver observationernas läge i det nya rummet. P kallas för loadings och är en matris med vikter som beskriver hur de nya komponenterna relateras till de ursprungliga variablerna. Matrisen E kallas för residualmatrisen och innehåller den del av data som inte förklaras med hjälp av de ingående principalkomponenterna, vanligen ses den som bruset i data (Yu m.fl., 2010). Med införda dimensioner blir ekv. (1) enligt Björk (2007) istället

$$\begin{bmatrix} X \\ n \end{bmatrix}^m = \begin{bmatrix} T \\ n \end{bmatrix}^{lv} \begin{bmatrix} P^T \\ m \end{bmatrix} + \begin{bmatrix} E \\ n \end{bmatrix}^m \quad (2)$$

där lv är antalet principalkomponenter.

2.2.2 PLS

Partial least squares projection to latent structures (PLS) är en metod för att beskriva förhållandet mellan två datamängder genom att ta fram en modell som söker sambandet mellan variabler i data X med variabler i data Y . Metoden är lämplig när X -variablerna är korrelerade och om det finns fler än en Y -variabel (Wold m.fl., 2001). Med PLS kan X -variablerna utnyttjas för att förklara och prediktera Y -variablerna.

I PLS antas, precis som i PCA, att de undersökta observerade data endast influeras av ett fåtal underliggande, latent, variabler som inte kan mätas direkt. PLS tar fram nya variabler (komponenter) som är uppskattningar av de latent variablerna, som, förutom att prediktera Y , även modellerar X (Wold m.fl., 2001). Det vill säga både X och Y antas delvis att beskrivas av samma latent variabler. PLS-komponenterna beräknas en i taget genom att maximera korrelationen mellan X och Y , se Wold m.fl (2001) för detaljer. Modellordningen det vill säga antalet PLS-komponenter tas fram iterativt genom att testa prediktionsförmågan hos de redan framtagna komponenterna och stoppa proceduren när ökningen i prediktionsförmåga inte är signifikant (Wold m.fl., 2001).

I modellen för PLS representeras data av två matriser X och Y . Matriserna har dimensionerna $n \times m$ respektive $n \times p$ där n är antalet observationer och m är antalet X -variabler och p är antalet Y -variabler. Med PLS-modellen beskrivs matriserna X och Y enligt Björk (2007) genom

$$X = TP^T + E \quad (3)$$

$$Y = UQ^T + F \quad (4)$$

där T och U kallas för scores för X respektive Y och är matriser som beskriver observationernas läge i det nya rummet. P och Q kallas för loadings för X respektive Y och är matriser med vikter som beskriver hur de nya komponenterna relateras till de ursprungliga variablerna. Matriserna E och F kallas för residualmatriser. Mellan T och U finns det ett inre förhållande enligt

$$U = bT + G \quad (5)$$

där b är lutningen som sammankopplar T med U och där G är en residualmatris. Vissa utvidgningar av PLS har istället ett icke-linjärt förhållande i ekv. (5) (Björk, 2007).

Med införda dimensioner blir ekv. (3) och ekv. (4) istället

$$\begin{bmatrix} X \\ n \end{bmatrix}^m = \begin{bmatrix} T \\ n \end{bmatrix}^{lv} \begin{bmatrix} P^T \\ m \end{bmatrix} + \begin{bmatrix} E \\ n \end{bmatrix}^m \quad (6)$$

$$\begin{bmatrix} Y \\ n \end{bmatrix}^p = \begin{bmatrix} U \\ n \end{bmatrix}^{lv} \begin{bmatrix} Q^T \\ p \end{bmatrix} + \begin{bmatrix} F \\ n \end{bmatrix}^m \quad (7)$$

där lv är antalet PLS-komponenter (Björk, 2007).

PLS finns i många olika varianter, exempelvis OPLS och O2PLS, men de behandlas inte i denna rapport.

2.2.3 Programvara för multivariat analys

Det finns flera olika programvaror för multivariat analys. I det här avsnittet behandlas programmet SIMCA från Umetrics som användes i projektet. All information i detta avsnitt är hämtad från användarhandboken Users Guide to SIMCA, version 13, MKS Umetrics (2012), som medföljer programmet.

I SIMCA tas multivariata modeller av olika modelltyp fram genom att använda data som importerats till programmet i ett *Workset*. Geometriskt representeras observationerna som punkter i ett multidimensionellt rum som spänns upp av variablerna. Det finns flera olika användarval när ett *Workset* finns tillgängligt i SIMCA. Först väljs vilka variabler som ska vara X -respektive Y -variabler. Därefter väljs en modelltyp till data och vilka observationer

som ska ingå. De modelltyper som används i modelleringen i denna rapport är PCA-X som skapar en modell av X -variablerna, PCA-Y som skapar en modell av Y -variabler och PLS som skapar en modell av både X - och Y -variabler, för övriga modelltyper som finns tillgängliga, se användarhandboken. I multivariat analys tas information fram från processdata genom att minska dimensionen i data med nya komponenter. Med scores beskrivs observationernas koordinater i det nya rummet. Dessa scores ger en god överblick över data i enkla grafer. Koefficienterna i projektionerna, det vill säga hur komponenterna i det nya rummet relateras till de ursprungliga variablerna kallas för loadings. Graferna med loadings visar bland annat variablernas betydelse, likheter och samband med varandra.

I SIMCA skalas alla variabler automatiskt till *Unit Variance* innan modelleringen startar. Då centreras variabeln genom att medelvärdet dras bort och skalas genom att variabeln divideras med standardavvikelsen. Variablerna kan även transformeras eller expanderas beroende på vad som ska undersökas. Variabelexpansion kan till exempel vara att kvadrera eller korsmultiplicera variabler med varandra.

För att anpassa en modell till data i SIMCA används *Autofit*, som bestämmer antalet komponenter (modellordningen) till modellen automatiskt. Metoden använder korsvalidering för att avgöra om tillägg av ytterligare en komponent i modellen kan anses vara signifikant. Antalet komponenter kan vid behov modifieras för att se hur modellen förändras. Korsvalideringen går ut på att dela upp observationerna i grupper och en och en hålla dem utanför modelleringsarbetet. De utslutna grupperna predikteras sedan från de kvarvarande grupperna. Därefter jämförs prediktionerna mot de utslutna värdena. Programmet använder ett antal regler för när ytterligare en komponent i modellen inte är relevant. Proceduren avbryts till exempel när den nya komponenten inte kan ge tillräckligt mycket mer ny information om variansen i datasetet; 2 % för PCA och 3 % för PLS. Som standard används sju grupper till korsvalidering.

När en modell skapats erhålls en sammanfattning där ett antal parametrar visas (Tabell 1). $Q^2(\text{cum})$ -värdet (härdaneftter kallat Q^2 -värdet) är en indikation på hur bra modellen predikterat. En modell med ett Q^2 -värde $> 0,5$ anses ha god prediktionsförmåga.

Tabell 1. Beskrivning av olika parametrar i SIMCA.

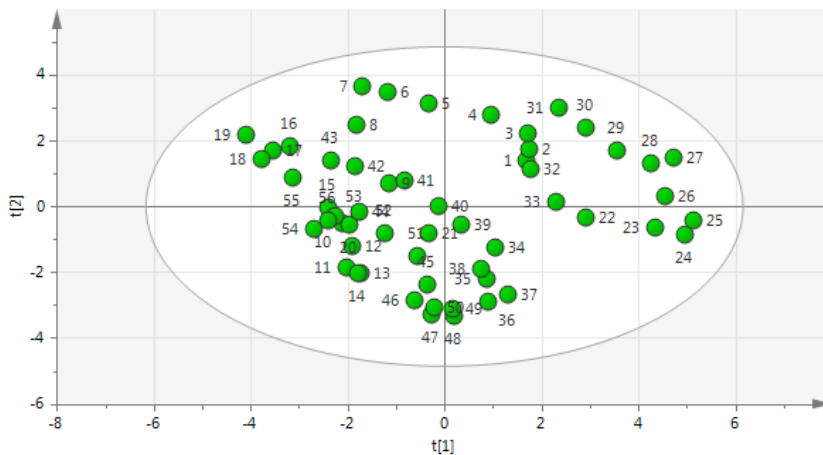
Parameter	Beskrivning
A	Antalet komponenter.
R^2X	Fraktionen av den totala variationen i X som förklaras av varje komponent.
$R^2X(\text{cum})$	Kumulativ R^2X i X som förklaras av alla de extraherade komponenterna.
R^2Y	Fraktionen av den totala variationen i Y som förklaras av varje komponent.
$R^2Y(\text{cum})$	Kumulativ R^2Y av Y som förklaras av alla de extraherade komponenterna.
$Q^2(X/Y)$	Fraktionen av den totala variationen i X (i PCA) och Y (i PLS) som kan predikteras av varje komponent.
$Q^2(\text{cum})$	Den kumulativa Q^2 som kan predikteras av alla extraherade komponenter.

När en modell har skapats finns en mängd olika grafer och tabeller som kan användas för fortsatt analys och utvärdering.

Scores Plot

För att hitta trender, grupperingar, avvikande observationer och andra mönster i data används scores-grafer. De byggs upp av ett antal scores-vektorer t_i som beräknats med hjälp av linjärkombinationer av X , se ekv. (1) och ekv. (3). Från t erhålls en sammanfattning av X som i PCA-modellen approximerar variationen i X och i PLS-modellen approximerar X och predikterar Y . För PLS-modellen finns även scores-vektorerna u , se ekv. (4), som används för att maximera korrelationen med t .

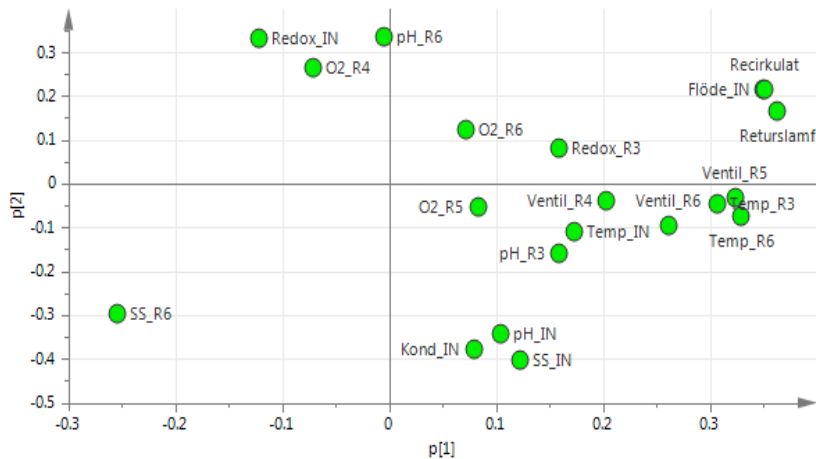
I graferna plottas scores mot varandra i två valda dimensioner, med t_1 mot t_2 eller u_1 mot u_2 i en 2D *Score Scatter Plot* (Figur 3). SIMCA ritar ut en ellips som motsvarar en signifikansnivå på 0,05. Observationer som hamnar utanför ellipsen kan anses vara avvikande.



Figur 3. En *Score Scatter Plot* som visar observationerna projicerade på komponenterna t_1 och t_2 i en PCA- X -modell.

Loadings Plot

För att hitta samband i korrelationen mellan variabler används loadings-grafer. Där p är loadings för X -delen av modellen, se (1) och (3). I en PCA-modell representeras loadings av koefficienterna som variablerna i X kombineras ihop med för att skapa scores t i det nya rummet. I en PLS-modell uttrycker p vikten av variablerna i approximationen av X för den valda komponenten. För PLS-modeller finns även vikter w och c som används i loadings-grafer för PLS. Vikterna kan kombineras på olika sätt. Variabler med stora absolutvärden på p_1 och/eller p_2 dominerar projektionen. Variabler placerade nära varandra är positivt korrelerade och variabler placerade motsatta varandra är negativt korrelerade (Figur 4).

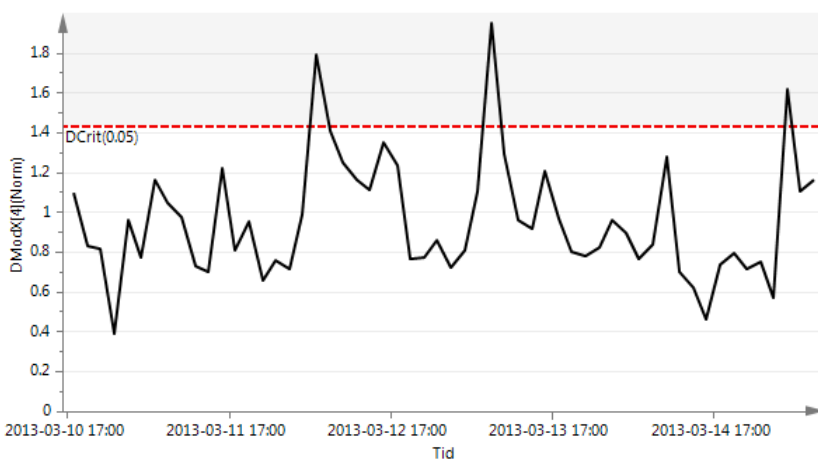


Figur 4. En *Loadings Scatter Plot* som visar variablerna projicerade på komponenterna p_1 och p_2 i en PCA-X-modell.

Värt att notera är att scores-grafer och loadings-grafer kompletterar varandra. Positionen hos en observation i en given plats i scores-grafen påverkas av variablerna som ligger i samma plats i loadings-grafen.

Distance to Model Plot

Residualer är avvikelser mellan observerad data och det skapade modellplanet. Med *Distance to Model* ($D_{ModX/Y}$) används de för att uppskatta hur långt ifrån modellplanet den sista komponenten i X - eller Y -rummet är belägen för en viss observation (Figur 5).

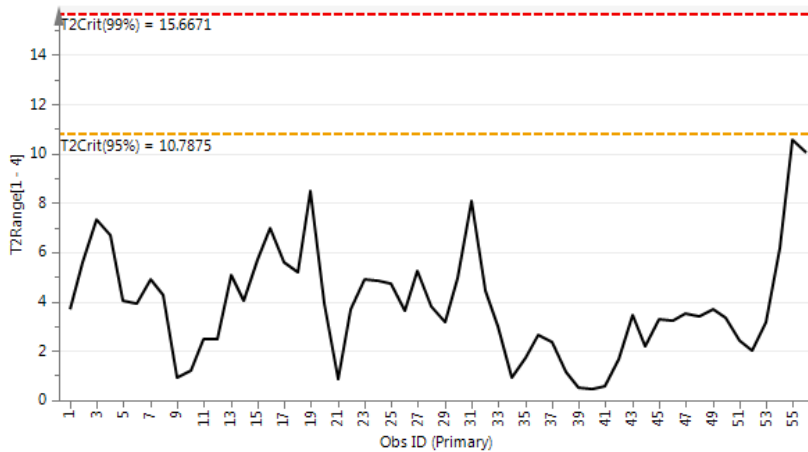


Figur 5. En D_{ModX} Plot för den sista variabeln i en PCA-X-modell.

När observationerna har ett värde på D_{ModX} eller D_{ModY} som är större än gränsen D_{crit} , som är beräknat med 0,05 signifikansnivå, anses de som avvikande. Om de är dubbelt så stora anses observationen vara mycket avvikande.

