

MEMÒRIA DESCRIPTIVA

MEMÒRIA DESCRIPTIVA

1. INTRODUCCIÓ	4
1.1. ANTECEDENTS.....	4
1.2. OBJECTE	4
1.3. ABAST	5
2. LES AIGÜES RESIDUALS	7
2.1 INTRODUCCIÓ	7
2.2 IMPORTÀNCIA DE LA DEPURACIÓ DE LES AIGÜES RESIDUALS	7
2.3 PROBLEMÀTICA DE LES AIGÜES RESIDUALS	8
2.4 LEGISLACIÓ	8
2.5 TRACTAMENT DE LES AIGÜES RESIDUALS	8
2.5.1 Procés d'una estació depuradora convencional	9
2.5.2 Tractament biològic de les aigües residuals.....	9
3. REACTOR DISCONTINU SEQÜENCIAL.....	11
3.1 INTRODUCCIÓ	11
3.2 PLANTA PILOT	14
3.3 INSTAL·LACIÓ EXPERIMENTAL	15
3.4 DESCRIPCIÓ DEL CICLE DE TREBALL	16
3.5 MÒDUL DE MONITORITZACIÓ I CONTROL.....	17
4. CONTROL ESTADÍSTIC MULTIVARIABLE DE PROCESSOS PER LOTS	21
4.1 CONTROL ESTADÍSTIC DE PROCESSOS CONVENCIONAL	21
4.2 CONTROL ESTADÍSTIC MULTIVARIABLE DE PROCESSOS.....	23
4.3 ANÀLISIS DE COMPONENTS PRINCIPALS	24
4.3.1 Elecció del número de components.....	30
4.4 CONTROL ESTADÍSTIC MULTIVARIABLE PER A PROCESSOS PER LOTS	32
4.4.1 Desdoblament de la matriu tridimensional a matriu bidimensional.....	33
4.4.2 Preprocessat de les dades	35
4.5 ANÀLISI ESTADÍSTIC MULTIVARIABLE A NIVELL DE LOT	39
4.6 ANÀLISIS ESTADÍSTIC MULTIVARIABLE A NIVELL D'OBSERVACIÓ	41
4.7 PROGRAMARI PCA I MPCA	42
4.7.1 SIMCA-P.....	42
4.7.2 PLS_Toolbox 3.x	44
4.8 MODELS MPCA REALITZATS EN EL GRUP EXIT	51

5. ENTORN PER A L'ANÀLISI ESTADÍSTIC EN LÍNIA D'UN PROCÉS SBR.....	55
5.1 INTRODUCCIÓ	55
5.2 MÒDUL PER A L'ANÀLISI ESTADÍSTIC EN LÍNIA	56
5.2.1 <i>Funció</i>	56
5.2.2 <i>Funcionament</i>	56
5.3 MÒDUL PER A LA CREACIÓ DE MODELS	59
5.3.1 <i>Funció</i>	59
5.3.2 <i>Funcionament</i>	60
5.4 MÒDUL PER A LA VALIDACIÓ DE LOTS	64
5.4.1 <i>Funció</i>	64
5.4.2 <i>Funcionament</i>	64
5.5 FUNCIONS ESPECÍFIQUES.....	67
5.5.1 <i>Preprocessat de les dades</i>	67
5.5.2 <i>Creació de models</i>	68
5.5.3 <i>Projecció de nous lots</i>	70
5.5.4 <i>Connexió amb la base de dades MySQL</i>	70
5.5.5 <i>Connexió amb MATLAB</i>	72
5.6 INTERFÍCIES D'USUARI.....	74
5.6.1 <i>Interfície del mòdul per a l'anàlisi en línia</i>	74
5.6.2 <i>Interfície del mòdul de creació de models</i>	84
5.6.3 <i>Interfície del mòdul validació de lots</i>	93
6. VALORACIÓ ECONÒMICA	101
7. CONCLUSIONS	102
8. RELACIÓ DE DOCUMENTS.....	104
9. BIBLIOGRAFIA.....	105
10. GLOSSARI	108

1. INTRODUCCIÓ

1.1. Antecedents

En el grup d'investigació eXIT del Departament d' Electrònica, Informàtica i Automàtica (EIA) s'està treballant en la monitorització d'un Reactor Discontinu Seqüencial (*SBR, en anglès Sequential Batch Reactor*) mitjançant tècniques basades en el Control Estadístic Multivariable de Processos (*MSPC, en anglès Multivariate Statistical Process Control*).

Al llarg de les últimes dècades, els processos industrials han anat evolucionant esdevenint cada vegada més complexos. Aquesta complexitat comporta la utilització de conjunts de dades de grans dimensions. Per tal de poder treballar amb aquestes dades i obtenir-hi la informació rellevant dels processos present han aparegut tècniques basades en MSPC. Aquestes tècniques es basen en una reducció de la dimensionalitat de les dades originals.

Un SBR és un sistema de tractament d'aigües residuals en el qual l'aigua residual és introduïda de forma discontinua en el reactor, on és tractada i posteriorment descarregada. El funcionament d'un SBR es basa en la repetició cíclica d'etapes com *omplerta, barreja, reacció, sedimentació i buidat*.

El grup d'investigació LEQUIA del Departament d'Enginyeria Ambiental i Tecnologia Agroalimentària (EQATA) està treballant en l'estudi de les condicions d'operació d'un SBR. Aquest grup disposa d'una planta pilot situada a l' EDAR del Terri. Aquesta planta pilot disposa d'un mòdul de monitorització i control que permet l'obtenció en línia de les variables de procés. Aquest mòdul s'ha realitzat amb la col·laboració del grup eXIT. Aquest mòdul està pendent d'actualització en l'entorn LabVIEW. En l'actualitat està implementat en un entorn LABWindows/CVI.

En el grup eXIT s'han desenvolupat uns models basats en tècniques MSPC per al control de sistemes SBR. Aquests models han estat creats a partir de les dades obtingudes a la planta pilot mitjançant el mòdul de monitorització instal·lat.

1.2. Objecte

L'objecte del present projecte és la realització d'una aplicació que permeti portar a terme el control estadístic multivariable en línia de la planta SBR.

Aquesta aplicació ha de ser implementada en l'entorn LABVIEW, entorn sobre el qual s'està implementant l'actualització del mòdul de monitorització.

Aquesta eina ha de permetre realitzar l'anàlisi estadístic del lot en procés, de l'últim lot finalitzat i de la resta de lots processats a la planta. Per a l'obtenció de les dades corresponents a aquests lots, l'aplicació ha d'establir una connexió amb la base de dades del mòdul de monitorització. En aquesta base de dades s'emmagatzemen els valors de les variables de procés mesurades per a cadascun dels lots.

L'aplicació ha de permetre l'elecció de diferents models MSPC. L'anàlisi estadístic es portarà a terme mitjançant la projecció de les dades dels lots sobre aquests models. L'aplicació també ha de permetre la creació de nous models a partir de dades de nous lots i l'edició de models ja creats. Per a la creació i la importació d'aquests models a l'entorn LABVIEW, l'aplicació ha d'establir una comunicació amb el programa MATLAB. La creació de models es portarà a terme mitjançant la *toolbox* de MATLAB PLS_Toolbox 3.5. Aquesta *toolbox* té implementat un conjunt d'algoritmes referents a les tècniques MSPC.

Tot i les connexions necessàries a certs programes externs, l'aplicació s'ha de construir de la manera més autònoma possible en termes de càlcul. Els càlculs corresponents als anàlisis estadístics s'han de realitzar de manera interna des de la pròpia aplicació.

L'aplicació ha de presentar un conjunt de gràfics de control que permetin visualitzar de manera clara els resultats obtinguts dels corresponents anàlisis realitzats dins l'espai multivariable. Aquests gràfics han d'estar recolzats per les taules i informacions necessàries.

1.3. Abast

Aquesta aplicació ha de ser incorporada a l'actualització que s'està realitzant del mòdul de monitorització.