Hotelling's T2 Range Plot

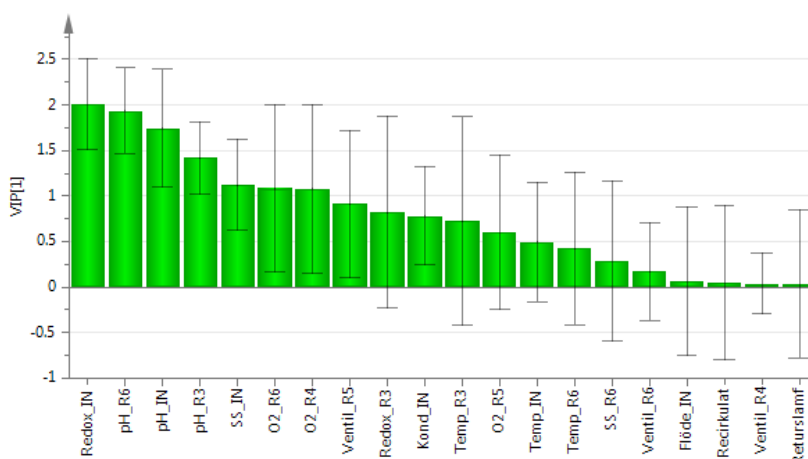
En *Hotelling's T2 Range Plot* visar avståndet från origo i modellplanet (scores) för alla observationer hos varje komponent samtidigt. Värden som är större än den gula gränsen (0,05 signifikansnivå) (Figur 6) kan misstänkas vara avvikande medan värden som är större än den röda gränsen (0,01 signifikansnivå) är tydligt avvikande och bör övervägas att exkluderas ur modellen om fler undersökningar indikerar detta.



Figur 6. En *Hotelling's T2 Range Plot* för en PCA-X-modell.

VIP Plot

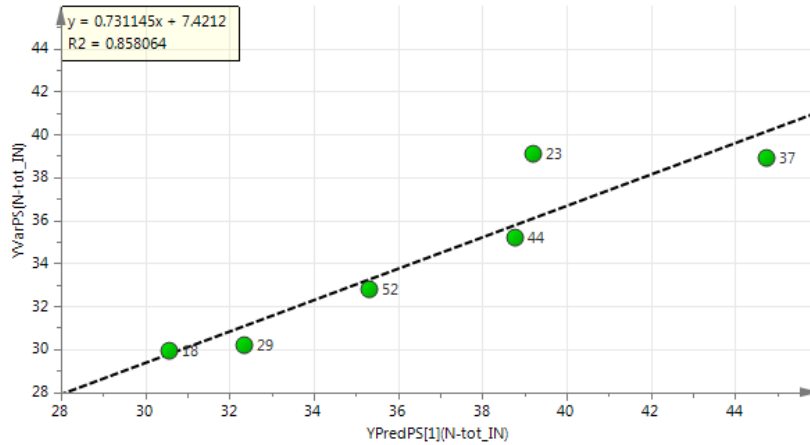
För PLS-modeller finns VIP-grafer (*Variable Influence on the Projection*) som visar hur betydelsefulla variablerna i X är för modellen för Y . En *VIP Plot* tar med hjälp av vikterna i PLS hänsyn till mängden förklarad varians för Y i varje dimension. Variabler med VIP-värde större än 1,0 är mest relevanta för att beskriva Y och variabler med VIP mindre än 0,5 saknar betydelse för modellen (Figur 7). Staplarna anger 95 % konfidensintervall.



Figur 7. En *VIP Plot* för en PLS-modell.

Y PS Plot

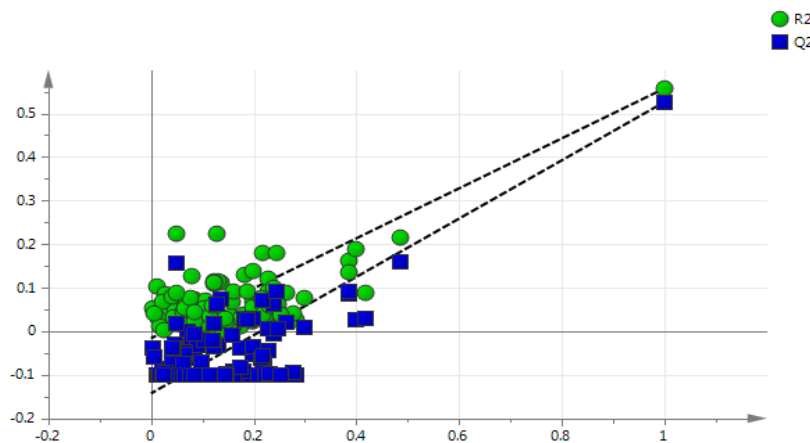
En *Y PS Plot* visar prediktionen för den valda *Y*-variabeln genom att plotta observerade värden för *Y* mot PLS-modellens predikterade värden för *Y* (Figur 8). Regressionslinjens R^2 -värde kan tas fram. När värdet är nära 1,0 anses prediktionen vara god.



Figur 8. En *Y PS Plot* för en PLS-modell.

Permutation Plot

För att studera modellens användbarhet och grad av överanpassning (overfit) används en *Permutation Plot*. Där utvärderas PLS-modellen för att se om den enbart passar data som använts för att skapa modellen, och därmed inte är lämplig för att prediktera *Y* för nya observationer. Valideringen sker genom att ursprungsmodellen jämförs med flera andra modeller vars ordning hos *Y*-variablerna slumpmässigt flyttats om (permutationer), medan *X*-matrisen behålls intakt. Grafen (Figur 9) visar ursprungsmodellens $R^2Y(\text{cum})$ - och Q^2 -värde till höger och de permuterade modellernas värden till vänster. Den horisontella axeln visar korrelationen mellan de permuterade *Y*-vektorer och den ursprungliga *Y*-vektorn. *Y* har korrelationen 1,0 med sig själv vilket syns uppe till höger. Den vertikala axeln visar $R^2Y(\text{cum})$ - och Q^2 -värdet.



Figur 9. En *Permutation Plot* för en PLS-modell. Den horisontella axeln visar korrelationen mellan de permuterade *Y*-vektorer och den ursprungliga *Y*-vektorn. Den vertikala axeln visar $R^2Y(\text{cum})$ - och Q^2 -värdet.

Kriterierna för giltighet är:

- Alla de blå Q^2 -värdena till vänster är lägre än punkterna till höger, eller
- Regressionslinjen för Q^2 -punkterna skär den vertikala axeln i eller under origo.

Skärningen med vertikala axeln är ett mått på överanpassningen.

3 MATERIAL OCH METOD

3.1 MÄTKAMPANJ

För att ta fram data till modelleringen genomfördes en mätkampanj i Linje 1 på Hammarby Sjöstadsvverk under tidsperioden 11-15 mars 2013. Mätningar skedde med ett varierande inflöde till försökslinjen och en provtagning av avloppsvattnet skedde varannan timme. Proverna analyserades sedan i laboratorium.

3.1.1 Förändringar i flöden

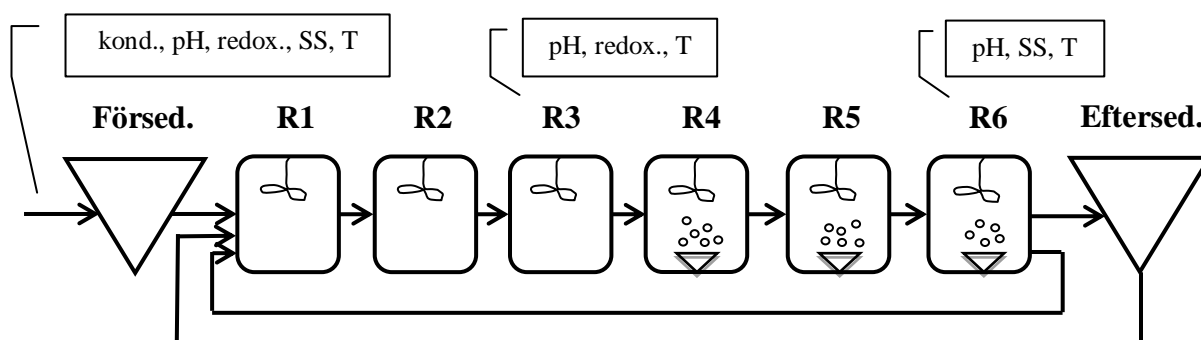
Flödesvariationer skapades manuellt genom att pumparna för det inkommande flödet, recirkulationsflödet och returslamflödet ändrades vid klockan 12, fyra gånger under tiden för mätkampanjen. Förhållandet mellan inflöde och recirkulationsflödet var alltid 1:4 och förhållandet mellan inflöde och returslamflödet var 1:1. Normalt driftläge för inflödet var 1,25 m³/tim och det inkommande inflödet ändrades under mätkampanjen kring detta värde (Tabell 2).

Tabell 2. Förändringarna som utfördes manuellt fyra gånger för de olika flödena på Linje 1 under mätkampanjen.

	Initialt	Ändring 1 (måndag)	Ändring 2 (tisdag)	Ändring 3 (onsdag)	Ändring 4 (torsdag)
Inflöde [m ³ /tim]	1,25	0,9	1,5	1,1	1,0
Recirkulationsflöde [m ³ /tim]	5,00	3,6	6,0	4,4	4,0
Returslamflöde [m ³ /tim]	1,25	0,9	1,5	1,1	1,0

3.1.2 Datainsamling och provtagning

Vid försökslinjen fanns ett antal on-linesensorer installerade där data samlades in (Figur 10) för totalt 20 olika lättmätbara variabler. På det inkommande vattnet till anläggningen registrerades konduktivitet, pH, redoxpotential, suspenderat material (SS) och temperatur. I R3 registrerades pH, redoxpotential och temperatur, medan det i R6 registrerades pH, SS och temperatur.



Figur 10. De olika on-linesensorernas placering under mätkampanjen. I inkommande vatten (in till försedimenteringsbassängen) mättes konduktivitet [$\mu\text{S}/\text{cm}$], pH, redoxpotential [mV], suspenderat material (SS) [mg/l] och temperatur (T) [$^{\circ}\text{C}$]. I R3 mättes pH, redoxpotential och temperatur medan R6 mätte pH, SS och temperatur. Dessutom registrerades syrehalten [mg/l] och ventilläget [%] till utblåset för R4-R6 samt inflöde [m³/tim], recirkulationsflöde [m³/tim] och returslamflöde [m³/tim].

I R4-R6 mättes även syrehalten och ventilläget på utblåset. Dessutom registrerades inkommande flöde, recirkulationsflöde och returslamflöde. Sensorerna samlade in data varje sekund och lagrade ett medelvärde av dessa var sjätte minut (Björk, pers. medd.). Innan försöket påbörjades rengjordes samtliga sensorer och de som var möjliga att kalibrera kalibrerades. De sensorer som inte kunde kalibreras utan professionell hjälp från leverantören inspekterades genom att vattnet vid den aktuella sensorn kontrollmättes antingen genom en analys eller med en bärbar enhet. Detta gällde för sensorerna för konduktivitet och SS på inkommande vatten samt sensorn för SS i R6.

Förutom mätvärdena från on-linesensorerna samlades vattenprover för de svårämätbara variablerna in från tre olika mätpunkter i processen: vid inkommande vatten, i R3 och i R6. Vattenprover togs med hjälp av tre olika provtagare (Figur 11), som programmerades att samla in en liten mängd vatten i en flaska var sjätte minut under 120 minuter, totalt cirka 300-400 ml från varje provtagningsplats.



Figur 11. En av tre provtagare som användes för insamling av vatten under mätkampanjen.

För att inte missa variationer över dygnet försköts provtagningen 30 minuter varje dag (Tabell 3). Då hann även provtagaren ställas om och nya flaskor ställas in. Under mätkampanjen samlades 53 vattenprover in från var och en av de tre provtagarna, totalt 159 prov analyserades i laboratorium.

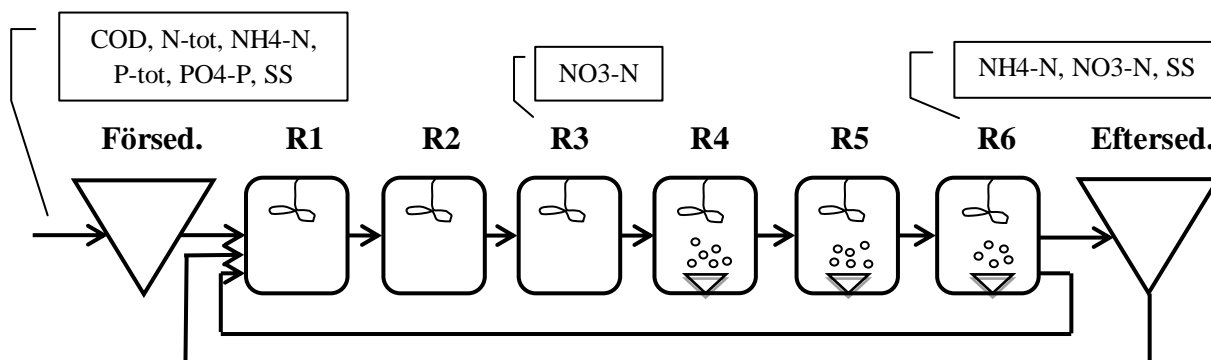
Tabell 3. Provtagnings-schemat för mätkampanjen.

mån, 11/3-2013	tis, 12/3-2013	ons, 13/3-2013	tors, 14/3-2013	fre, 15/3-2013
00:00	00:30	01:00	01:30	00:00
02:00	02:30	03:00	03:30	02:00
04:00	04:30	05:00	05:30	04:00
06:00	06:30	07:00	07:30	06:00
08:00	08:30	09:00	09:30	08:00
10:00	10:30	11:00	11:30	10:00
12:30	13:00	13:30	14:00	
14:30	15:00	15:30	16:00	
16:30	17:00	17:30	18:00	
18:30	19:00	19:30	20:00	
20:30	21:00	21:30	22:00	
22:30	23:00	23:30		

3.1.3 Analyser

Vattenprover tagna på det inkommande vattnet analyserades med avseende på kemisk syreförbrukning (COD), totalkväve (N-tot), ammoniumkväve (NH₄-N), totalfosfor (P-tot), fosfatfosfor (PO₄-P) och SS. I R3 analyserades nitratkväve (NO₃-N) och i R6 analyserades NH₄-N, NO₃-N och SS (Figur 12).

Analyserna inleddes med att vattnet filtrerades genom ett glasfiberfilter (Ø 55 mm). För det inkommande vattnet filtrerades 50 ml och för vattnet från de två bioreaktorerna filtrerades 25 ml vardera. SS bestämdes sedan genom att filtren vägdes innan filtrering och igen när filtren med filterrest hade torkats helt torra i 105°C. Differensen dividerat med den filtrerade volymen vatten ger mängden suspenderat material.



Figur 12. De olika variablerna som analyserades under mätkampanjen, alla i [mg/l]. I inkommande vatten analyserades kemisk syreförbrukning (COD), totalkväve (N-tot), ammoniumkväve (NH₄-N), totalfosfor (P-tot), fosfatfosfor (PO₄-P) och suspenderat material (SS). I R3 analyserades nitratkväve (NO₃-N) och i R6 analyserades NH₄-N, NO₃-N och SS.