Primerament, l'aplicació inserida dins el mòdul de monitorització ha de ser provada a la planta experimental SBR que disposa el grup Lequia a les seves instal·lacions. En aquesta planta es realitzaran proves mitjançant els models de referència ja creats.

Un cop portades a terme les comprovacions d'ajustatge i bon funcionament conjunt dels dos entorns, aquesta aplicació s'instal·larà en la planta pilot SBR situada en l'actualitat a l'EDAR

del Terri. En aquesta planta serà necessari una actualització dels models de referència degut a la diferent naturalesa de les aigües tractades.

L'aplicació podrà ser utilitzada en futures instal·lacions de plantes SBR per a portar a terme el control estadístic multivariable de les mateixes.

2. LES AIGÜES RESIDUALS

2.1 Introducció

Tota comunitat genera residus tant sòlids com líquids. La fracció líquida dels mateixos és essencialment l'aigua de que es desprèn la comunitat una vegada ha estat contaminada durant els diferents usos pels quals ha estat utilitzada. Des del punt de vista de fonts de generació, es pot definir l'aigua residual com la combinació de líquids o aigües que transporten residus procedents de residències, instal·lacions públiques i centres comercials i industrials, a les que, eventualment, s'hi poden afegir subterrànies, superficials i pluvials.

2.2 Importància de la depuració de les aigües residuals

La creixent industrialització del món, ha comportat un agreujament del problema de les aigües residuals. El volum d'aigua residual generada ha anat augmentant així com la seva concentració i l'aparició de nous compostos de síntesi.

La natura presenta un complex ecosistema que interrelaciona els canvis que afecten a la vida dels éssers que hi habiten, i que li permet actuar i corregir les petites perturbacions que modifiquin l'equilibri natural. Dins aquest ecosistema global, el cicle hidrològic és el que garanteix la circulació i disponibilitat de les aigües.

L'home ha anat adquirint protagonisme dins aquest cicle fins a esdevenir un factor determinant. Els ingents volums d'aigües residuals produïdes en determinades ciutats del món, així com les elevades concentracions d'estranyes combinacions de compostos característics de la indústria moderna han superat amb escreix la capacitat d'autodepuració del medi natural.

Per tant l'home resta com a principal responsable de la contaminació dels recursos hidràulics, però alhora és l'únic que pot actuar sobre els mateixos per tal de reduir l'impacte d'aquests residus fins a límits assimilables per l'entorn natural.

2.3 Problemàtica de les aigües residuals

La composició característica de l'aigua residual suposa que el seu abocament directe a un medi aquàtic pugui produir una sèrie de conseqüències que dificultin la vida al seu voltant. L'elevada presència de microorganismes patògens afavoreix la transmissió de malalties com la gastroenteritis, el còlera, la disenteria, el tifus o l'hepatitis A. Aquesta conseqüència es fa encara palesa en indrets subdesenvolupats on l'expansió d'epidèmies molt perilloses és constant degut a la nul·la canal·lització existent.

L'alt contingut en matèria orgànica facilita tant l'aparició d'olors desagradables com una disminució del nivell d'oxigen dissolt a l'aigua que dificulta la vida aquàtica. Una elevada concentració de nutrients pot provocar l'eutrofització de l'aquífer, ocasionant mortaldat de peixos i plantes. La presència a nivell de traça de metalls o compostos tòxics, poden arribar a ser letals si es permet un augment significatiu de la seva concentració per acumulació.

2.4 Legislació

El Pla de Sanejament de la Generalitat de Catalunya, basant-se en la directiva del Consell 91/271/CEE del Diari Oficial de les Comunitats Europees, insta a la depuració dels principals components de l'aigua residual urbana que s'aboqui a llera pública.

S'estableix un límit de fins a 25 mg/l per a la DBO₅, amb un percentatge mínim de reducció del 70-90%, mentre que pels SS el límit queda fixat en 35mg/ amb un percentatge mínim de reducció del 90%. En quant als abocaments a zona sensible (amb risc d'eutrofització), s'hi addiciona un límit pels nutrients establert en 1 mg/l (i 80% de reducció) pel fòsfor, i en 10mg/l pel nitrogen total (amb 70-80% mínim d'eliminació). Aquests límits són lleugerament més permissibles en cas que la depuradora tracti cabals petits (d' entre 10.000 i 100.000 habitants equivalents).

2.5 Tractament de les aigües residuals

La depuració de les aigües residuals pretén aconseguir un vessament lliure de tots aquells compostos contaminants que puguin provocar efectes negatius sobre el medi receptor.

Diverses operacions i processos unitaris són requerits per tal de tractar adequadament les aigües residuals. Diferents combinacions d'aquests sistemes, tant físics, químics com

biològics, conformen el diagrama de procés de cada estació depuradora. El procés global sempre segueix una seqüència lògica de tractament.

2.5.1 Procés d'una estació depuradora convencional

Un cop arriba el volum d'aigües residuals a l'estació depuradora es realitza una primera etapa, anomenada pretractament, que realitza un primer desbast groller dels sòlids més grans arrossegats per l'aigua residual que arriba del col·lector. La finalitat del pretractament és evitar possibles obturacions posteriors, així com eliminar l'efecte abrasiu d'aquests materials sobre mecanismes com les bombes i vàlvules que es troben al llarg del procés.

Seguidament es realitza una segona etapa on l'aigua es deixa reposar unes hores a un tanc de sedimentació primària, per tal que decanti la matèria orgànica sedimentable, així com la resta de sorres o partícules inorgàniques que no han quedat retingudes al pretractament. Els sòlids sedimentats són enviats cap a una línia de tractament específic, la línia de fangs, essent habitual el seu pas previ per un gavell que separi les petites partícules inorgàniques contingudes.

L'etapa següent és la més important del procés anomenat tractament secundari. El fonament d'aquesta etapa no és altra que accelerar el procés biològic que es donaria a la natura, és a dir, la degradació, per part d'una població multiespecífica de microorganismes, de la matèria orgànica dissolta a l'aigua residual. Aquesta reacció té lloc a uns bioreactors, fortament airejats en cas que el procés sigui aerobi.

Seguint el camí que recorre l'aigua dins l'EDAR, la última de les etapes habituals és una nova decantació uns sedimentadors secundaris. L'objectiu és assolir una bona separació entre l'aigua tractada i la biomassa present. El sobrenedant, ben clarificat, sol ser abocat directament cap al llit receptor on segueix el seu cicle natural.

2.5.2 Tractament biològic de les aigües residuals

El procés més habitual de tractament biològic de les aigües residuals és l'anomenat sistema de fangs actius. Consisteix en una oxidació bacteriana del residu orgànic, seguit d'una separació entre els sòlids en suspensió o l'aigua tractada. Simplificant el procés, es pot considerar que els microorganismes utilitzen l'oxigen present a l'aigua per tal de consumir el substrat o aliment, en aquest cas les molècules orgàniques biodegradables contingudes a l'aigua residual. Com a resultat d'aquest consum els microorganismes obtenen l'energia

necessària per mantenir les seves funcions vitals, alhora que generen nous individus. A la Figura 1 es pot observar el diagrama d'un sistema convencional de fangs actius.