För analys av de övriga variablerna användes färdiga kyvetttester från Hach Lange som analyserades fotometriskt. Kyvetttesterna är godkända för ackreditering och konstruerade så att kyvetterna är färdiga att använda med medföljande reagenser. En beteckning över vilket test som användes till vilken analys finns beskrivet i Bilaga A.

För att data från on-linesensorer och analysresultat skulle stämma överens i tidsupplösning medelvärdesbildades data från on-linesensorerna i Matlab så att en upplösning på 120 minuter erhöles.

3.2 TIDSFÖRSKJUTNING

För att kunna modellera data som samlats in, behövde det organiseras i en tidsserie där alla mätvärden vid samma tidpunkt tagits på "samma" vatten. Eftersom data samlats in på olika ställen i processen så kunde till exempel mätningar som tagits på onsdagen vid midnatt på inkommande vatten, inte placeras tillsammans med mätningar gjorda vid samma tid i någon av bioreaktorerna eftersom vattnet inte hunnit färdas dit ännu. Dessa mätvärden behövde tidsförskjutas så att de hamnade tillsammans med data som matchade dem. Uppskattningar på hur lång tid det tog för vattnet att passera genom olika delar av processen gjordes genom att använda beräknade värden på omsättningstiderna i försedimenteringsbassängen och de olika bioreaktorerna. Tankarna antogs vara fyllda till 90 % under mätkampanjen. Därefter beräknades hur lång tid vattnet tog på sig att passera mellan olika delar i processen genom

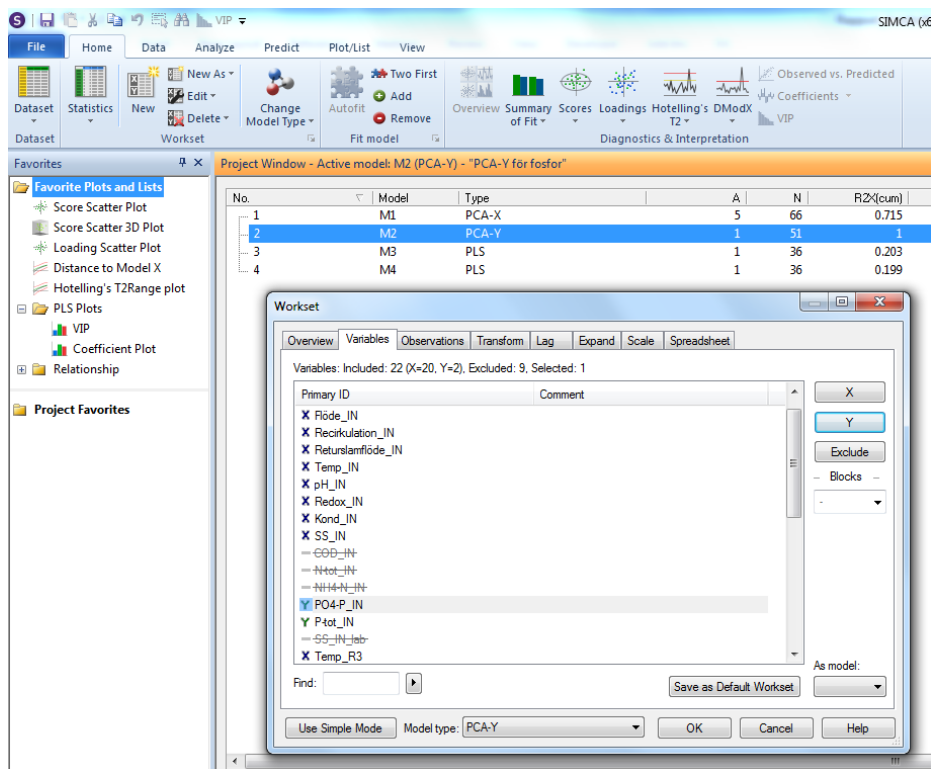
$$T = \frac{V}{Q} \quad (8)$$

där T är omsättningstiden [tim], V är den summerade volymen på bassängerna fram till mätpunkten [m^3] och Q är det aktuella vattenflödet [m^3/tim]. När omsättningstiderna var kända genomfördes tidsförskjutningar av data.

3.3 MODELLERING I SIMCA

När data har tidsförskjutits och uppenbart felaktiga värden från analysresultaten har tagits bort, importerades värdena där X - och Y -variabler definierats till SIMCA. Data innehöll 20 X -variabler från on-linesensorerna och 10 Y -variabler från analysresultaten i 56 observationer (tidpunkter). På grund av tidsförskjutningen fanns det luckor i data; 5,3 % av platserna saknade mätvärden.

Innan PLS-modellerna skapades togs först ett antal PCA-modeller fram. En PCA- X -modell för att ge en bättre samlad överblick över data. Sedan ett antal PCA- Y -modeller, en för varje enskild Y -variabel som modellerades genom att exkludera de övriga Y -variablerna från modellen (Figur 13). För inkommande vatten skapades två modeller för fosfor (en för totalfosfor och en för fosfatfosfor), två modeller för kväve (en för totalkväve och en för ammoniumkväve) och en modell för COD. För bioreaktorerna skapades en modell för nitratkväve. Ingen modell för ammoniumkväve togs fram på grund av att för få observationer fanns att tillgå. Observationer som saknade mätvärden för den aktuella Y -variabeln exkluderades ur modellen. Ett antal observationer togs slumpmässigt bort från PCA- Y inför modelleringsprocessen, för att senare användas till valideringen av de slutgiltiga PLS-modellerna.



Figur 13. Gränssnittet i SIMCA där variabler kunde definieras som X eller Y eller exkluderas.

Ur PCA-Y-modellen skapades en första PLS-modell. Denna ursprungliga PLS-modell modifierades till nya modeller iterativt genom att bland annat testa olika transformationer och expansioner på X-variablerna för att erhålla ett högre Q^2 -värde. Q^2 -värdet beskriver hur stor andel av observationerna i Y som kan predikteras av modellen. Processen innefattade även att exkludera de X-variabler som saknade korrelation med den aktuella Y-variabeln. Det gjordes genom att studera PLS-modellens *VIP Plot* och utesluta den variabel som bidrog minst till modellen. För varje ny PLS-modell som skapades noterades hur många PLS-komponenter modellen byggde på. Modelleringsarbetets syfte var att finna modeller med så få variabler som möjligt med ett högt Q^2 -värde.

När en bra modell arbetats fram testades den genom att prediktera de observationer som inte var inblandade när modellen skapades. En *Y PS Plot* togs fram och R^2 -värdet för varje Y-variabel noterades. Till sist togs en *Permutation Plot* fram för att undersöka modellens giltighet. Om Q^2 -värdets regressionslinje skar den positiva vertikala axeln förkastades modellen. Egenskaperna hos ett antal modeller jämfördes därefter med varandra för att ta fram den modell som bäst förklarade den aktuella Y-variabeln. För den valda modellen togs modellekvationen för PLS-regressionen fram. Ekvationen baseras på koefficienter där X-variablerna skalats och centrerats till *Unit Variance* medan Y-variabeln enbart skalats.

4 RESULTAT OCH OBSERVATIONER

4.1 MÄTKAMPANJ

Insamlad data från on-linesensorerna medelvärdesbildades och presenteras i Bilaga B. Ett medelvärde för varje variabel beräknades (Tabell 4).

Tabell 4. Medelvärden för varje variabel från on-linesensorerna.

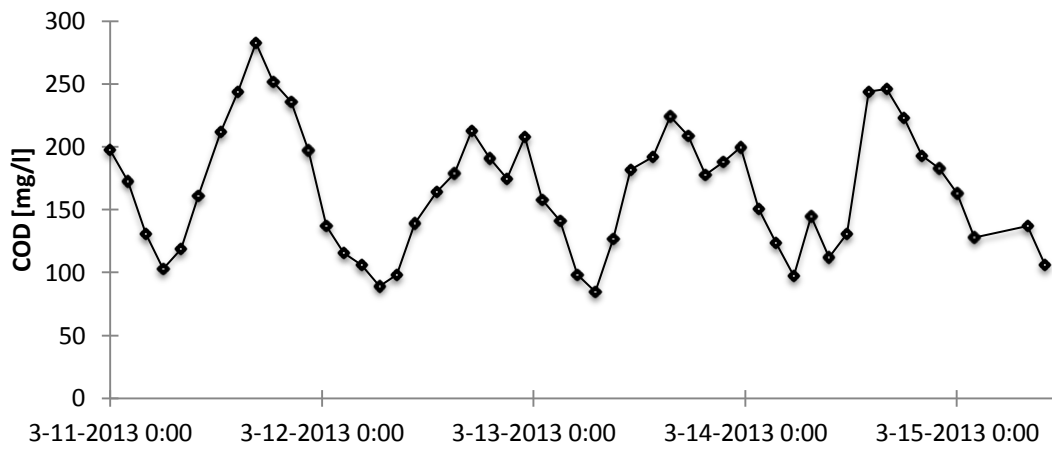
Variabel	Medelvärde
Temperatur_IN [°C]	15,6
pH_IN	7,3
Redoxpotential_IN [mV]	-289,1
Konduktivitet_IN [μS/cm]	535,7
SS_IN [mg/l]	316,7
Temperatur_R3 [°C]	15,7
pH_R3	6,6
Redoxpotential_R3 [mV]	-241,4
Temperatur_R6 [°C]	15,3
pH_R6	6,4
SS_R6 [mg/l]	3 618,4
O2_R4 [mg/l]	2,0
O2_R5 [mg/l]	2,0
O2_R6 [mg/l]	1,0
Ventil_R4 [% , ventilläge]	69,2
Ventil_R5 [% , ventilläge]	35,0
Ventil_R6 [% , ventilläge]	15,7
Inflöde [m ³ /tim]	1,2
Recirkulationsflöde [m ³ /tim]	4,7
Returslamflöde [m ³ /tim]	1,1

Kraftigt avvikande värden för analysresultaten hos de insamlade vattenproverna togs bort och medelvärden för varje variabel beräknades sedan (Tabell 5), se Bilaga C för samtliga mätvärden.

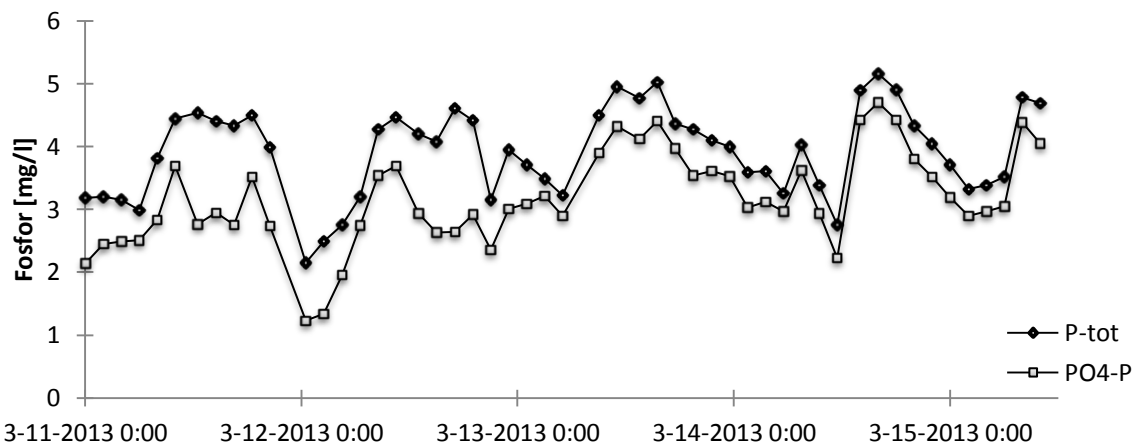
Tabell 5. Medelvärden för varje variabel från vattenproverna.

Variabel	Medelvärde [mg/l]
COD_IN	166,5
N-tot_IN	35,9
NH4-N_IN	26,1
P-tot_IN	3,9
PO4-P_IN	3,2
SS_IN	198,6
NO3-N_R3	0,3
NH4-N_R6	1,0
NO3-N_R6	1,3
SS_R6	2767,5

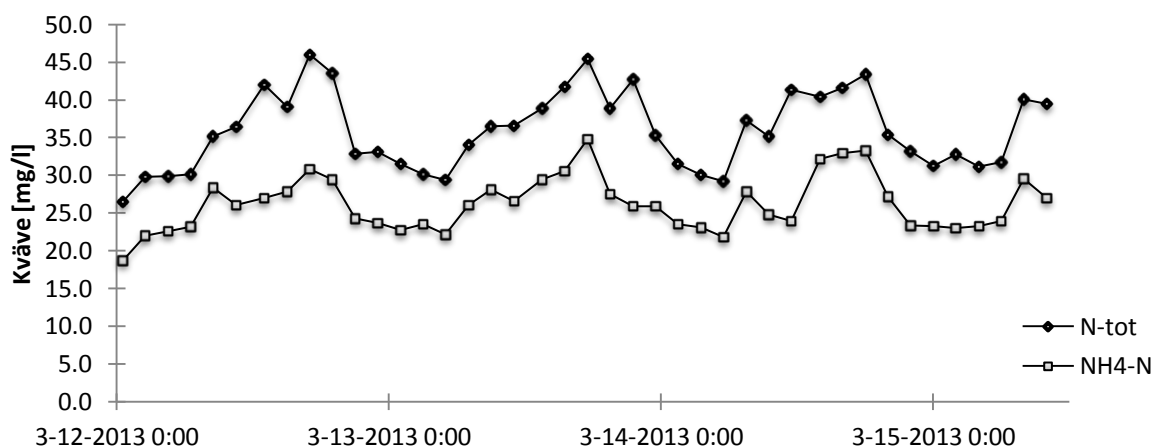
För inkommande vatten plottades variablernas förändringar under mätperioden. Både för COD (Figur 14), fosfor (Figur 15) och kväve (Figur 16) kan dygnsvariationer urskiljas. För inkommande kväve har den första mätdagen tagits bort på grund av mätfel.



Figur 14. COD [mg/l] i inkommande vatten under mätkampanjen.



Figur 15. Fosfor [mg/l] i inkommande vatten under mätkampanjen. Fyllda punkter visar totalfosfor och ihåliga punkter visar fosfatfosfor.



Figur 16. Kväve [mg/l] i inkommande vatten under mätkampanjen. Fyllda punkter visar totalkväve och ihåliga punkter visar ammoniumkväve.

4.2 TIDSFÖRSKJUTNING

Omsättningstiden i försedimenteringsbassängen och en bioreaktor beräknades med ekv. (8) för varje inflöde (Tabell 6). För bioreaktorn inkluderades även tillförseln från recirkulationsflödet och returslamsflödet med i beräkningarna.

Tabell 6. Omsättningstiden i försedimenteringsbassängen och en bioreaktor för varje inflöde.

Inflöde [m^3/tim]	$T_{\text{Försed.}}$ [tim]	$T_{\text{Bioreaktor}}$ [tim]
1,25	1,82	0,605
0,90	2,53	0,840
1,50	1,52	0,504
1,10	2,07	0,687
1,00	2,28	0,756

Omsättningstiden för varje bioreaktor med on-linesensorer beräknades genom att summera omsättningstiderna för försedimenteringsbassängen med det aktuella antalet bioreaktorer (Tabell 7).

Tabell 7. Omsättningstid för varje bioreaktor med on-linesensorer.

Inflöde [m^3/tim]	T_{R3} [tim]	T_{R4} [tim]	T_{R5} [tim]	T_{R6} [tim]
1,25	3,64	4,24	4,85	5,45
0,90	5,05	5,89	6,73	7,57
1,50	3,03	3,53	4,04	4,54
1,10	4,13	4,82	5,51	6,19
1,00	4,55	5,30	6,06	6,81

Tidsförskjutningen förenklades sedan genom att ta ett medelvärde för varje bioreaktor (Tabell 8). Tidsupplösningen på 120 minuter ansågs så hög att de förhållandevis små tidsskillnaderna kunde försummas. Medelvärdet avrundades därefter till jämna tvåtimmarsintervaller som användes vid tidsförskjutningen av data.

Tabell 8. Medelvärde för omsättningstiden i varje bioreaktor, samt värdet på tidsförskjutningen som användes till att förskjuta data.

	T_{R3} [tim]	T_{R4} [tim]	T_{R5} [tim]	T_{R6} [tim]
Medelvärde	4,08	4,76	5,44	6,11
Tidsförskjutning	4,00	4,00	6,00	6,00

Mätvärden tagna i R3 och R4 försköts 4 timmar från tiden för det inkommande vattnet och mätvärden från R5 och R6 i 6 timmar.

4.3 MODELLERING I SIMCA

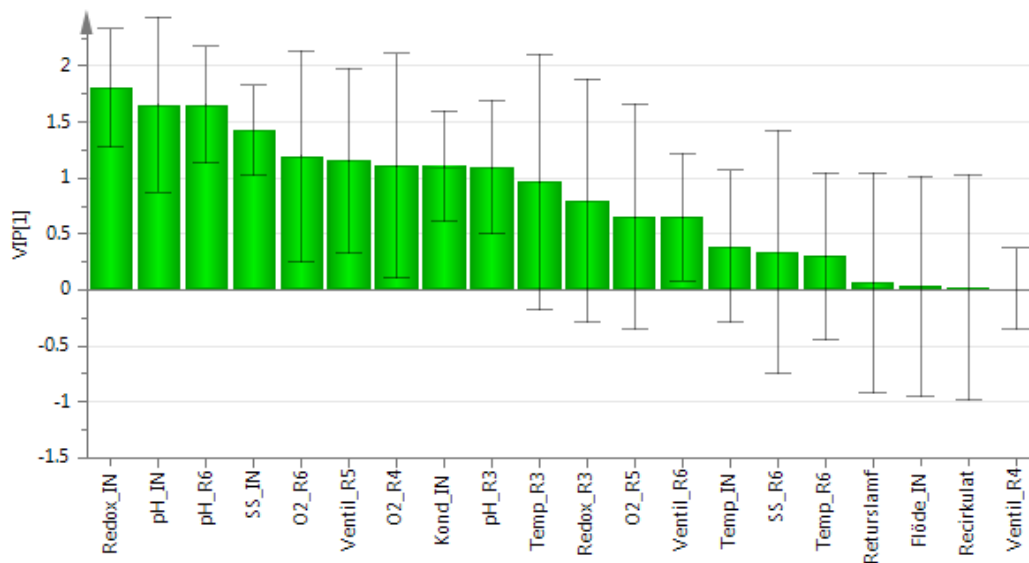
En översikt av data erhöles med hjälp av en PCA-X-modell. Med *Autofit* behövdes fyra principalkomponenter för PCA-X-modellen. Modellen beskrev 70,2 % av variationen i data ($R^2X(\text{cum})$). För att upptäcka eventuella avvikande observationer i data analyserades alla *Score Plots* utan att några kunde identifieras. *Hottelling's T2Range Plot* visade att alla observationer i det fyrdimensionella projektionsplanet befann sig under den gula gränsen för 0,05 signifikansnivå. I en *DModX Plot* fanns ett fåtal observationer som stack något över den kritiska gränsen D_{crit} , men inte så pass mycket att det ansågs nödvändigt att exkludera dem ur modellen. Således upptäcktes inga avvikande observationer i PCA-X-modellen och samtliga observationer användes därmed till PLS-modelleringen. Olika PLS-modeller med inkommande fosfor, inkommande kväve, inkommande COD och nitrathalten i bioreaktoerna som *Y*-variabler skapades för att undersöka om det var möjligt att förklara någon av de svårsmätbara variablerna i avloppsvattnet baserat på de mer lättsmätbara variablerna.

4.3.1 Inkommande fosfor

Två PLS-modeller, en med totalfosfor och en med fosfatfosfor som *Y*-variabler togs fram i SIMCA. Båda modellerna hade 51 observationer vardera. Åtta av dessa sparades undan för validering.

Totalfosfor

En *VIP Plot* togs fram som beskrev hur betydelsefull varje *X*-variabel var för PLS-modellen (Figur 17) där variablerna längst till höger inte är betydelsefulla för modellen.



Figur 17. En *VIP Plot* över de 20 *X*-variablerna för inkommande totalfosfor med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till 0,301. Ett trettiotal modeller skapades och analyserades, varav tre valdes ut för validering (Tabell 9). Ju fler *X*-variabler som exkluderas desto mer minskade Q^2 -värdet, vilket inte är önskvärt.

Tabell 9. Utvalda PLS-modeller för totalfosforhalten i inkommande vatten.

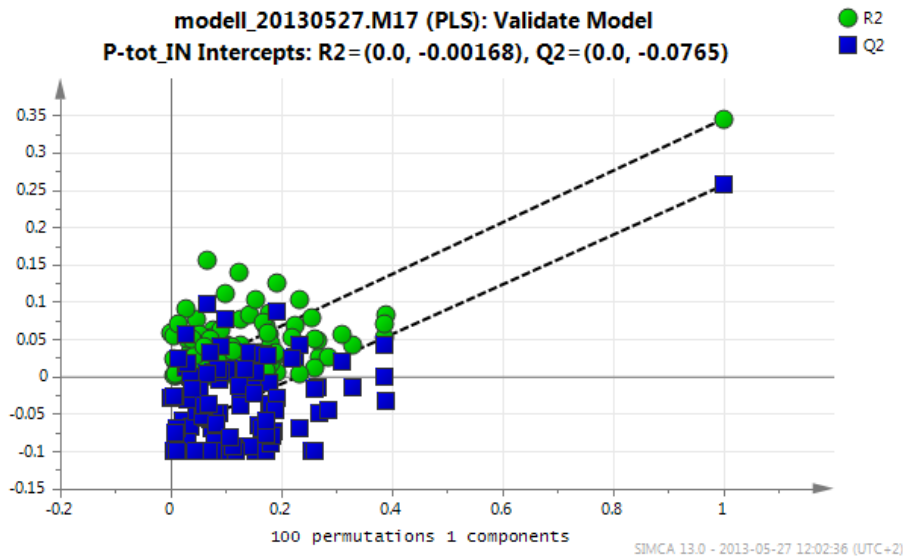
Modell nr.	Antal <i>X</i> -variabler	Antal PLS-komponenter	$R^2X(\text{cum})$	$R^2Y(\text{cum})$	Q^2
1	4	1	0,661	0,384	0,282
2	3	1	0,758	0,361	0,268
3	2	1	0,817	0,345	0,256

Modellernas prediktionsförmåga testades därefter på valideringsdata. En *Y PS Plot* togs fram och R^2 -värdet för varje *Y*-variabel noterades (Tabell 10).

Tabell 10. Prediktionsförmåga för totalfosforhalten hos de utvalda PLS-modellerna.

Modell nr.	R^2 -värde
1	0,69
2	0,47
3	0,51

Modellerna undersöktes vidare med *Permutations Plot* med 100 permutationer (Figur 18, visar modell 3). Samtliga visade indikationer på att vara användbara då skärningen för Q^2 skar de negativa vertikala axlarna. Modellekvationen valdes att tas fram på modell 3.



Figur 18. *Permutations Plot* för totalfosfor med modell 3.

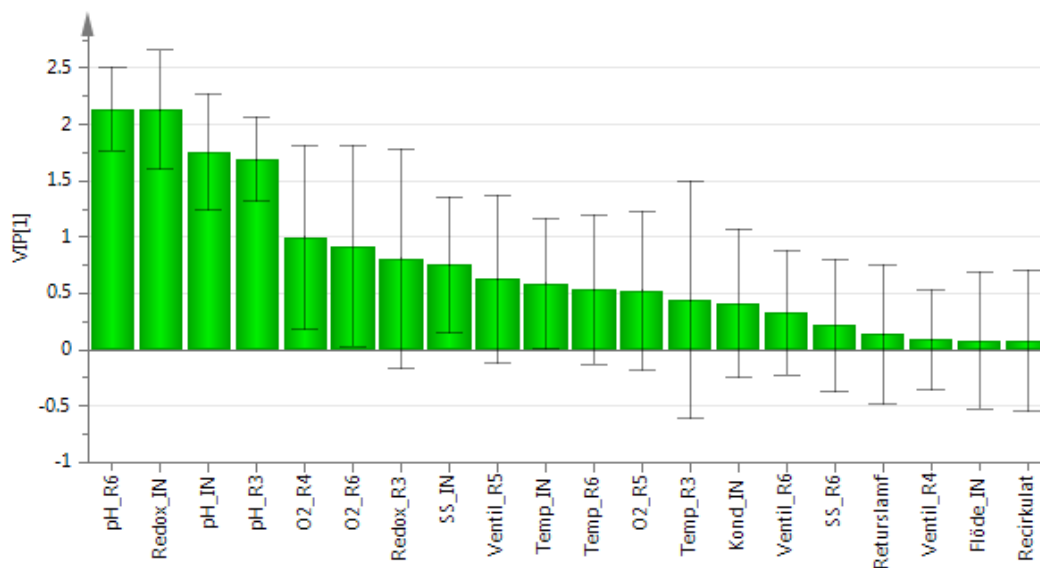
Modellekvationen för inkommande totalfosfor med modell 3 beskrivs enligt

$$P_{tot_IN} = 5,34 + 0,31 \cdot pH_{IN} - 0,34 \cdot Redox_{IN} \quad (9)$$

där halten av inkommande totalfosfor beror av pH och redoxpotentialen i inkommande avloppsvatten.

Fosfatfosfor

En *VIP Plot* användes för att se vilka *X*-variabler som var betydelsefulla för modellen (Figur 19).



Figur 19. En *VIP Plot* över de 20 *X*-variablerna för inkommande fosfatfosfor med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till 0,413. Ett tjugotal modeller skapades och analyserades, varav tre valdes ut för validering (Tabell 11). Ju fler X -variabler som exkluderas desto mer ökade Q^2 -värdet. Av de ursprungliga 20 X -variablerna verkade 2-4 X -variabler vara tillräckligt för att förklara inkommande fosfatfosforhalt till reningsverket.

Tabell 11. Utvalda PLS-modeller för fosfatfosforhalten i inkommande vatten.

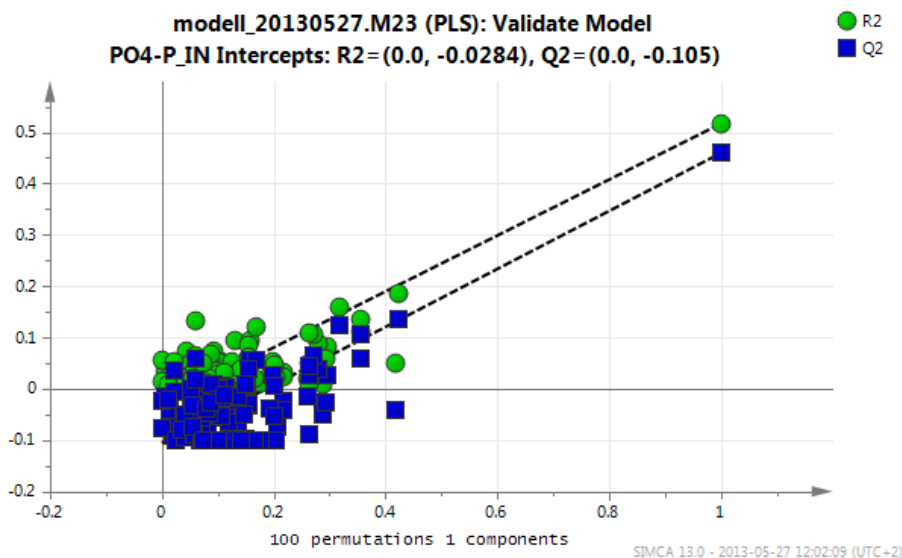
Modell nr.	Antal X -variabler	Antal PLS-komponenter	$R^2X(\text{cum})$	$R^2Y(\text{cum})$	Q^2
1	4	1	0,699	0,496	0,441
2	3	1	0,732	0,504	0,453
3	2	1	0,807	0,517	0,460

Modellernas prediktionsförmåga testades därefter på valideringsdata. En Y PS $Plot$ togs fram och R^2 -värdet för varje Y -variabel noterades (Tabell 12).

Tabell 12. Prediktionsförmåga för fosfatfosforhalten hos de utvalda PLS-modellerna.

Modell nr.	R^2 -värde
1	0,67
2	0,67
3	0,68

Modellerna undersöktes vidare med $Permutations$ $Plot$ med 100 permutationer (Figur 20, visar modell 3). Samtliga visade indikationer på att vara användbara då skärningen för Q^2 skar de negativa vertikala axlarna. Modellekvationen valdes att tas fram på modell 3.



Figur 20. $Permutations$ $Plot$ för fosfatfosfor med modell 3.

Modellekvationen för inkommande fosfatfosfor med modell 3 beskrivs enligt

$$PO4_IN = 3,95 - 0,40 \cdot Redox_IN - 0,40 \cdot pH_R6 \quad (10)$$

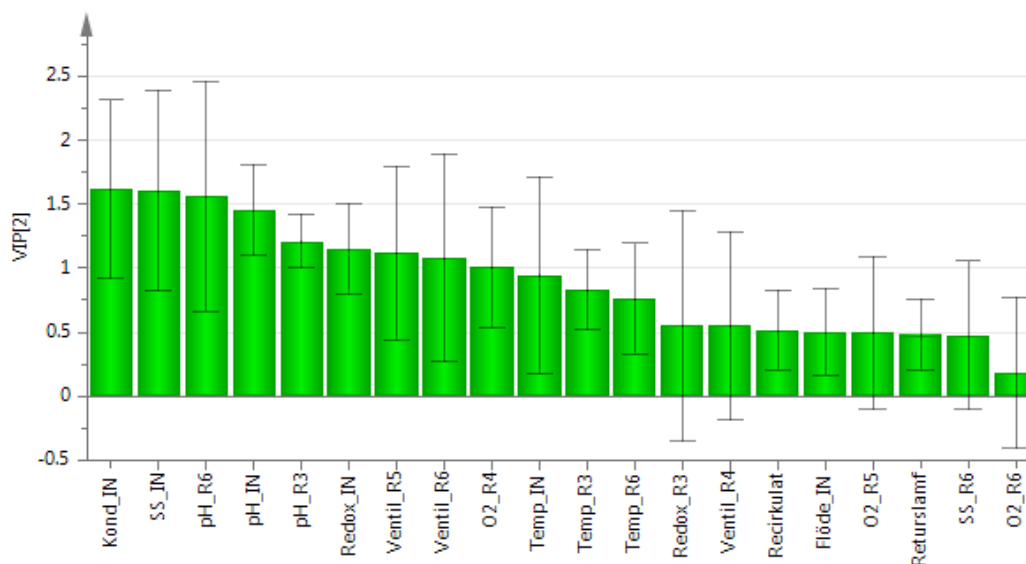
där halten av inkommande fosfatfosfor beror av redoxpotentialen i inkommande vatten samt pH i R6.