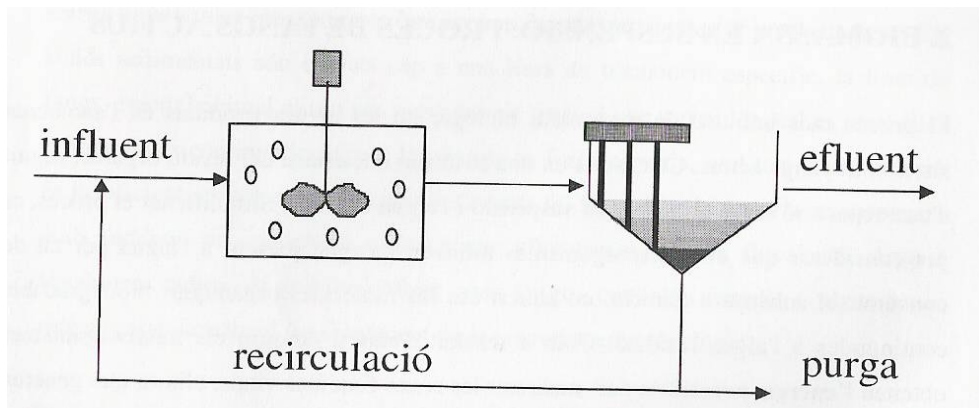


Figura 1. Diagrama d'un sistema convencional de fangs actius

En certa manera, el que s'aconsegueix és transformar la fracció soluble de matèria orgànica en una material insoluble, fet que facilita la seva posterior separació amb una simple decantació. La major part dels microorganismes separats al decantador és retornada cap al reactor biològic per tal de mantenir el nivell de depuració necessari, mentre que una petita fracció és apartada diàriament del sistema i enviada cap a la línia de fangs, per tal d'evitar un augment i envelliment excessiu de la biomassa present al sistema. Aquestes dues accions, clau per garantir el correcte desenvolupament del procés de fangs actius, són les anomenades recirculació i purga.

3. REACTOR DISCONTINU SEQÜENCIAL

3.1 Introducció

El sistema de fangs actius descrit en el capítol anterior, va ser presentat per primer cop per Ardern i Lockett el 1914. Els sistemes en lot, no van ser aplicats mai en aquelles èpoques com a conseqüència de les dificultats d'operació que suposaven en aquells moments (obrir i tancar vàlvules, controlar el temps,...).

Anys més tard, amb l'evolució del maquinari i els sistemes computaritzats de control (PLC), pels anys setanta, aquests sistemes es varen començar a introduir.

És a partir dels anys vuitanta, que l'Agència dels Estats Units per la Protecció Ambiental (USEPA) va fer una de les majors contribucions al desenvolupament d'aquesta tecnologia amb la promoció de diferents estudis de reintroducció al mercat del reactors biològics en lot.

Com el seu nom indica, un Reactor Discontínu Seqüencial (a partir d'ara SBR, de l'anglès Seqüencial Batch Reactor) és un sistema de tractament de fangs actius en el qual l'aigua residual és introduïda de forma discontinua en el reactor, on és tractada i posteriorment descarregada. Mentre que en el sistema convencional aquestes etapes tenen lloc en reactors separats i es produeixen de manera simultània, en els SBRs es produeixen en un mateix reactor però de manera seqüencial en el temps.

L'operació de la major part dels SBRs que s'utilitzen en l'actualitat consta de cinc etapes tal i com es mostra a la Figura 2 que es realitzen de manera seqüencial.

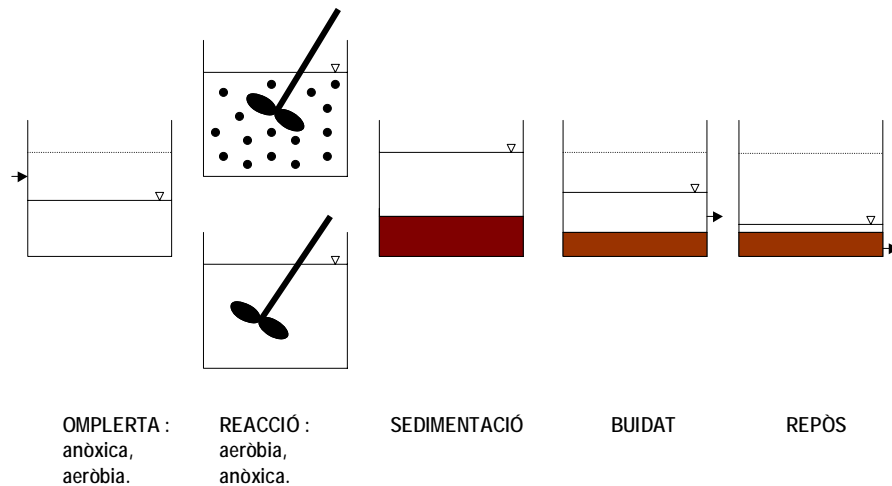


Figura 2. Etapes típiques d'un SBR

Omplerta: Durant aquesta etapa l'aigua residual a tractar entra en el reactor. El grau d'interacció que es produeix entre la biomassa que està en el reactor del cycle anterior i l'aigua residual que entra depèn de l'estat dels sistemes d'aeració i d'agitació.

Reacció: En aquesta etapa tenen lloc les transformacions bioquímiques portades a terme pels microorganismes. De la mateixa manera que durant l'etapa d'omplerta, en aquesta fase també es poden imposar diverses condicions ambientals en el reactor, el que permet que aquesta etapa es pugui dividir en diverses subetapes:

Barreja: El contingut del reactor es manté barrejat però no està airejat. Es poden produir, per tant, reaccions anòxiques o anaeròbies.

Barreja aeròbica: En aquest cas el contingut del reactor es manté barrejat i airejat, el que permet que tinguin lloc reaccions aeròbies.

Sedimentació: En aquesta etapa es deté tant l'aeració com l'agitació, per a permetre la separació entre els sòlids i poder obtenir un efluent clarificat. L'eficiència de la decantació sol ser major que en els sistemes en continu.

Buidat: En aquesta etapa l'aigua clarificada es retira del reactor, de manera que aquest queda disponible per a rebre un nou volum d'aigua residual a tractar.

Repòs: És el temps que transcórrer des de que acaba la fase de descàrrega fins que comença un nou cycle. En sistemes amb diversos SBRs en paral·lel, aquesta fase permet que un reactor acabi la seva etapa d'omplerta abans d'iniciar l'omplerta de l'altre reactor. No

és una fase necessària, i en cas que es pugui aconseguir un cabal d'entrada constant, per exemple mitjançant la instal·lació d'un tanc de regularització, es pot suprimir.

Un SBR presenta diverses opcions pel que respecte a la seva alimentació, com poden ser l'alimentació única, l'alimentació esglaonada o múltiple i l'alimentació continuada.

Alimentació única: Es realitza al principi del cicle i pot anar combinada de reacció, ja sigui aeròbia, per eliminar la matèria orgànica i nitrificar, o anòxica, per desnitrificar. És capaç d'assolir fàcil i ràpidament la nitrificació però presenta inconvenients com l'absència de matèria orgànica per poder desnitrificar i l'alta variabilitat de la càrrega dins del reactor.

Alimentació esglaonada o múltiple: Consisteix en realitzar més d'una alimentació al llarg del cicle, aquestes es solen fer en condicions anòxiques per tal d'afavorir la desnitrificació del sistema. D'aquesta manera s'aconsegueix treballar amb unes menors variacions de càrrega a l'interior del reactor. El producte residual final serà inversament proporcional al número d'alimentacions realitzades.

Alimentació continuada: En aquest cas el reactor treballa amb una alimentació contínua al llarg de totes les etapes del cicle, alternant fases aeròbies i anòxiques. D'aquesta manera s'aconsegueix tenir matèria orgànica per poder desnitrificar. El fet d'alimentar contínuament provoca que a la sortida sempre es tinguin unes concentracions residuals de nitrats i amoni.

La principal avantatge que presenten els sistemes SBR és la seva flexibilitat. Aquesta característica rau en la possibilitat de controlar de manera senzilla el temps que es dedica a cadascuna de les etapes del procés. A cada etapa s'imposen unes condicions ambientals determinades, i per a modificar la duració de cadascuna d'elles tan sols és necessari actuar sobre els controladors que connecten i desconnecten les bombes. D'altra banda, en els sistemes de tractament convencional en continu, per a un cabal d'aigua residual donat, la única forma de modificar el temps de cada etapa és alterant el tamany dels corresponents reactors.

3.2 Planta Pilot

Els grups de recerca Lequia i eXIT disposen d'una planta pilot d'un SBR que treballa en fase experimental i que ha estat emplaçada en diverses EDARs de les comarques gironines.