4.3.2 Inkommande kväve

Två PLS-modeller, en med totalkväve och en med ammoniumkväve som Y-variabler togs fram i SIMCA. Båda modellerna hade 41 observationer vardera. Sex av dessa sparades undan för validering innan modelleringen startade.

Totalkväve

En *VIP Plot* användes för att se vilka X-variabler som var betydelsefulla för modellen (Figur 21).



Figur 21. En *VIP Plot* över de 20 X-variablerna för inkommande totalkväve med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till 0,477. Ett tjugotal modeller skapades och analyserades, varav tre valdes ut för validering (Tabell 13). Från de ursprungliga 20 X-variablerna verkade 2-4 X-variabler vara tillräckligt för att förklara inkommande totalkvävehalt till reningsverket.

Tabell 13. Utvalda PLS-modeller för totalkvävehalten i inkommande vatten.

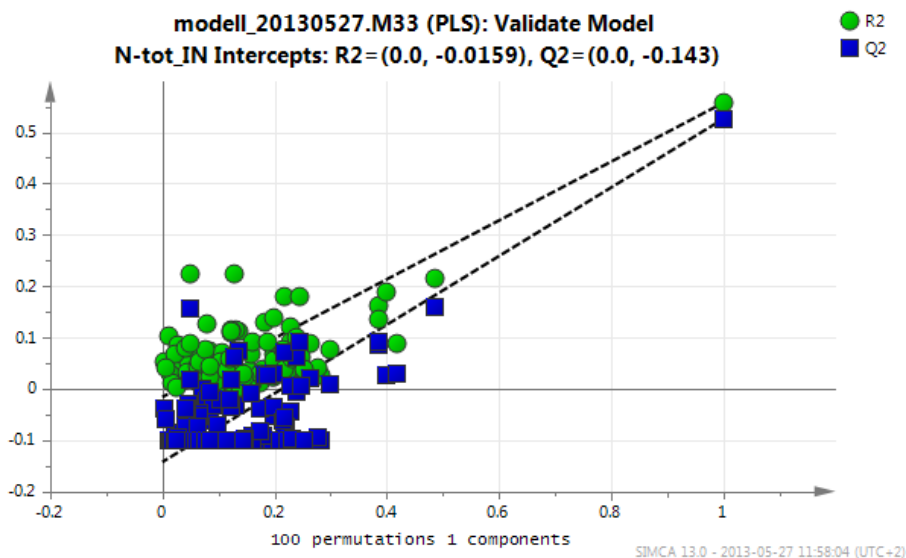
Modell nr.	Antal X-variabler	Antal PLS-komponenter	$R^2X(\text{cum})$	$R^2Y(\text{cum})$	Q^2
1	4	1	0,607	0,566	0,527
2	3	1	0,639	0,558	0,526
3	2	1	0,845	0,440	0,406

Modellernas prediktionsförmåga testades därefter på valideringsdata. En *Y PS Plot* togs fram och R^2 -värdet för varje *Y*-variabel noterades (Tabell 14).

Tabell 14. Prediktionsförmåga för totalkvävehalten hos de utvalda PLS-modellerna.

Modell nr.	R^2 -värde
1	0,90
2	0,86
3	0,87

Modellerna undersöktes vidare med *Permutations Plot* med 100 permutationer (Figur 22, visar modell 2). Samtliga visade indikationer på att vara användbara då skärningen för Q^2 skar de negativa vertikala axlarna. Modellekvationen valdes att tas fram på modell 2.



Figur 22. *Permutations Plot* för totalkväve med modell 2.

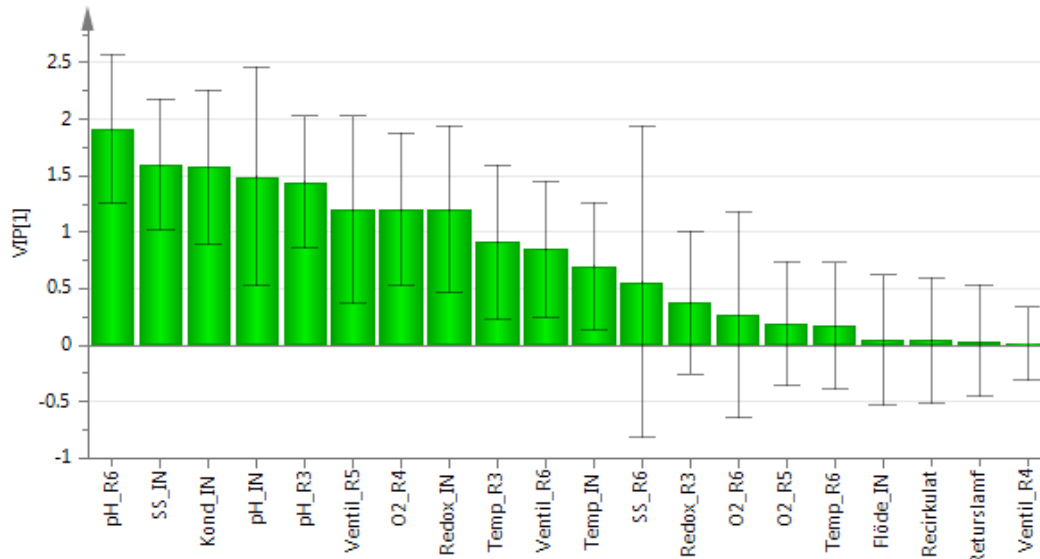
Modellekvationen för inkommande totalkväve med modell 2 beskrivs enligt

$$N_{tot_IN} = 6,82 + 0,32 \cdot Kond_IN + 0,32 \cdot SS_IN - 0,30 \cdot pH_R6 \quad (11)$$

där halten av inkommande totalkväve beror på konduktiviteten och SS i inkommande vatten samt pH i R6.

Ammoniumkväve

En *VIP Plot* användes för att se vilka *X*-variabler som var betydelsefulla för modellen (Figur 23).



Figur 23. En *VIP Plot* över de 20 *X*-variablerna för inkommande ammoniumkväve med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till 0,381. Ett tjugotal modeller skapades och analyserades, varav tre valdes ut för validering (Tabell 15). Från de ursprungliga 20 *X*-variablerna verkade 2-4 *X*-variabler vara tillräckligt för att förklara inkommande totalkvävehalt till reningsverket.

Tabell 15. Utvalda PLS-modeller för ammoniumkvävehalten i inkommande vatten.

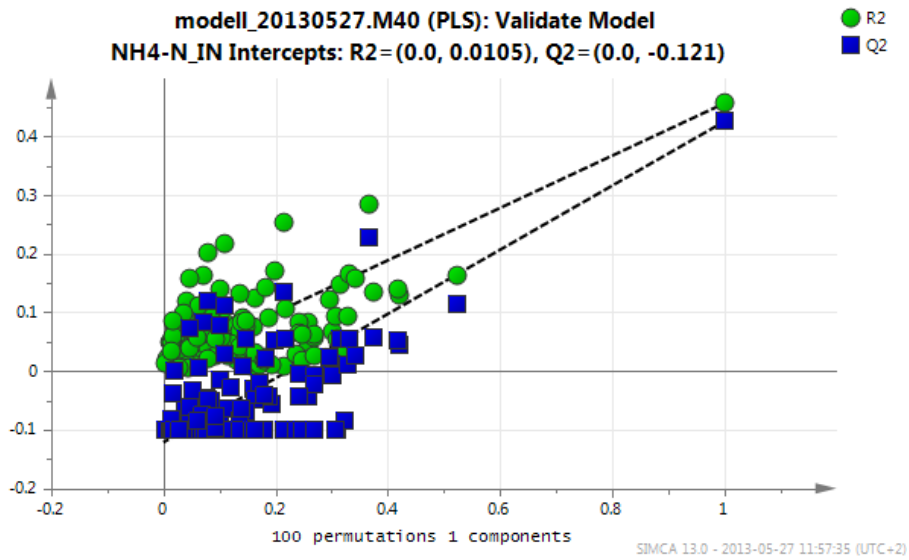
Modell nr.	Antal <i>X</i> -variabler	Antal PLS-komponenter	$R^2X(\text{cum})$	$R^2Y(\text{cum})$	Q^2
1	4	1	0,575	0,464	0,430
2	3	1	0,632	0,457	0,426
3	2	1	0,692	0,438	0,408

Modellernas prediktionsförmåga testades därefter på valideringsdata. En *Y PS Plot* togs fram och R^2 -värdet för varje *Y*-variabel noterades (Tabell 16).

Tabell 16. Prediktionsförmåga för ammoniumkvävehalten hos de utvalda PLS-modellerna.

Modell nr.	R^2 -värde
1	0,90
2	0,86
3	0,87

Modellerna undersöktes vidare med *Permutations Plot* med 100 permutationer (Figur 24, visar modell 2). Samtliga visade indikationer på att vara användbara då skärningen för Q^2 skar de negativa vertikala axlarna. Modellekvationen valdes att tas fram på modell 2.



Figur 24. *Permutations Plot* för ammoniumkväve med modell 2.

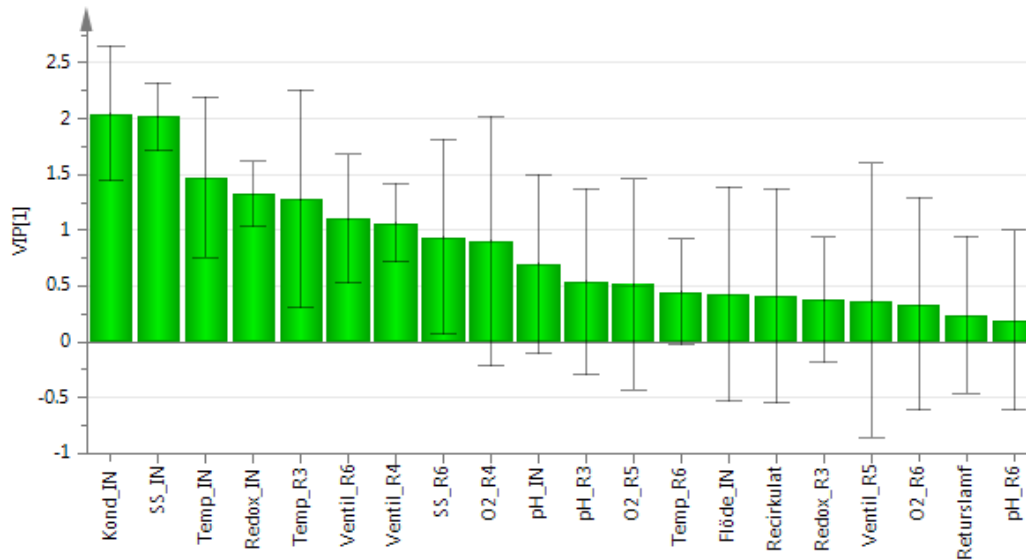
Modellekvationen för inkommande ammoniumkväve med modell 2 beskrivs enligt

$$NH4_IN = 6,98 + 0,27 \cdot Kond_IN + 0,27 \cdot SS_IN - 0,32 \cdot pH_R6 \quad (12)$$

där halten av inkommande ammoniumkväve beror på konduktiviteten och SS i inkommande vatten samt pH i R6.

4.3.3 Inkommande COD

En PLS-modell med kemisk syreförbrukning (COD) som Y -variabel och 51 observationer togs fram i SIMCA. Åtta observationer sparades undan för validering innan modelleringen startade. En *VIP Plot* användes för att se vilka X -variabler som var betydelsefulla för modellen (Figur 25).



Figur 25. En *VIP Plot* över de 20 X -variablerna för inkommande COD med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till 0,499. Ett tiotal modeller skapades och analyserades, varav tre valdes ut för validering (Tabell 17). Från de ursprungliga 20 X -variablerna verkade 2-4 X -variabler vara tillräckligt för att förklara inkommande COD till reningsverket.

Tabell 17. Utvalda PLS-modeller för COD-halten i inkommande vatten.

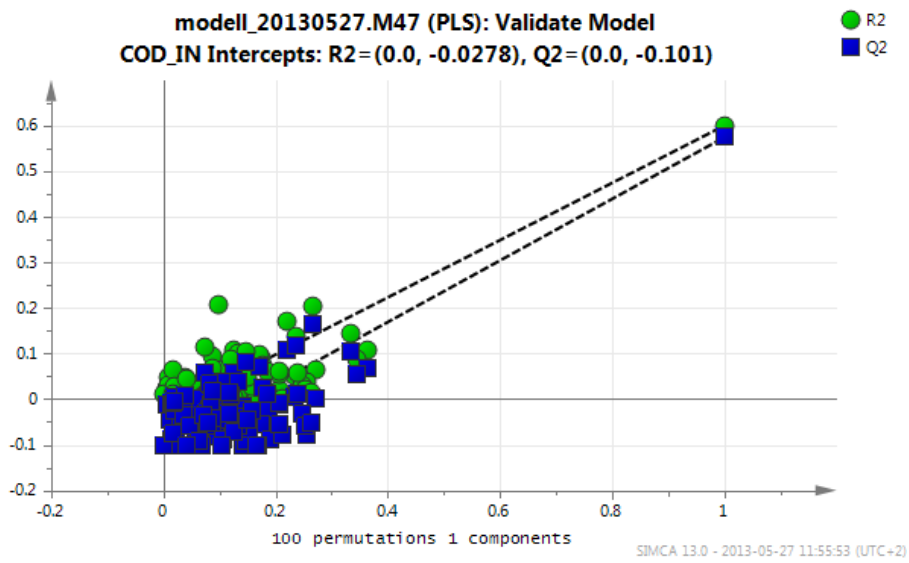
Modell nr.	Antal X -variabler	Antal PLS-komponenter	R^2X (cum)	R^2Y (cum)	Q^2
1	4	1	0,571	0,665	0,645
2	3	1	0,649	0,656	0,647
3	2	1	0,854	0,600	0,575

Modellernas prediktionsförmåga testades därefter på valideringsdata. En *Y PS Plot* togs fram och R^2 -värdet för varje Y -variabel noterades (Tabell 18).

Tabell 18. Prediktionsförmåga för COD-halten hos de utvalda PLS-modellerna.

Modell nr.	R^2 -värde
1	0,18
2	0,25
3	0,32

Modellerna undersöktes vidare med *Permutations Plot* med 100 permutationer (Figur 26, visar modell 3). Samtliga visade indikationer på att vara användbara då skärningen för Q^2 skar de negativa vertikala axlarna. Modellekvationen valdes att tas fram på modell 3.



Figur 26. *Permutations Plot* för COD med modell 3.