En l'actualitat es troba en funcionament a l'EDAR del Terri. En aquesta EDAR es tracta l'aigua residual dels municipis de Banyoles i altres pobles de la comarca del Pla de l'Estany amb una població servida de 185.560 habitants equivalents. Un cop depurada s'aboca al riu Terri. En l' EDAR del Terri, es tracten aigües residuals d' origen domèstic i també les produïdes per algunes indústries de la comarca. Aquestes indústries estan obligades a disposar de sistemes de depuració propis per tal que les aigües que aboquin a la xarxa de sanejament tinguin característiques similars a les aigües residuals domèstiques

La planta pilot s'alimenta de les mateixes aigües que arriben a l'EDAR del Terri, sense que aquestes pateixin cap pretractament o tractament primari anterior. A la Figura 3 es mostra la planta pilot ubicada dins un mòdul prefabricat.



Figura 3. Planta pilot SBR emplaçada a l'EDAR del Terri

El reactor és un cub d'acer inoxidable d'1m³ de capacitat. En cada cicle són tractats 200 litres d'aigües residuals. Hi ha un reactor de pretractament de 300 litres on l'aigua a tractar hi roman, en condicions d'agitació, abans de passar al reactor biològic. La planta pilot disposa d'un equip d'aireació per les etapes aeròbiques, d'un agitador, de diferents bombes i vàlvules per l'alimentació i l'extracció de l'aigua, així com les instal·lacions elèctriques necessàries pel seu funcionament. La planta pilot té instal·lats quatre sensors que recullen

diferents senyals a temps real referents al pH, l'Oxigen Dissolt (OD), el potencial RedOx i la Temperatura.

En aquesta planta pilot hi ha instal·lat un mòdul de monitorització que permet la visualització a temps real de les variables mesurades per les sondes.

3.3 Instal·lació experimental

En el grup de recerca de Lequia disposen d'una instal·lació experimental on es porten a terme diferents assajos per tal d'analitzar el comportament del reactor davant diverses condicions determinades així com a banc de proves pels diferents controls que s'hi implementen. Un cop portades a terme les comprovacions pertinents els controls són aptes per ser instal·lats a la planta pilot.

La instal·lació experimental consta d'un reactor de vidre cilíndric que acaba amb una forma cònica per tal facilitar-ne la sedimentació. Aquest té un volum màxim aproximat de 40 litres i el volum de treball emprat habitualment està entre els 20 i els 30 litres. El volum alimentat es manté sempre a 10 litres per cicle. La biomassa es manté en suspensió mitjançant agitació mecànica amb una hèlix marina. L'aeració es realitza a través d'un tub difusor col·locat a la part inferior del reactor i connectat a un compressor. L'alimentació, buidat i purga es realitzen amb tres bombes peristàltiques diferents. A la Figura 4 es mostra la planta experimental SBR i el diagrama de connexions corresponents.

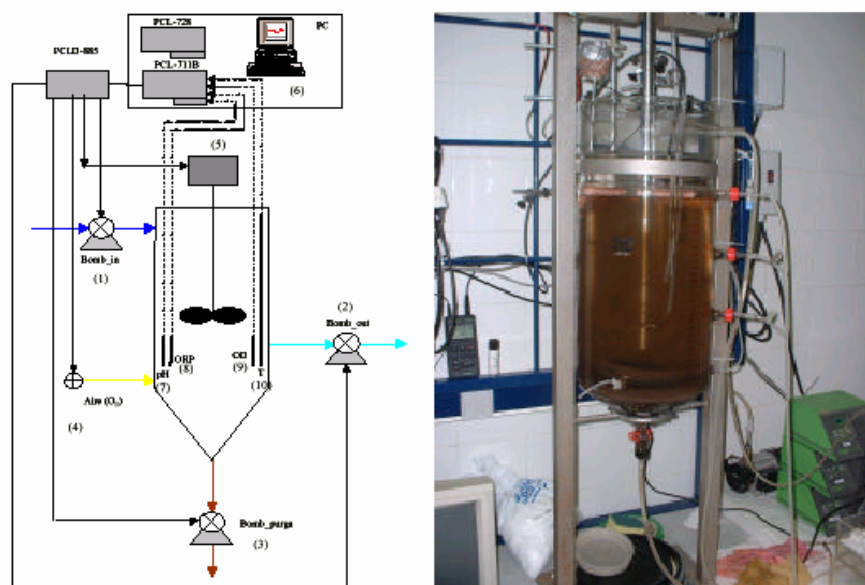


Figura 4. Planta experimental SBR i diagrama de connexions

A l'interior del reactor es troben col·locades, amb un sistema de flotació, quatre sondes: Temperatura, pH, RedOx i OD, i dins el reactor una resistència calefactora, connectat tot a un ordinador. Aquestes sondes permeten la monitorització de manera que es pot veure l'evolució d'aquests paràmetres en línia i mitjançant la resistència calefactora es pot augmentar la temperatura fins a la temperatura desitjada. Mitjançant un programa informàtic i les plaques de relés, es poden fixar totes les condicions del cicle, l'aireig, l'agitació, les bombes o la temperatura així com el seu temps de durada en cada cicle. Amb tot això el que s'aconsegueix és automatitzar la planta.

3.4 Descripció del cicle de treball

El funcionament de la planta pilot i la instal·lació experimental SBR segueixen una pauta d'operacions anomenada seqüència de treball. Una seqüència de treball està format per una successió contínua de cicles. Un cicle és una combinació de les etapes descrites en l'apartat 3.1. Un cicle de treball també és anomenat pla de treball.

El cicle de treball de la planta pilot i la instal·lació experimental del SBR té una durada de vuit hores i està dividit en nou etapes tal i com es mostra a la Figura 5. Les sis primeres etapes estan formades alhora per altres tres etapes (alimentació, etapa anòxica i etapa aeròbica). La setena etapa és la purga de fangs, la vuitena la sedimentació dels fangs actius i la última etapa correspon al buidat del reactor.

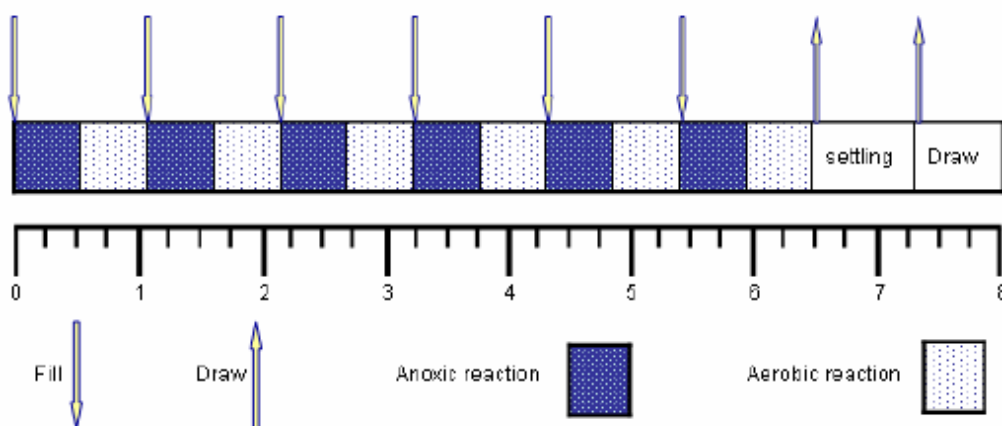


Figura 5 . Cicle de treball per a la planta pilot i instal·lació experimental

L'alimentació del reactor en les cinc primeres etapes té una durada de quinze minuts, excepte l'alimentació de la sisena etapa que dura cinc minuts. Cada etapa anòxica dura vint minuts i cada etapa aeròbica trenta minuts. La purga es realitza durant un temps de tres minuts. Seguidament es deixa que els fangs actius sedimentin en el temps d'una hora i finalment es deixa buidar el reactor durant vint-i-cinc minuts.