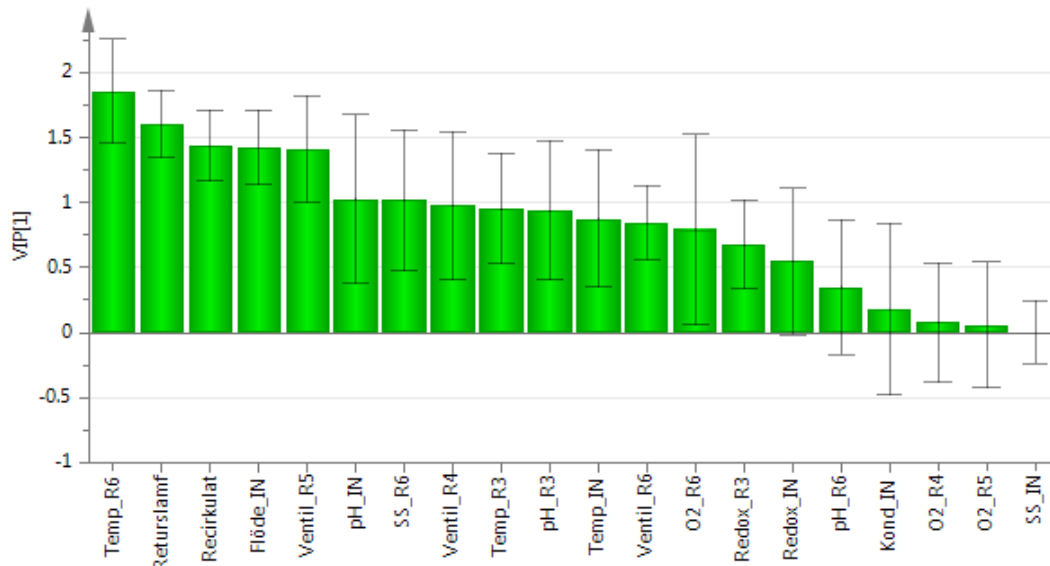
Modellekvationen för inkommande COD med modell 3 beskrivs enligt

$$COD_{IN} = 3,14 + 0,42 \cdot Kond_{IN} + 0,42 \cdot SS_{IN} \quad (13)$$

där halten av inkommande COD beror på konduktiviteten och SS i inkommande vatten.

4.3.4 Nitrat i bioreaktorerna

En PLS-modell med nitrathalterna i R3 och R6 som Y -variabler samt 54 observationer togs fram i SIMCA. Nio observationer sparades undan till validering. En VIP Plot användes för att se vilka X -variabler som var betydelsefulla för modellen (Figur 27).



Figur 27. En VIP Plot över de 20 X -variablerna för nitrat i bioreaktorerna med 95 % konfidensintervall. Ett värde över 1,0 indikerade att variabeln var betydelsefull för modellen och ett värde under 0,5 indikerade att variabeln var oanvändbar för modellen.

Från början var Q^2 -värdet beräknat till -0,1. Ingen PLS-modell kunde tas fram där Q^2 -värdet var nära att överstiga 0,2, som är det lägsta värde som erhållits under tidigare analyser, varför modelleringen avbröts.

5 DISKUSSION

Projektet har kortfattat inneburit att genom egna provtagningar samla in data som använts till framtagandet av multivariata modeller på avloppsvatten. Resultatet visar att det är möjligt att utifrån lättmätbara variabler utvinna information om de mer svårsmätbara variablerna i avloppsvattnet. För fortsatta studier och vidareutvecklingar av projektet, skulle mer data behöva samlas in. Detta för att få fler observationer till modelleringen och möjligtvis skulle bättre modeller då kunna utvinnas. Nu erhöles cirka femtio observationer, önskvärt hade varit att ha minst det dubbla antalet för att eventuellt avvikande värden lätt ska kunna tas bort utan att riskera att gå miste om värdefull information. Även samlingsintervallet för provtagningarna skulle kunna kortas ner för att fånga mer av variationen i provtagningen och samtidigt ge mer data. Detta skulle även underlätta den manuella tidsförskjutningen av data då en upplösning på 120 minuter är förhållandevis lång tid och många avrundningar kan påverka resultaten negativt. För att skapa mer varians i data borde inflödet till reningsverket kunna variera helt fritt istället för att manuellt varieras och hållas på konstanta nivåer periodvis. Då hade försöket liknat ett verkligt reningsverk mer, vilka har en kraftigt växlande belastning under dygnet.

5.1 MÄTKAMPANJ

Under mätkampanjen beräknades medelvärden från mätdata som registrerats av on-linesensorerna. Inga tydliga avvikelser upptäcktes hos on-linesensorerna, utan värdena kan ses som rimliga (Tabell 4). Analysresultaten (Tabell 5) för de inkommande parametrarna då de avvikande värdena togs bort har jämförts med Bengtsson (2004) och ligger inom samma intervall och antas därför vara relevanta. För bioreaktorerna jämfördes värden från den tidigare mätkampanjen och även de ligger inom samma intervall.

I projektet har provtagningen bestått av samlingsprov över två timmar istället för enstaka stickprov. Det negativa med samlingsprov är att man förlorar information om lokala toppar och dalar medan stickprov kräver fler analyser eftersom man förlorar information mellan provtagningarna. Syftet med den här mätkampanjen var att fånga variationer över dygnet, och eftersom den variationen ansågs vara ganska långsam valdes samlingsprov istället för stickprov. Men om syftet varit att fånga mer snabba, frekventa variationer hade täta stickprov varit att föredra.

Under insamlingen av data fanns ett antal felkällor som kan ha påverkat resultaten i olika grad.

- Första morgonen under mätkampanjen var fortfarande efterfällning av fosfor aktiv, vilken egentligen skulle vara avstängd under hela försöket. En omgång vatten från backspolning av sandfiltret hamnade i Linje 1 innan utloppet hann ändras och efterfällningen stängdes av. Detta kan till viss del ha påverkat resultaten men felet bedöms inte påvisbart eftersom det enbart handlade om ett enstaka tillfälle.

- Eftersom pumparna för flödena ändrades manuellt tog det en stund (cirka 10-15 min) innan önskvärda flöden erhöles. Detta antas dock inte ha gett någon större effekt på resultaten.
- Provtagarna tog tid att ställa om, speciellt att tömma och rengöra alla provflaskor innan ny start av maskinen var möjlig. Ibland hann momentet inte slutföras under de 30 min som pausades varje dag. En gång tog det upp till 45-60 min innan den nya provtagningen startade, vilket kan ha förskjutit mätpunkterna något. Eftersom tidsupplösningen är 120 minuter borde detta inte märkas.
- Vid ett tillfälle blev det stopp mellan R5 och R6 vilket ledde till höga vattennivåer i bioreaktorerna. Till slut svämmade försedimenteringen över och en del vatten var nödvändigt att tas ut för att få ner nivåerna igen. Det kan ha påverkat resultatet något, men då problemet löstes snabbt så avtas det inte ge några betydande effekter på resultatet.
- I provtagaren för inkommande vatten blev det en gång stopp i provtagningsslangen, vilket hindrade vattnet från att hamna i provflaskan. Lyckligtvis hände det under dagtid och upptäcktes och löstes i tid för att få ett prov under den aktuella 120-minutersperioden. Mängden vatten blev dock inte lika stor i det provet.

5.2 TIDSFÖRSKJUTNING

Omsättningstiderna för bioreaktorerna (Tabell 7) är en grov uppskattning av tiden det tar för vattnet att passera genom processen. Den egentliga tiden kan skilja sig från omsättningstiden. Tidsförskjutningen av data förenklades därefter genom att ta ett medelvärde på omsättningstiden för varje bioreaktor (Tabell 8). Det medför att en del av informationen går förlorad, men eftersom tidssteget var 120 minuter ansågs det försumma variationerna i omsättningstiderna. Till exempel varierar omsättningstiden i R3 mellan tre och fem timmar och den resulterande tidsförskjutningen blev fyra timmar. Att R3 och R4 samt R5 och R6 fick samma tidsförskjutning är inte helt realistiskt men eftersom vattnet passerar en bioreaktor så pass snabbt (Tabell 6) ansågs även denna variation kunna försummas av den stora tidsupplösningen.

5.3 MODELLERING I SIMCA

5.3.1 Inkommande fosfor

I PLS-modellerna för inkommande fosfor visade *VIP Plots* (Figur 17, Figur 19) att det fanns flera *X*-variabler som verkade vara betydelsefulla för modellerna. När de oanvändbara *X*-variablerna exkluderats ur PLS-modellerna och 2-4 variabler kvarstod var Q^2 -värdet något högre för fosfatfosfor än för totalfosfor (Tabell 9, Tabell 11). Att Q^2 -värdet var så lågt (mindre än 0,5) för samtliga utvalda modeller tyder det på att prediktionen av *Y*-variablerna från modellerna inte är tillräckligt bra. Då modellerna testades på valideringsdata erhöles R^2 -värden kring 0,67, för fosfatfosfor, men något lägre värden för totalfosfor (Tabell 10, Tabell 12). Modell 3 ansågs vara bäst för både totalfosfor och fosfatfosfor då de var enklast med minst antal variabler valdes de ut till vidare analyser. *Permutatins Plot* (Figur 18, Figur 20)

visade för både totalfosfor och fosfatfosfor en skärning på negativa vertikala axeln vilket indikerar att modellerna skulle kunna vara användbara för att förklara Y med andra observationer och behövde därför inte förkastas. För att förklara totalfosfor använder PLS-modellen värden för pH och redoxpotential i inkommande vatten ekv. (9). Medan fosfatfosfor förklaras med värden för redoxpotentialen i inkommande vatten samt med pH i R6 ekv. (10). Modellen för totalfosfor skulle kunna användas för prediktion, medan modellen för fosfatfosfor innehåller en variabel i R6 och kan därför bara användas till att få en förståelse av fosfatfosforhalten i vattnet. När pH i R6 testas att bytas ut mot pH i inkommande vatten istället blev tyvärr modellen för fosfatfosfor tyvärr sämre. Sammanfattningsvis kan det konstateras att ingen av PLS-modellerna för inkommande fosfor kunde beskriva Y -variablerna felfritt. Modellen för totalfosfor var något sämre än den för fosfatfosfor, för att kunna avgöra om modellen är användbara behöver de testas på en större mängd valideringsdata.

5.3.2 Inkommande kväve

I PLS-modellerna för inkommande kväve visade *VIP Plots* (Figur 21, Figur 23) att det fanns några X -variabler som verkade vara betydelsefulla för modellerna. När X -variablerna som inte bidrog till modellerna tagits bort och 2-4 variabler kvarstod var Q^2 -värdet för totalkväve högre än för ammoniumkväve (Tabell 13, Tabell 15). För totalkväve har 2 av 3 utvalda modeller ett värde över 0,5 och indikeras därför att inneha en god prediktionsförmåga. Då modellerna testades på valideringsdata erhöles R^2 -värden kring 0,87, för både totalkväve och ammoniumkväve (Tabell 14, Tabell 16) vilket kan anses som bra. Modell 2 valdes för både totalkväve och ammoniumkväve för att Q^2 -värdet var högre än i modell 3. *Permutatins Plot* (Figur 22, Figur 24) visade för både totalkväve och ammoniumkväve en skärning på negativa vertikala axeln vilket indikerar att modellerna skulle kunna vara användbara för att förklara Y med andra observationer och behöver därför inte förkastas. För att förklara totalkväve och ammoniumkväve använder PLS-modellerna värden för konduktivitet och SS i inkommande vatten, samt pH i R6 ekv. (11), ekv. (12). Dessa modeller kan bara användas för att ge en förståelse för kvävehalten och kan ej användas vid prediktion då de beror av värden i R6. För totalkväve kunde pH-värdet i R6 tas bort, det är modell 3 i Tabell 13 och Tabell 14. Modell 3 för totalkväve skulle kunna användas för prediktion av inkommande totalkväve. Sammanfattningsvis kan det fastställas att PLS-modellerna för inkommande kväve kunde beskriva Y -variablerna ganska bra, modell 3 för totalkväve kan även användas för prediktion. Modell 2 för både totalkväve och ammoniumkväve var likvärda med nästan exakt samma modellekvation. Därför skulle en modell räcka för att förklara inkommande kvävehalt. Men för att vara helt säker på att modellerna är användbara skulle de behöva testas på en större mängd valideringsdata.

5.3.3 Inkommande COD

För inkommande COD visar PLS-modellens *VIP Plot* (Figur 25) att speciellt två X -variabler har stor betydelse för modellen; Kond_IN och SS_IN. När de X -variabler som inte var användbara för modellen uteslöts och 2-4 X -variabler kvarstod var Q^2 -värdet ungefär lika med 0,647 (Tabell 17). Då modellerna testades på valideringsdata erhöles R^2 -värden kring 0,25 (Tabell 18), vilket inte ses som tillräckligt tillfredsställande. Modell 3 valdes ut för vidare

analyser för att den hade minst antal variabler men ändå ett Q^2 -värde större än 0,5. *Permutations Plot* (Figur 26) visade en skärning med den negativa vertikala axeln vilket indikerar att modellen skulle kunna vara användbar för att förklara Y med andra observationer och behöver därför inte förkastas. För att förklara COD i med den utvalda PLS-modellen användes värden för konduktivitet och SS i inkommande vatten ekv. (13) och skulle kunna användas vid prediktion. Sammanfattningsvis kan det konstateras att PLS-modellen för COD inte presterade bra på valideringsdata, men innan den förkastas bör den testas på en större mängd valideringsdata.

5.3.4 Nitrat i bioreaktorerna

För nitrat i bioreaktorerna lyckades ingen bra PLS-modell tas fram.

5.3.5 Allmänt

En gemensam felkälla för alla PLS-modeller i denna rapport är valideringen, då antalet valideringspunkter var ganska få. Egentligen var ytterligare en mätkampanj planerad för att användas till enbart validering, men den kunde inte genomföras inom projektet. För att skapa en bra modell med multivariat analys krävs minst 20-30 mätvärden och ju fler värden som finns tillgängliga desto bättre blir oftast modellen. Till följd av detta kunde därför inte data från mätkampanjen delas på hälften och använda den ena halvan till att skapa modeller och den andra till validering utan ett fåtal punkter valdes istället slumpmässigt ut till validering.

Syftet med PLS-modellerna måste innan modelleringen börjar klargöras. Om tanken är att kunna använda modellen för att skapa förståelse över processen kan alla X -variabler användas till modelleringen. Men om syftet istället är att kunna använda modellen till styrning kan exempelvis inte detta ske med en Y -variabel för inkommande vatten vars PLS-modell innehåller variabler som mäts i exempelvis R6.

Det finns flera tidigare studier där svärmätbara variabler i avloppsvattnet har försökt förklarats med lättillgängliga on-linevärden. Jansson m.fl. (2002) har skapat mjukvarusensorer för fosfor på avloppsvattnet i Borlänge's reningsverk med flera olika multivariata analysmetoder och modellstrukturer. Både modeller för totalfosfor och fosfatfosfor togs fram. Variabler som kunde mätas on-line var bland annat vattenflöde, konduktivitet, pH, COD och SS. Modeller skapades genom att använda olika kombinationer av de lättmätbara variablerna, allt ifrån tre variabler till samtliga variabler. Alla beräkningar skedde med PLS-Toolbox i Matlab. Bäst resultat gav modeller som byggde på PLS istället för MLR (multiple linear regression) eller PCR (principal component regression), samt modeller med modellstrukturen FIR (finite impulse response) istället för statistiska modeller. Slutligen visar mjukvarusensorerna en bättre uppskattning för fosfatfosfor än totalfosfor. Den bästa statistiska modellen var en PLS med två komponenter som byggde på fem lättmätbara variabler. I Göteborg har Hedegård & Wik (2011) använt prediktionsmodeller med observatör för aktivslamprocessen i Ryaverket. Förhoppningen var att kunna använda resultatet i den etablerade modellen Activated Sludge Model NO.1 (ASM1) för att förbättra driften för reningsverk genom att erhålla en bättre övervakning och kunna använda nya styrstrategier. Hedegård & Wik studerade i fall halterna av S_S (readily biodegradable dissolved substrate),

X_S (slowly biodegradable substrate) och X_{BH} , X_{BA} (active biomass concentrations) som är svåra att mäta on-line skulle kunna uppskattas med hjälp av lättmätbara variabler så som syrehalt, vattenflöden, luftflöden och totalt suspenderat material. Genom simuleringar kunde slutsatsen att uppskattningarna för variablerna i aktivslamprocessen inte fungerade bra för system med bara en luftad tank dras. Men om minst två tankar i serie användes och syrehalten mättes i båda konvergerar prediktionerna. I Kina har Zhao m.fl. (2010) använt mjukvarusensorer för att uppskatta reningsgraden för utflödet hos avloppsvattnet i Shenyang. Genom 17 olika variabler i processen, exempelvis syrehalt, flöden, slamkoncentrationer, COD, SS, pH används för att ta reda på utgående värden av BOD, COD, SS, pH och NH_4 . De jämförde flera olika multivariata modeller som alla baserades på PLS. Resultaten visade att mer avancerade modeller exempelvis NNPLS (non-linear neutral network PLS) är bättre på att förutspå utflödet än en traditionell linjär PLS. Deras slutsatser var att linjära PLS-modeller ger en bra prediktionsförmåga för linjära och stabila modeller men är inte lika tillämpbara på olinjära och dynamiska system som de flesta avloppsreningsverksprocesser är.

De framtagna modellerna i det här projektet behöver bekräftas med hjälp av en större mängd valideringsdata. Om modellerna inte kan förklara det nya datasetet bör de förkastas. Eftersom halterna av de studerade ämnena i inkommande vatten varierar kraftigt kan det vara lämpligt att se över metoden för insamling av mätdata, i första hand genom att utvidga mätkampanjen. Ett annat alternativ kan vara att använda mer avancerade multivariata analysmetoder som kanske bättre lämpar sig för den komplicerade process som kommunala avloppsreningsverk är.

6 SLUTSATSER

Från den här studien kan följande slutsatser dras:

- Det är möjligt att förklara svårsmätbara variabler i avloppsvattnet i Hammarby Sjöstadsverk med modeller skapade med multivariata analysmetoder som bygger på lättmätbara variabler.
- PLS-modeller som väl kan prediktera kvävehalt i inkommande vatten erhöles. Modellerna för inkommande fosforhalt och COD hade en sämre prediktionsförmåga. Ingen modell som acceptabelt kunde prediktera nitrathalten i bioreaktorerna lyckades tas fram.

7 REFERENSER

7.1 TRYCKTA REFERENSER

Aguado, D., Ferrer, A., Seco, A., Ferrer, J., (2006). *Comparison of different predictive models for nutrient estimation in a sequencing batch reactor for wastewater treatment*. Chemometrics and intelligent laboratory systems, 1, 75-81.

Andersson, M., Olsson, J., Röttorp, J., Ek, M., (2003). *Tillämpning av multivariata metoder för övervakning och optimering av skogsindustriella biologiska reningsanläggningar*. IVL Rapport B1539.

Baresel, C., Ek, M., Ekengren, Ö., Levlin, E., Bengtsson, L., Trela, J., Plaza, E., (2011). *Hammarby Sjöstadswerk – Uppstart av försöks- och demonstrationsanläggningen för framtidens kommunala VA-teknik*. IVL Rapport B1954.

Bengtsson, J., (2004). *Förbehandling av kommunalt avlopp före anaerob behandling*. Lunds Tekniska Högskola. Master's Thesis no 2003-03.

Björk, A., (2007). *Chemometric and signal processing methods for real time monitoring and modeling using acoustic sensors – applications in the pulp and paper industry*. Royal Institute of Technology, School of Chemical Science and Engineering, Department of Chemistry, ISSN 1654-1081.

Galicia, H. J., He, Q.P., Wang, J., (2011). *A reduced order soft sensor approach and its application to a continuous digester*. Journal of Process Control, 4, 489-500.

Hedegård, M., & Wik, T., (2011). *An online method for estimation of degradable substrate and biomass in an aerated activated sludge process*. Department of Signals and Systems, Chalmers University of Technology. Water Research, 45, 6308-6302.

Jackson, J. E., (1991). *A User's Guide To Principal Component*, John Wiley & Sons, INC., ISBN 0-471-62267-2.

Jansson, Å., Röttorp, J., Rahmberg M., (2002). *Development of a software sensor for phosphorus in municipal wastewater*. Journal of chemometrics, 16, 542-547.

Jolliffe, I. T., (2002). *Principal component analysis*, Springer, Secaucus.

Lumley, D., (2002). *On-line instrument confirmation: how can we check that our instruments are working?*. Water Science and Technology, 4-5, 469-476.

MKS Umetrics, (2012). *Users Guide to SIMCA, version 13*.

Nilsson, Å., Rahmberg, M., Nilsson, S., Röttorp, J., (2007). *Resurseffektiv avloppsvattenrening*. IVL Rapport B1760.

Shlens, J., (2005) *A Tutorial on Principal Component Analysis*. Systems Neurobiology Laboratory, Salk Institute for Biological Studies & Institute for Nonlinear Science, University of California.

Svenskt Vatten, (2010). *Publikation U2: Avloppsteknik 2 – Reningsprocessen*. ISSN: 1654-5117.

Wold, S., Sjöström, M., Eriksson, L., (2001). "PLS-regression: a basic tool of chemometrics". *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130.

Yu, H., Wu, Z. R., Bao, T. F., Zhang, L., (2010). "Multivariate analysis in dam monitoring data with PCA". *Sci China Tech Sci*, 53, 1088-1097.

Zhao, L-J., Xiao, H., Diao, X-K., Chai, T., (2010). *Nonlinear Dynamic Soft Sensor modeling of Wastewater Treatment Effluent Quality*. College Of Information Engineering, Institute of Chemical Technology. ISBN: 978-1-4244-5845-5.

7.2 ÖVRIGA REFERENSER

Baresel, Christian, (2013). muntligen, 2013-01-21.

Björk, Anders, (2013). e-post, 2013-04-04.

BILAGA A

HACH LANGE KYVETTESTER

Undersökt variabel	Beteckning
Ammoniumkväve	LCK 303, 304
Fosfor	LCK 348
Kemisk syreförbrukning (COD)	LCK 514, 614
Nitrat	LCK 339
Totalkväve	LCK 338

BILAGA B

DATA FRÅN ON-LINESENSORERNA UNDER MÄTKAMPANJEN, DEL 1/2

Tid	Temp_IN	pH_IN	Redox_IN	Kond_IN	SS_IN	Temp_R3	pH_R3	Redox_R3	Temp_R6	pH_R6	SS_R6
201303092200	16.166870	7.042181	-269.170250	525.171650	290.471150	16.497940	6.614885	-239.426850	15.929270	6.451554	3585.274500
201303100000	15.948970	6.989699	-263.737950	565.330000	343.679150	16.446300	6.641458	-254.942400	15.983255	6.465244	3594.958500
201303100200	15.217260	7.145319	-269.627950	580.972750	258.617950	16.382705	6.659954	-258.574200	15.973365	6.473464	3595.686000
201303100400	14.807710	7.209570	-274.163050	565.574050	206.728300	16.309820	6.684469	-264.763450	15.911830	6.470899	3593.264000
201303100600	14.268255	7.244241	-266.484650	553.891650	178.879250	16.224245	6.688925	-248.435900	15.778965	6.453323	3590.600500
201303100800	13.842100	7.146845	-272.789750	540.406450	217.169700	16.146240	6.671468	-252.134900	15.580190	6.418673	3580.103000
201303101000	14.352620	7.291524	-275.957550	532.247900	307.976500	16.104250	6.639893	-226.554200	15.337605	6.396843	3552.503000
201303101200	15.631210	7.388613	-290.081500	540.890700	377.461100	16.081315	6.616554	-224.521650	15.133985	6.393140	3551.477500
201303101400	16.419695	6.889817	-277.001300	493.273700	359.841850	16.083885	6.597653	-234.434000	15.039570	6.381156	3571.872500
201303101600	16.495375	6.732222	-270.073750	508.215550	340.228500	16.080100	6.568600	-246.311900	15.108640	6.379427	3581.057500
201303101800	16.341455	6.736678	-272.350150	515.094900	356.098200	16.065320	6.563066	-241.611800	15.222385	6.402233	3582.668000
201303102000	16.366715	6.740686	-273.564900	490.586000	353.103150	15.976555	6.578793	-260.594550	15.305440	6.419853	3577.170000
201303102200	16.388925	6.746875	-275.737850	471.908900	312.167750	15.891110	6.569415	-251.280300	15.339725	6.433444	3557.160500
201303110000	16.241955	6.738584	-268.926000	485.059650	280.265800	15.817745	6.566769	-261.851950	15.365515	6.459121	3564.722500
201303110200	15.592510	6.834461	-273.302450	502.174500	215.135200	15.734720	6.565977	-253.733900	15.383390	6.473606	3577.683500
201303110400	15.056370	6.877396	-264.848800	498.685600	175.778500	15.611075	6.559851	-255.833600	15.355615	6.474317	3581.544000
201303110600	14.331120	6.946445	-258.269050	489.062200	171.709250	15.504510	6.566220	-254.502950	15.258340	6.461867	3581.002000
201303110800	14.340155	7.269581	-277.855800	496.427300	300.586950	15.419420	6.542314	-240.397450	15.083440	6.441826	3576.159000
201303111000	15.478615	7.190179	-306.543150	514.999800	349.066700	15.384880	6.515519	-243.729850	14.875725	6.420281	3572.741000
201303111200	15.846660	7.033851	-300.708000	536.980400	380.667600	15.412230	6.520645	-244.218300	14.757945	6.398693	3583.109000
201303111400	15.756345	7.351116	-304.382450	577.000200	405.074150	15.487905	6.527728	-247.068650	14.778960	6.377005	3587.595500
201303111600	15.700805	7.280394	-290.997000	610.501400	405.717050	15.577870	6.520970	-253.257850	14.844085	6.365775	3618.386500
201303111800	15.653065	7.230953	-281.035800	634.692300	425.281350	15.642330	6.517452	-258.964800	14.926035	6.364942	3654.560000

201303112000	15.742430	7.394818	-293.530100	620.931650	396.219800	15.691160	6.497797	-252.250700	14.977895	6.370862	3673.160500
201303112200	16.095210	6.840154	-244.444100	580.511950	349.522200	15.661745	6.473504	-252.897700	15.012470	6.371574	3686.092500
201303120000	16.208515	6.688193	-200.436400	517.505400	279.696100	15.622305	6.483495	-255.082850	15.064325	6.376050	3692.644500
201303120200	16.038335	6.804290	-206.680400	501.200500	264.534350	15.576520	6.473259	-235.984400	15.104450	6.380038	3689.282500
201303120400	15.447490	7.036699	-201.486300	509.509550	205.580900	15.524150	6.478224	-244.059550	15.150765	6.391389	3684.055500
201303120600	14.722860	7.047116	-190.603450	500.069100	201.348850	15.473735	6.483312	-238.877500	15.176680	6.399793	3680.281000
201303120800	14.564660	7.355713	-231.009900	506.327500	350.962900	15.439065	6.488011	-223.941750	15.158085	6.392814	3675.626000
201303121000	15.803580	7.557643	-293.316550	513.049700	329.868450	15.469605	6.488540	-225.272350	15.102620	6.380138	3677.534000
201303121200	16.133420	7.666514	-301.843400	556.084900	366.311700	15.551015	6.506648	-235.251850	15.029815	6.363557	3680.452000
201303121400	15.805910	7.700329	-294.207600	581.919750	395.218900	15.607425	6.508235	-224.253000	15.112955	6.342417	3663.776000
201303121600	15.745970	7.544115	-289.336900	619.998750	411.389500	15.670525	6.513647	-218.021100	15.219900	6.331738	3614.341000
201303121800	15.736200	7.526760	-301.001150	621.419900	394.787450	15.728265	6.530981	-222.550050	15.301870	6.323557	3588.107500
201303122000	15.972055	7.461553	-301.635750	552.736900	366.588350	15.756340	6.560381	-236.527650	15.365660	6.330028	3583.451500
201303122200	16.483045	7.231462	-297.784450	480.847650	364.798050	15.786130	6.564205	-235.746300	15.434500	6.331003	3587.509500
201303130000	16.561050	7.177709	-289.715350	509.920450	306.438300	15.790025	6.553179	-236.515350	15.541965	6.352387	3583.664000
201303130200	16.187385	7.457503	-295.245200	516.983050	231.647750	15.786240	6.556821	-219.553200	15.669920	6.376395	3582.368000
201303130400	15.557605	7.488530	-292.718350	506.461750	199.200400	15.772475	6.563473	-220.432100	15.750925	6.384065	3578.865500
201303130600	14.846890	7.463934	-267.265850	484.708150	190.883150	15.748040	6.565182	-224.375250	15.752900	6.373040	3572.215000
201303130800	14.755460	7.562435	-276.384800	518.606600	289.087550	15.733265	6.542457	-210.049850	15.661575	6.356619	3569.963000
201303131000	15.844830	7.602324	-302.355950	491.926950	332.977350	15.742795	6.520401	-201.071250	15.494645	6.332733	3572.698500
201303131200	16.122815	7.407452	-304.492350	521.436150	361.355500	15.809445	6.516415	-209.872700	15.412460	6.319793	3583.878500
201303131400	15.769160	7.416548	-312.128000	576.663550	369.461150	15.841055	6.543352	-235.941750	15.456995	6.319529	3613.174000
201303131600	15.627305	7.238645	-316.089400	608.674400	406.766900	15.851925	6.567400	-239.371950	15.508985	6.283822	3628.698500
201303131800	15.579575	7.389954	-322.870600	630.830850	480.027700	15.847540	6.570655	-241.178600	15.530000	6.273343	3638.269000
201303132000	15.548705	7.377870	-323.206350	608.824850	345.827600	15.809820	6.577981	-244.981250	15.515960	6.274524	3672.277000
201303132200	15.879640	7.042122	-314.838100	550.713750	308.725350	15.749385	6.574461	-244.334250	15.492440	6.274525	3728.105500
201303140000	16.290175	6.982120	-311.023400	515.491050	289.999000	15.676135	6.556840	-229.770850	15.458515	6.282948	3645.488500
201303140200	15.996955	6.974000	-311.200200	502.758900	235.806250	15.607280	6.547013	-230.350750	15.445765	6.307301	3637.954000
201303140400	15.444205	7.021836	-310.394750	497.444600	198.915550	15.546490	6.543067	-232.547950	15.454275	6.326344	3636.944000
201303140600	14.693195	7.098575	-290.819950	481.442650	191.631700	15.461155	6.539588	-230.112650	15.423000	6.323821	3635.049500