En l' etapa d'alimentació funcionen tant l'agitador com la bomba d'alimentació del reactor. En l'etapa anòxica només funciona l'agitador. En l'etapa aeròbica s'obre la vàlvula de tres vies i s'activen l'agitador i el soplant. En l'etapa de purga es fa actuar a la bomba de purga i l'agitador. Finalment en l'etapa de buidat només s'obre la vàlvula de buidat. En l'etapa de sedimentació tota la instal·lació roman en repòs.

Totes aquestes etapes es detallen a la Taula 1 juntament amb els dispositius que s'activen a cadascuna de les etapes del procés.

		PROCÉS	TEMPS (min)	DISPOSITIUS ACTIUS
REACTOR FANGS ACTIUS	ETAPA 1	alimentació	15	bomba reactor, agitador
		anòxic	20	agitador
		aeròbic	30	vàlvula 3v, agitador, soplant
	ETAPA 2	alimentació	15	bomba reactor, agitador
		anòxic	20	agitador
		aeròbic	30	vàlvula 3v, agitador, soplant
	ETAPA 3	alimentació	15	bomba reactor, agitador
		anòxic	20	agitador
		aeròbic	30	vàlvula 3v, agitador, soplant
	ETAPA 4	alimentació	15	bomba reactor, agitador
		anòxic	20	agitador
		aeròbic	30	vàlvula 3v, agitador, soplant
	ETAPA 5	alimentació	15	bomba reactor, agitador
		anòxic	20	agitador
		aeròbic	30	vàlvula 3v, agitador, soplant
ETAPA 6	alimentació	5	bomba reactor, agitador	
	anòxic	30	agitador	
	aeròbic	32	vàlvula 3v, agitador, soplant	
ETAPA 7	purga fangs	3	bomba purga, agitador, soplant, vàlvula 3v	
ETAPA 8	sedimentació	60	-----	
ETAPA 9	buidat	30	vàlvula de buidat	

Taula 1. Detall de les etapes que conformen el cicle de treball

3.5 Mòdul de monitorització i control

Tant a la planta pilot com a la instal·lació experimental es troba instal·lat un mòdul de monitorització i control. El mòdul de monitorització és el responsable de l'adquisició de

dades en línia procedents de les sondes de pH, Oxigen Dissolt, Redox i Temperatura. Aquest mòdul també incorpora un mòdul de càlcul d'una variable que es pot calcular de forma contínua com són el valors de la velocitat de consum d'oxigen (OUR, de l'anglès *oxygen uptake rate*).

A part d'aquestes mesures procedents de les sondes, també es pren registre del funcionament mecànic dels diferents elements elèctrics de la planta pilot com són les bombes d'alimentació, buidat i purga, aireació, cabalímetre màssic d'aire, agitació.

Altres funcions que porta a terme el mòdul de monitorització i control són la definició dels plans i seqüències de treball, el control de les etapes, la calibració de les variables d'entrada i el control d'oxigen dissolt en l'etapa aeròbica.

En aquests moments, el mòdul de monitorització i control que es troba instal·lat a la planta pilot i a la planta experimental està implementat amb el programa LABWindows/CVI. En l'actualitat s'està desenvolupant un nou mòdul de monitorització i control. Aquest nou mòdul de monitorització i control s'està implementat amb el programa LabVIEW.

L'emmagatzematge de les dades del procés és una de les funcions més importants del mòdul de monitorització. En l'actualització del mòdul de monitorització implementat en LabVIEW aquestes dades són emmagatzemades en una base de dades del tipus MySQL. En aquesta base de dades són emmagatzemades tant les dades procedents de les sondes anteriorment esmentades com els valors de la variable calculada OUR.

El conjunt de dades mesurat durant un cicle o pla de treball és emmagatzemat en una taula de la base de dades. En cas que una seqüència de treball estigui formada per la combinació de cicles diferents, en una mateixa taula es guarden tants lots com tipologies s'hagin definit en la seqüència de treball. Actualment, els plans de treball estan formats per la repetició d'un mateix tipus de cicle, definit a l'apartat anterior. En aquest cas a cada taula es guarden les dades d'un únic lot.

Cadascuna de les taules de les taules segueix la nomenclatura que es detalla a la Figura 6; on s'inclou informació referent a la data i hora d' inici del processat del lot, el nom de la seqüència de treball a què pertany i l'ordre del cicle dins la seqüència de treball.

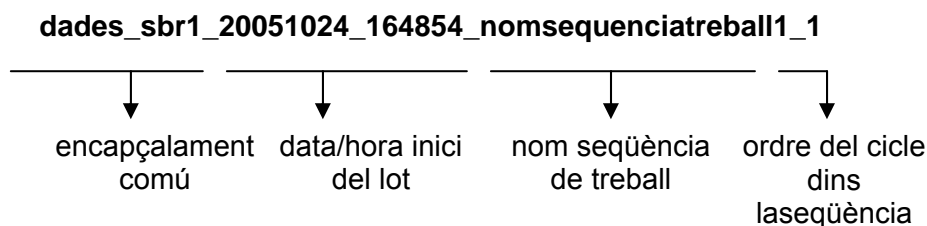


Figura 6. Nomenclatura de les taules de la base de dades

Les taules de la base de dades estan formades per nou camps tal i com es detalla a la Taula 2. En aquesta taula s'indica la tipologia dels camps i una breu descripció de la informació que contenen.

Nom del camp	Tipologia	Descripció
ID_Seq_Treball	Int(11)	Nº seqüència a què pertany el lot
ID_Pla_Treball	Int(11)	Nº de pla de treball dins la seqüència
ID_Etapa	Char(11)	Etapa actual
Data_Hora	Date/time	Data i hora de la presa de mesures
pH	Float	Mesura de la sonda pH
Oxigen	Float	Mesura de la sonda d'Oxigen
Temperatura	Float	Mesura de la sonda de Temperatura
Redox	Float	Mesura de la sonda de Redox
Flag	Int(1)	Mesura necessària (o no) per a MSPC

Taula 2. Llistat dels camps d'una taula de la base de dades

Actualment, l'emmagatzemament de les dades del procés es porta a terme cada cinc segons. Tenint en compte que la duració total d'un lot és de vuit hores, cada cicle està format per 5760 registres, cadascun d'ells contenint la informació dels camps que s'ha detallat a la Taula 2.

El fet que els processos biològics siguin processos lents i les variacions entre instants de temps siguin mínimes ha comportat la introducció a les taules del darrer camp anomenat *flag*. Aquest camp permet filtrar les dades necessàries per portar a terme anàlisis que no requereixen la totalitat de les dades que formen un lot. Aquest camp permet alleugerir el

procés d'importació de dades des de la base de dades cap a altres aplicacions. Aquest camp *flag*, està configurat de tal manera que es posa a '1' una vegada per minut. D'aquesta manera el volum de les dades per a cada cicle es redueix aproximadament en un 90%.

Pel que respecte a l'emmagatzemament dels valors resultants del càlcul de l'OUR, es guarden en taules que segueixen una nomenclatura similar a les referents a les variables de procés canviant l'encapçalament per *our_sbr1*. La composició de camps d'aquestes taules també és similar a les anteriors emmagatzemant en aquest cas només el valor de la OUR resultant.

4. CONTROL ESTADÍSTIC MULTIVARIABLE DE PROCESSOS PER LOTS

4.1 Control estadístic de processos convencional

El control estadístic de processos (SPC, de l'anglès *Statistical Process Control*) neix de la idea que tot procés experimenta variacions en el seu funcionament. Aquestes variacions poden afectar en major o menor mesura a la qualitat del producte final. La variabilitat que presenta un procés en el seu comportament pot ser deguda a dues causes:

-*Causes comuns*: són aquelles que sempre estan presents en el procés, són innates al mateix, i són conseqüència de com s'ha dissenyat el procés i com opera normalment. Quan el procés està afectat per aquest tipus de variabilitat es diu que està sota control estadístic.

-*Causes especials*: són aquelles que apareixen de manera esporàdica i que són conseqüència d'anomalies o falles en el procés, originant que aquest tingui un comportament diferent a l'habitual. Quan el procés està afectat per causes especials de variabilitat, es diu que està fora de control estadístic.