201303140800	14.940525	7.119226	-291.668400	508.886950	381.872050	15.386200	6.516454	-221.268400	15.334385	6.311126	3630.391500
201303141000	15.832020	8.125403	-329.092550	520.658950	479.172950	15.383160	6.508601	-223.679250	15.203265	6.299753	3629.082500
201303141200	15.858745	7.926043	-325.965100	548.716600	393.493450	15.412205	6.529883	-226.944900	15.115750	6.282806	3631.431500
201303141400	15.874520	7.955239	-337.624000	573.343150	425.061750	15.468730	6.562985	-235.026100	15.124160	6.264047	3630.720000
201303141600	15.779295	7.740308	-338.594500	566.120450	454.766450	15.537105	6.575762	-230.814500	15.172805	6.256945	3635.062000
201303141800	15.691900	7.807143	-340.939100	586.466650	455.726900	15.578360	6.578611	-258.446000	15.227005	6.260243	3640.958500
201303142000	15.914315	7.909339	-340.704800	589.588500	361.754250	15.582160	6.588153	-253.111300	15.262295	6.270496	3647.425000
201303142200	16.300795	7.497768	-328.370000	532.064850	354.104150	15.533435	6.597594	-252.427800	15.266485	6.278797	3652.439000
201303150000	16.162005	7.409955	-314.447500	544.162250	287.028600	15.484125	6.582436	-252.275200	15.274540	6.287647	3652.708500
201303150200	15.737415	7.403649	-320.648850	509.558400	227.781900	15.422595	6.573422	-243.137950	15.299085	6.307200	3650.431000
201303150400	15.175505	7.416405	-315.338600	490.978750	187.147600	15.332385	6.570919	-244.035200	15.299005	6.326915	3651.699500
201303150600	14.624345	7.355144	-312.878800	485.249400	191.827150	15.242905	6.569557	-281.151800	15.250140	6.337149	3650.859500
201303150800	14.537065	7.789326	-316.706600	484.771700	277.319700	15.190540	6.575762	-343.702550	15.164155	6.336171	3645.346000

DATA FRÅN ON-LINESENSOERNA UNDER MÄTKAMPANJEN, DEL 2/2

Tid	O2_R4	O2_R5	O2_R6	Ventil_R4	Ventil_R5	Ventil_R6	Inflöde	Recirkulation	Returslam
201303092200	1.987002	1.954050	0.984647	72.525430	35.502440	17.313270	1.254867	5.023660	1.143594
201303100000	1.999759	1.968623	1.013517	74.591290	37.258100	18.261690	1.254561	5.028801	1.146465
201303100200	1.992517	1.969169	0.992004	75.941085	39.012030	18.733765	1.258603	5.025388	1.146085
201303100400	2.007225	1.988525	0.998454	75.732035	42.084720	19.022895	1.262456	5.025519	1.141397
201303100600	2.008059	1.999908	1.007678	73.082980	42.194040	18.906660	1.265459	5.028493	1.130573
201303100800	2.013577	2.008211	1.002052	69.547590	42.685435	19.352570	1.264375	5.022224	1.121988
201303101000	2.002993	2.083079	1.022741	66.696855	37.911100	17.335850	1.255219	5.026224	1.120875
201303101200	2.002709	1.987676	1.007388	65.978840	35.677675	15.254040	1.252978	5.013450	1.120040
201303101400	1.991922	1.982678	0.983512	67.164080	37.959755	15.353395	1.237905	5.017155	1.119761
201303101600	2.000513	2.019724	1.025071	67.709190	36.823375	15.610945	1.246107	5.013111	1.119352
201303101800	1.996384	2.027876	1.002118	68.609250	34.206415	15.282010	1.246870	4.996163	1.116201
201303102000	1.997158	1.984411	0.984156	70.047690	35.072425	16.550470	1.247440	5.020201	1.112523

201303102200	1.988709	2.041016	1.006972	70.707445	33.120920	16.865595	1.245741	5.009610	1.115015
201303110000	1.995649	1.982655	0.990305	73.027575	32.942110	16.669915	1.248231	5.015925	1.117768
201303110200	1.992616	1.992075	0.992429	74.720040	34.584435	17.469705	1.250167	5.007852	1.116949
201303110400	2.008813	1.987125	1.013248	73.212455	35.104040	16.632970	1.251147	5.004470	1.113945
201303110600	2.018029	2.018984	1.006735	69.858885	34.445425	16.195750	1.253433	5.012468	1.115806
201303110800	2.003807	2.019148	0.988795	67.484050	32.585245	14.613700	1.252349	5.018078	1.112144
201303111000	2.005229	1.960083	1.004967	66.491060	33.556590	13.661730	1.251645	5.005436	1.108671
201303111200	2.011602	2.037619	1.006960	65.278160	34.433745	13.428040	0.961963	3.851179	0.966782
201303111400	1.993795	1.998131	0.996310	64.979435	31.892060	12.309835	0.893919	3.579307	0.912657
201303111600	1.992757	1.998046	0.990556	65.858225	31.285530	12.951675	0.894739	3.582393	0.914088
201303111800	1.989428	1.991272	0.992404	66.849090	31.601680	13.358860	0.894857	3.581343	0.915277
201303112000	2.013054	1.977314	0.980348	67.938280	32.025235	15.013080	0.896995	3.587940	0.914999
201303112200	1.989331	1.986084	0.988749	68.050985	34.077360	15.945810	0.895618	3.596389	0.915702
201303120000	1.990865	2.015203	0.991442	68.522860	33.570275	16.240220	0.894007	3.595766	0.909564
201303120200	1.992676	2.054659	0.991849	68.477510	29.835825	15.652665	0.895604	3.586939	0.906089
201303120400	2.007426	2.008713	1.005005	67.707525	26.600970	14.809945	0.897244	3.583144	0.904636
201303120600	2.004132	2.006696	0.999064	66.839180	24.902615	14.382565	0.898475	3.589007	0.903007
201303120800	1.986012	1.992992	0.983380	65.076805	24.534035	13.667015	0.886155	3.584156	0.903268
201303121000	2.005336	1.986758	0.997515	64.553335	24.470140	12.991085	0.879827	3.589344	0.907493
201303121200	1.984785	1.834671	0.978879	66.487270	33.124320	14.905035	1.400931	5.497828	1.382837
201303121400	1.991254	1.980672	0.968586	66.952225	41.157275	17.484685	1.492811	5.992508	1.517991
201303121600	1.995690	1.994192	1.001272	67.177030	41.699365	19.096110	1.494596	5.992128	1.519281
201303121800	1.988363	1.962534	0.973543	69.249430	42.994920	22.054720	1.494582	5.999069	1.521536
201303122000	1.991074	1.999981	0.978780	71.184435	44.020645	24.361245	1.493717	6.007098	1.520584
201303122200	1.991316	1.970870	0.994245	72.482525	45.949795	26.285175	1.493248	6.012415	1.513757
201303130000	1.997620	2.027307	1.024128	73.461975	45.360625	21.024370	1.496751	6.011991	1.499607
201303130200	1.996542	2.022963	0.982747	73.600915	42.242680	18.520540	1.501409	6.013851	1.492767
201303130400	1.996847	2.000743	1.013749	72.428885	40.737975	17.868680	1.504397	6.016795	1.488854
201303130600	2.003296	2.007621	0.981484	69.805135	38.791815	17.373520	1.501805	6.023973	1.486058
201303130800	2.003419	2.011117	0.994486	67.331450	37.008435	15.646750	1.491594	6.019255	1.488796

201303131000	1.999451	1.963632	1.011440	65.894880	41.163285	14.422070	1.469516	6.042108	1.485177
201303131200	1.994975	1.978454	1.017686	66.042495	45.452855	11.777066	1.140153	4.523376	1.167825
201303131400	1.983296	2.111695	0.970219	66.566525	32.435935	11.356220	1.105977	4.411277	1.096072
201303131600	1.992271	1.954243	0.996910	67.265380	32.395580	13.580174	1.110840	4.414442	1.089012
201303131800	1.995870	1.968835	0.987229	68.589035	34.882860	14.911740	1.110020	4.400056	1.086683
201303132000	1.988586	1.974206	0.976671	69.386855	36.700425	16.110845	1.110371	4.406443	1.084999
201303132200	1.999272	1.979411	1.005008	69.803300	38.082750	16.689715	1.108862	4.399104	1.080121
201303140000	1.988953	2.033889	1.001360	70.242135	36.486750	15.925745	1.108262	4.398855	1.073003
201303140200	1.988953	2.037793	0.991325	70.907390	33.011285	15.033340	1.110707	4.400994	1.065752
201303140400	1.997846	1.996058	1.003386	70.460095	31.751625	14.791650	1.116757	4.397757	1.062236
201303140600	2.002115	2.003500	0.987798	69.014255	30.963075	14.173650	1.115249	4.392175	1.059876
201303140800	2.002709	2.023546	1.000087	67.188905	29.036060	12.814855	1.103823	4.394416	1.058764
201303141000	1.996870	1.977261	0.989160	66.638620	28.333075	11.861324	1.101377	4.398723	1.058061
201303141200	1.993389	1.994449	0.990113	66.150935	29.133565	11.117778	0.982339	3.951977	0.970941
201303141400	1.990786	1.987283	0.984271	66.483185	30.122405	11.636770	0.997443	4.014483	0.942143
201303141600	1.986411	1.982750	1.010989	67.177595	31.026380	12.034485	1.006979	4.014542	0.940121
201303141800	1.995689	1.986185	0.982767	68.596685	32.582965	13.491360	1.004474	4.011099	0.933354
201303142000	1.987286	1.972341	0.979001	69.911525	35.076920	15.200710	1.006115	4.027535	0.930380
201303142200	1.998842	1.986455	0.994108	70.403960	36.881505	15.796595	1.002381	4.021735	0.924989
201303150000	1.993733	2.022191	1.004379	71.557625	35.505325	15.585805	1.007345	4.014994	0.918192
201303150200	1.997177	2.025027	0.976727	71.401920	32.826210	14.708310	0.996154	4.018277	0.920595
201303150400	2.000753	2.020673	1.004099	70.801485	30.214285	14.531200	1.016237	4.025702	0.914486
201303150600	2.000591	2.019356	1.000337	69.731450	28.876750	13.866845	1.023283	4.023169	0.916317
201303150800	1.974164	1.988406	0.986984	70.839410	29.083000	12.940000	1.001939	4.021324	0.917299

BILAGA C

ANALYSRESULTAT FRÅN MÄTKAMPANJEN

Tid	COD_IN	N-tot_IN	NH4-N_IN	PO4-P_IN	P-tot_IN	SS_IN	NO3-N_R3	NH4-N_R6	NO3-N_R6	SS_R6
201303110000	198	-	-	2.14	3.18	78	0.189	0.699	1.03	1356
201303110200	173	-	-	2.45	3.20	118	0.150	0.828	1.08	2096
201303110400	131	-	-	2.49	3.15	104	0.183	0.830	1.27	2560
201303110600	103	-	-	2.51	2.99	102	0.158	0.602	1.52	2008
201303110800	119	-	-	2.83	3.81	206	0.168	0.707	1.61	2560
201303111000	161	-	-	3.69	4.44	278	0.166	0.734	2.26	2600
201303111230	212	-	-	2.76	4.54	266	0.181	0.765	0.930	2164
201303111430	244	-	-	2.95	4.40	220	0.148	0.979	0.963	2992
201303111630	283	-	-	2.75	4.33	112	0.162	0.842	1.00	2744
201303111830	252	-	-	3.52	4.50	196	0.170	0.817	1.07	3004
201303112030	236	-	-	2.73	3.99	178	0.166	0.788	1.08	3324
201303112230	197	-	-	-	-	226	0.169	0.755	1.03	3140
201303120030	137	26.5	18.7	1.23	2.15	182	0.241	1.16	1.25	1852
201303120230	116	29.8	22.0	1.34	2.49	104	0.217	2.11	1.37	3036
201303120430	106	29.9	22.6	1.96	2.75	40	0.178	1.51	0.885	3268
201303120630	89.2	30.2	23.2	2.74	3.20	228	-	1.46	0.733	3148
201303120830	98.6	35.2	28.4	3.54	4.27	236	0.207	1.48	1.32	3240
201303121030	139	36.4	26.1	3.69	4.46	274	0.212	1.54	2.18	3264
201303121300	164	42.0	27.0	2.94	4.20	290	0.191	1.21	1.30	3140
201303121500	179	39.1	27.9	2.63	4.08	288	0.368	1.03	0.92	3000
201303121700	213	46.0	30.8	2.64	4.61	320	0.216	0.991	1.48	2716
201303121900	191	43.6	29.5	2.92	4.41	236	0.228	0.828	1.45	2836
201303122100	175	32.9	24.3	2.36	3.15	250	0.240	0.778	1.71	2976

201303122300	208	33.1	23.7	3.01	3.94	208	0.354	0.708	1.36	2984
201303130100	158	31.5	22.8	3.09	3.71	134	0.539	-	1.44	1584
201303130300	141	30.2	23.5	3.21	3.49	86	0.510	-	1.27	2848
201303130500	98.2	29.4	22.2	2.90	3.22	126	0.220	-	-	2948
201303130700	84.8	34.1	26.1	-	-	280	0.396	-	1.95	1464
201303130900	127	36.5	28.1	3.89	4.49	272	0.391	-	2.2	2848
201303131100	182	36.6	26.6	4.32	4.95	252	0.441	-	2.54	2996
201303131330	192	38.9	29.5	4.12	4.77	258	0.239	-	0.819	3200
201303131530	224	41.8	30.6	4.40	5.02	308	0.380	-	0.839	2976
201303131730	209	45.4	34.8	3.97	4.36	392	0.265	-	0.929	3136
201303131930	178	38.9	27.5	3.54	4.27	268	0.285	-	-	3228
201303132130	188	42.7	25.9	3.61	4.10	198	0.254	-	1.24	3324
201303132330	200	35.3	25.9	3.53	4.00	88	0.232	-	1.55	3040
201303140130	151	31.5	23.5	3.03	3.59	130	0.222	-	1.16	2896
201303140330	124	30.1	23.1	3.12	3.61	54	0.194	-	0.629	2688
201303140530	97.3	29.2	21.8	2.97	3.25	44	0.289	-	1.26	3068
201303140730	145	37.3	27.9	3.62	4.03	220	0.242	-	1.32	2744
201303140930	112	35.2	24.8	2.94	3.38	224	0.226	-	1.22	3220
201303141130	131	41.4	24.0	2.22	2.75	176	0.277	-	0.579	2424
201303141400	244	40.4	32.2	4.42	4.89	74	0.221	-	0.507	2084
201303141600	246	41.6	33.0	4.70	5.16	404	0.305	-	0.471	3300
201303141800	223	43.4	33.3	4.42	4.90	270	0.257	-	0.587	3100
201303142000	193	35.4	27.2	3.80	4.33	196	0.293	-	0.854	2564
201303142200	183	33.2	23.4	3.52	4.05	232	0.290	-	0.642	3212
201303140000	163	31.3	23.3	3.19	3.71	134	0.256	-	1.72	1992
201303150200	128	32.8	23.0	2.90	3.32	110	0.220	-	1.78	2976
201303150400	-	31.2	23.3	2.97	3.38	82	0.322	-	1.14	3140
201303150600	-	31.8	24.0	3.05	3.52	214	0.242	-	1.74	1852
201303150800	137	40.1	29.6	4.38	4.78	284	0.217	-	1.71	3332
201303151000	106	39.5	27.0	4.05	4.69	278	0.255	-	1.82	2484