L'objectiu del SPC és establir un sistema permanent i intel·ligent de monitorització d'un procés continu al llarg del temps per tal de detectar estadísticament situacions significatives o anormals que afectin al procés i identificar l'origen de les mateixes.

El SPC es fonamenta en la presa periòdica de dades del procés i la representació gràfica d'unes caixes de control com la de la Figura 7 on es delimita la regió de correcte funcionament del procés a partir d'un límit superior i inferior. Aquestes caixes faciliten la detecció visual de senyals d'avís que indiquen anomalies. La identificació de les causes que provoquen comportaments anòmals del procés (causes especials) és el primer pas en la presa de mesures pertinents per a la seva eliminació.

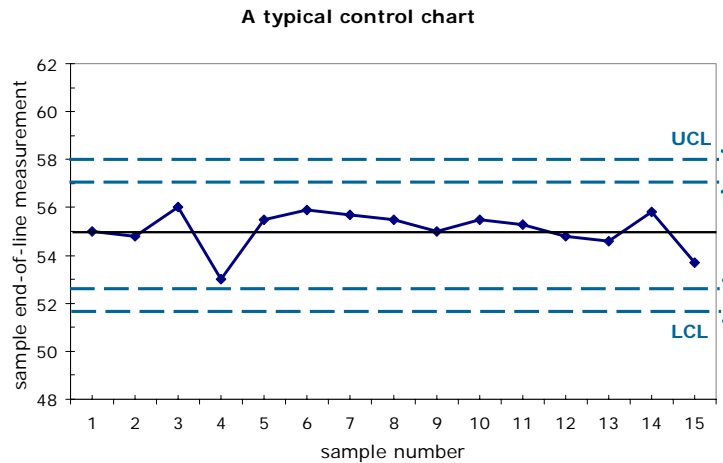


Figura 7 . Caixa de control per a un procés univariable

En aquesta caixa de control es defineixen uns límits, superiors i inferiors, que delimiten el correcte funcionament del procés. Aquests límits són calculats a partir de la mitjana i la desviació estàndard del conjunt de dades utilitzades. El límit superior es calcula a partir de l'Equació 1 i el límit inferior es calcula a partir de l'Equació 2.

$$LCS = \text{mitjana} + n \cdot \text{desviació estàndard} \quad (\text{Eq. 1})$$

$$LCI = \text{mitjana} - n \cdot \text{desviació estàndard} \quad (\text{Eq. 2})$$

El paràmetre n varia segons al nivell de confiança que volem establir el límit. Per norma general es treballa amb nivells de confiança del 95% i 99% que corresponen als valors $n=2$ i $n=3$, respectivament.

En el cas d'haver de monitoritzar diverses variables que no estiguin correlacionades entre si, es construeix per a cadascuna d'elles un gràfic de control univariant. Aquest procediment, per altra banda, no és adequat quan les variables de estan correlacionades, ja que el procés està fora de control sense que els gràfics univariants ho detectin. En aquest cas es creen regions errònies de correcte funcionament, degut a que els gràfics de control no tenen en compte la correlació entre les variables. És necessari doncs, l'establiment d'uns límits de control que tinguin en compte la correlació existent entre les variables de què es disposen; aquests límits seguiran una el·lipse en el cas de monitoritzar dues variables correlacionades. A la Figura 8, s'observen les falses regions de correcte funcionament creades en tractar com a variables independents variables que estan correlacionades. La regió de correcte

funcionament dins l'espai multivariable és la delimitada per l'el·lipse, que és combinació dels límits individuals de les variables tractades.

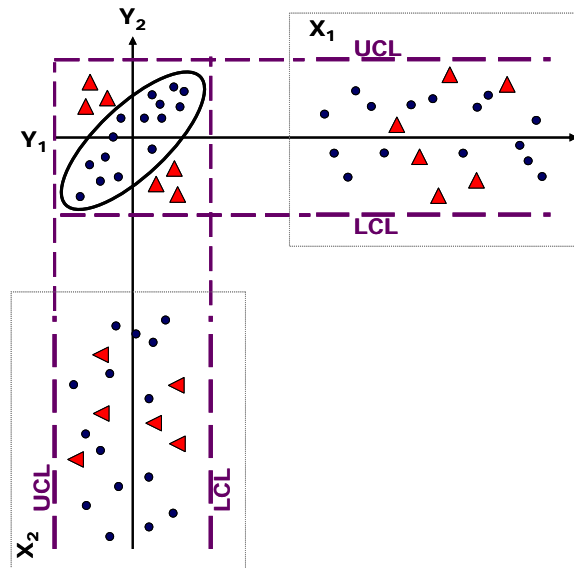


Figura 8 . Anàlisi multivariable vs Anàlisi univariante

Per tant, els gràfics de control constitueixen l'eina bàsica del SPC. Segons el número de variables que permeten monitoritzar simultàniament es denominen gràfics de control univariants (una sola variable) o gràfics de control multivariants (diverses variables).

4.2 Control estadístic multivariable de processos

El desenvolupament tecnològic de l'electrònica i les telecomunicacions ha comportat que la monitorització de processos hagi de manipular una gran quantitat d'informació registrada en temps real per variable de procés, treballant d'aquesta manera amb matrius de grans dimensions, dades mal condicionades i dades mancants. Per a treballar amb tota aquesta informació registrada s'han creat mètodes de projecció sobre estructures latents.

Aquest mètodes, són tècniques de reducció de la dimensionalitat de les dades que simplifiquen i afinen el procediment per a la supervisió de processos. Permeten sintetitzar tota la informació continguda en una base de dades en unes poques variables no observables (variables latents) que expliquen gran part de la variabilitat de les variables

originals (variables observables que han estat mesurades), i de les relacions que existeixen entre les mateixes. En el nou subespai creat de menors dimensions, és possible utilitzar tècniques de SPC, mitjançant l'aplicació dels gràfics de control a les variables latents. La monitorització d'aquestes variables latents permet de manera indirecta controlar la gran quantitat de variables que es mesuren en el procés. El conjunt de tècniques basades en aquestes mètodes formen el control estadístic multivariable de processos (MSPC, de l'anglès *Multivariate Statistical Process Control*).

Una de les tècniques de MSPC més utilitzades és l'anomenada *Anàlisi de Components Principals* (PCA). En l'àmbit dels processos actuals també s'utilitzen altres tècniques com són les anomenades *Regressió en mínims quadrats parcials* (PLS) i *Regressió en Components Principals* (PCR).

4.3 Anàlisi de components principals

L' Anàlisi de Components Principals (a partir d'ara PCA) és una tècnica estadística de síntesis de la informació, o reducció de la dimensió (respecte el número de variables de procés). El propòsit d'aquesta tècnica és descomposar una matriu de dades per a descobrir els fenòmens subjacents, la informació rellevant del procés que contenen. Per tant, una suposició bàsica del PCA és el fet que les direccions de major variància estan en certa mesura relacionades amb aquesta informació. Aquesta tècnica permet analitzar l'estructura de correlació de les variables i també identificar patrons en les dades, amb tendències, agrupacions i dades anòmales.

Les dades s'organitzen en una matriu com la de la Figura 9. Aquesta matriu està formada per m variables i n mostres per variable.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1m} \\ x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{2m} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot \\ x_{n1} & x_{n2} & \cdot & \cdot & \cdot & x_{nm} \end{bmatrix}$$

Figura 9. Matriu on s'organitzen les dades d'un procés

En realitzar un anàlisi PCA sobre una matriu de dades X , s'està realitzant la modelització definida per l'Equació 3.

$$X = \sum_{a=1}^N t_a p_a^T + E = \hat{X} + E \quad (\text{Eq. 3})$$

En l'expressió anterior p_a són els vectors de pesos que defineixen les direccions principals de màxima variància en l'espai X . Aquestes direccions determinen un subespai de menor dimensió A que l'espai original. Aquesta dimensió s'escull de manera que no existeixi informació important de X en E que és la matriu de residus, i que per tant, representa el soroll. \hat{X} és l'estimació de X a partir del model amb a components. A la Figura 10 es representen gràficament les matrius del model PCA.

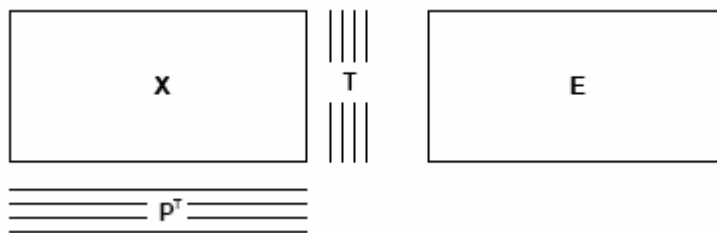


Figura 10. Representació esquemàtica de les matrius del model PCA

Aquest conjunt de vectors ortonormals que defineixen el subespai de menor dimensió A i que determinen les direccions de màxima variància es denominen *loading vectors* o *loadings*. Els *loadings* contenen la informació de com es relacionen les variables entre elles, entenent per variables el conjunt de dades referents a les variables de procés corresponents a una mostra. Aquest conjunt de vectors forma la matriu p on els vectors són ordenats de major a menor segons la variància retinguda en cadascun d'ells. Cada *loading vector* té associat un *eigenvalor* λ_i proporcional a la variància capturada per la direcció corresponent l'hiperplà creat.

Per a una component donada (columna de la matriu p), els pesos defineixen la direcció sobre la que s'estan projectant les observacions i s'està posant de manifest la contribució de les variables originals en la formació d'aquesta component. Les projeccions de les observacions en el subespai de dimensió a definit per les direccions principals són les variables latents, també denominades *scores* (t) i s'obtenen a partir de l'Equació 4 .

$$t_a = Xp_a \quad (\text{Eq. 4})$$

Aquestes variables latents obtingudes, els *scores*, no són observables ni mesurables directament, sinó que s'obtenen per combinació lineal de les variables originals. Els *scores* contenen la informació de com es relacionen les n mostres entre elles. Aquestes variables latents no estan correlacionades ja que són ortogonals i per tant independents entre elles. Cadascun dels elements que formen el vector d'*scores* també s'anomena component principal (PC, de l'anglès *Principal Component*). Segons l'ordre que ocupin en el vector *scores* s'anomenen *PC1, PC2,...*

Per tal de portar a terme la monitorització en aquest nou espai de reduïdes dimensions a partir de les noves variables latents obtingudes es defineixen dos paràmetres estadístics multivariables complementaris. Aquests dos paràmetres es denominen *Q-Statistic* i *Hotelling's T²-statistic*.

El paràmetre *Q-Statistic* o *Q*, també és conegut com l'error quadràtic de predicció (*SPE*, de l'anglès *Squared Prediction Error*). Aquest paràmetre representa la distància perpendicular d'una observació a l'hiperplà definit per les a components principals del model. Aquest paràmetre *Q*, permet detectar les anomalies que no respecten l'estructura de correlació del model. Per a una observació donada, el paràmetre *Q* es calcula a partir de l'Equació 5.

$$Q_i = e_i e_i^T = x_i (I - P_k P_k^T) x_i^T \quad (\text{Eq. 5})$$

On e_i és la i -èssima fila de la matriu d'errors E , P_k és la matriu dels k *loading vectors* retinguts en el model PCA, I és la matriu identitat de mida $(n \times n)$ i x el vector de les variables originals de l'observació.

El paràmetre *Q* indica la bondat d'una mostra alhora d'ajustar-se al model PCA. El valor de *Q* és una mesura de la diferència entre la mostra i la seva projecció en les a components principals retingudes en el model.

Les variables latents s'obtenen per combinació lineal de les variables originals, i en virtut del teorema central del límit es pot considerar que es distribueixen segons una normal

multivariant. Per a la seva monitorització s'utilitza el paràmetre *Hotelling's T²-statistic* o T^2 . El càlcul d'aquest estadístic es realitza a partir de l'equació 6.

$$T^2 = \sum_{a=1}^A \frac{t_a^2}{s_a^2} \quad (\text{Eq. 6})$$

on t_a és la variable latent a-éssima i s_a^2 la seva variància.

Aquest paràmetre permet comprovar, si un vector format per dades de les variables de procés, es projecta sobre l'hiperplà definit per les a components principals del model dins un límits determinats.

Els gràfics per a la monitorització dels estadístics Q i T^2 permeten detectar desviacions respecte les condicions normals definides per les dades de referència utilitzades per a la construcció del model. A la Figura 11 es mostra un exemple dels gràfics de control corresponents als paràmetres Q i T^2 .

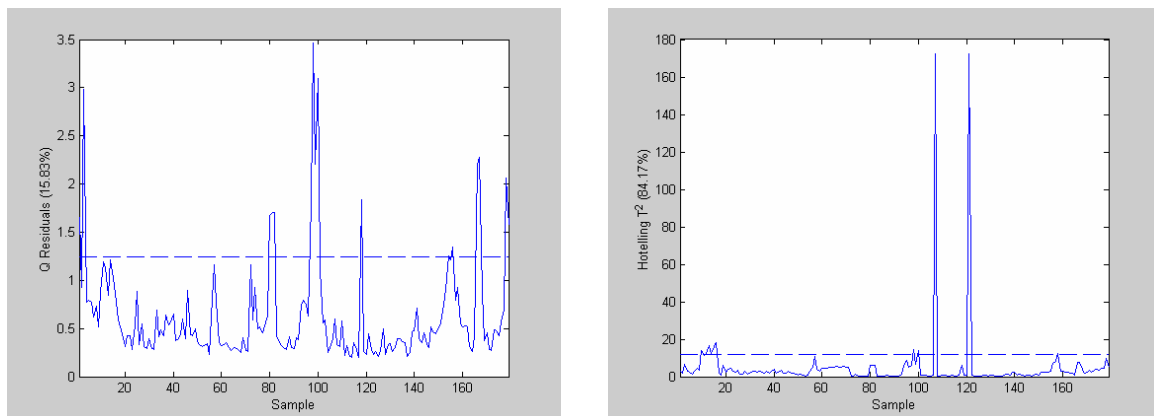


Figura 11. Gràfics de control per als paràmetres Q i T^2

Per tal de portar a terme els gràfics de control de les variables latents que formen el vector *scores* és necessari el càlcul dels límits de control corresponents. Aquests límits es calculen a partir de l'equació 7.

$$t_{a \text{ lim}} = \pm \sqrt{T_{\alpha}^2 \cdot s_a^2} \quad (\text{Eq. 7})$$

On T_{α}^2 és el valor del límit de T^2 per a un nivell de confiança α i s_a^2 és la variància corresponent de t_a .

A la Figura 12 es mostra els gràfics de control corresponents a la primera i a la segona component principal. A la gràfica de l'esquerra es porta a terme el control unitari de la primera component. A la gràfica de la dreta es porta a terme el control simultani de les dues primeres components. En aquest segon cas, el límit resultant és una el·lipse. Cada eix de l'el·lipse creada pren el valor corresponent del límit de la component principal. A partir de la creació d'aquest tipus de límits s'evita la creació de regions suposadament correctes quan en realitat són zones que es troben fora de control estadístic. També es poden realitzar gràfics de control amb la presència de tres components principals. En aquest cas, el límit de control obtingut esdevé una el·lipsoide. Cadascun dels eixos de l'el·lipsoide s'ajusta al valor del límit de la corresponent component principal.

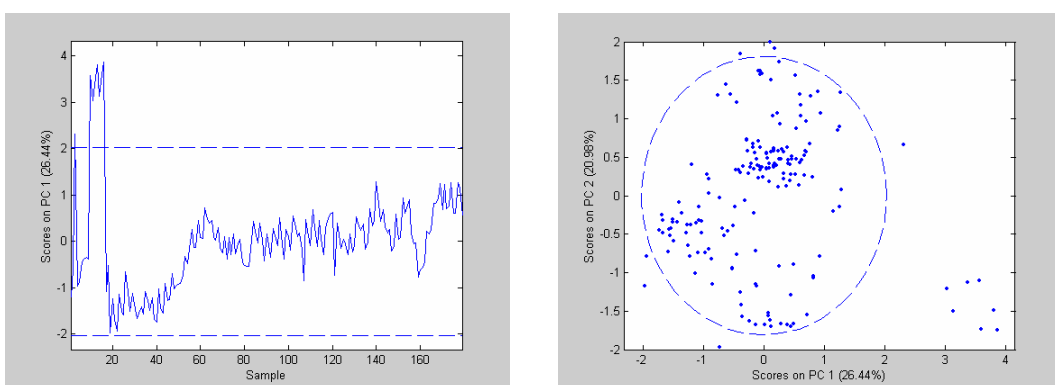


Figura 12. Gràfics de control per a les dues primeres components

El fet que no s'estiguin monitoritzant les variables originals sinó unes variables latents i uns estadístics calculats a partir d'elles i que s'està sintetitzant la major part de la informació que contenen, compliquen de manera important la tasca d'identificar quines variables i en quins instants de temps són les causants de la sortida per fora dels límits de control del procés.

Els gràfics de contribució constitueixen la principal eina per a analitzar les causes d'una sortida de control en els gràfics T^2 i Q. Diversos autors proposen l'ús dels gràfics de contribució com a eina de diagnòstic ja que indiquen com contribueixen cadascunes de les variables de procés en la projecció sobre el model.

En primera instància, cal detectar totes aquelles variables latents o components principals responsables que l'estadístic T^2 es situï fora dels límits de control establerts. Per tal d'efectuar aquesta elecció s'ha d'establir un llindar a partir del qual una component principal ha de ser investigada. Aquest límit es calcula a partir de l'Equació 8.

$$\left(\frac{t_i}{s_i} \right)^2 > (T_\alpha^2)^{1/a} \tag{Eq. 8}$$

Per a cadascuna de les components principals responsables que T^2 estigui fora de control és necessari calcular la contribució de cada una de les variables de procés en cadascun dels instants de temps del procés. Aquesta contribució es calcula a partir de l'equació 9.

$$cont_{a,j} = \frac{t_a}{s_a^2} p_{a,j} (z_j - \mu_j) \quad (\text{Eq. 9})$$

On $(z_j - \mu_j)$ és el vector de dades originals preprocessades.

El valor $cont_{aj}$ obtingut de l'Equació 9 cal igualar-lo a zero en aquells casos en què el valor sigui negatiu (és a dir, sigui de valor oposat al corresponent *score* t_a).

Un cop calculades les contribucions de les variables de procés en tots els instants de temps per cadascuna de les components principals es sumen aquests valors per a cada instant de temps tal i com descriu l'equació 10.

$$CONT_j = \sum_{a=i}^r (cont_{a,j}) \quad (\text{Eq. 10})$$

A la figura 13 es mostra un exemple de gràfiques de contribucions realitzada a partir dels valors obtinguts de les equacions descrites. A la gràfica de l'esquerra s'observa quines variables latents són les responsables del fora de control. A la gràfica de la dreta s'observa quines variables de procés i en quins instants de temps són les responsables que les variables latents es situïn fora de control en projectar-se sobre el model obtingut a partir de les dades de referència.

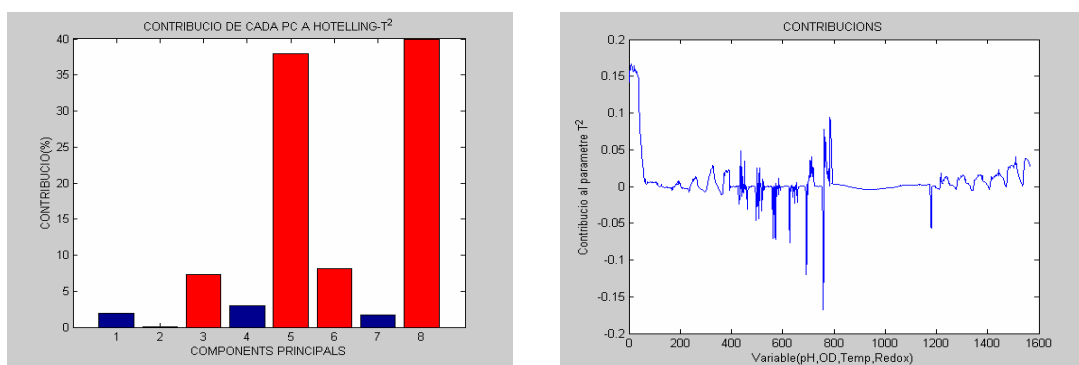


Figura 13. Gràfiques de contribució

4.3.1 Elecció del número de components

Una de les preguntes més importants alhora de crear un nou model PCA és determinar el número de components principals adequades per a la seva creació. En cas d'escollir un número elevat de components comporta que la dimensionalitat del procés no es veu reduïda. Per altra banda, si el número de components escollides és massa petit, certa informació rellevant del procés no serà considerada alhora de crear el model. En aquest cas, part de la informació important present a les dades originals restarà dins la matriu de residus E i no serà presa en consideració alhora de desenvolupar el model.

Per aquesta raó s'han plantejat diversos mètodes per tal d'establir un criteri alhora d'escollir el número de components principals adients que s'han de retenir alhora de crear el model PCA.

Un d'aquests criteris es basa en realitzar la gràfica dels *eigenvalors* de les a primeres components principals com s'observa a la Figura 14 . Els *eigenvalors* són proporcionals a la variància capturada per la direcció corresponent de l'hiperplà creat alhora de realitzar el model. Les components principals s'ordenen de major a menor segons la variància retinguda. Segons aquest criteri, es pot escollir el número adient de components en aquell punt on s'observa un "genoll" a la gràfica. L'aparició d'un "genoll" es correspon amb una disminució considerable del pendent de la recta entre dues components. Aquest canvi de pendent denota que la variància retinguda per la següent component és molt menor en comparació amb la component anterior.

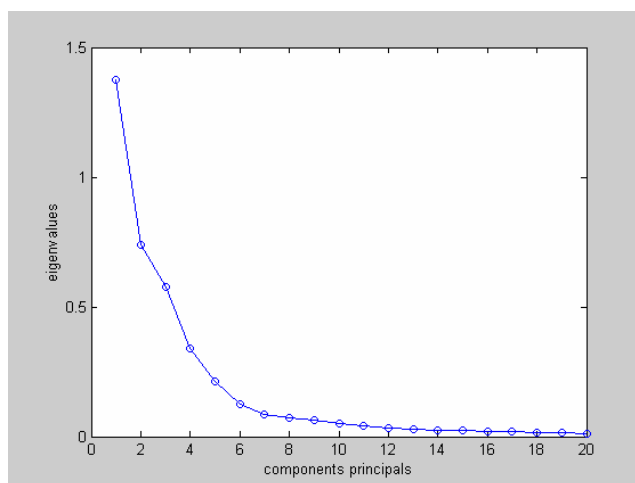


Figura 14. Gràfica dels *eigenvalors* de les a primeres components principals

En aquest cas i seguint aquest criteri, serien tres les components adequades a escollir per a la creació del model. El percentatge de variància retinguda que representen aquestes primeres components principals equival al percentatge de la suma dels seus corresponents *eigenvalors* respecte el total de components principals.

Hi ha altres mètodes plantejats alhora d'escollir el número de components. Un altre criteri plantejat es basa en els gràfics dels *loadings* de les components principals. El número de components escollit es limita a aquell número de components que capturen la informació de totes les variables de procés. En aquest sentit es realitzen els gràfics dels loadings per a les primeres components tal i com s'observa a la Figura 15 i s'observa quines components recullen la informació de cadascuna de les variables de procés.

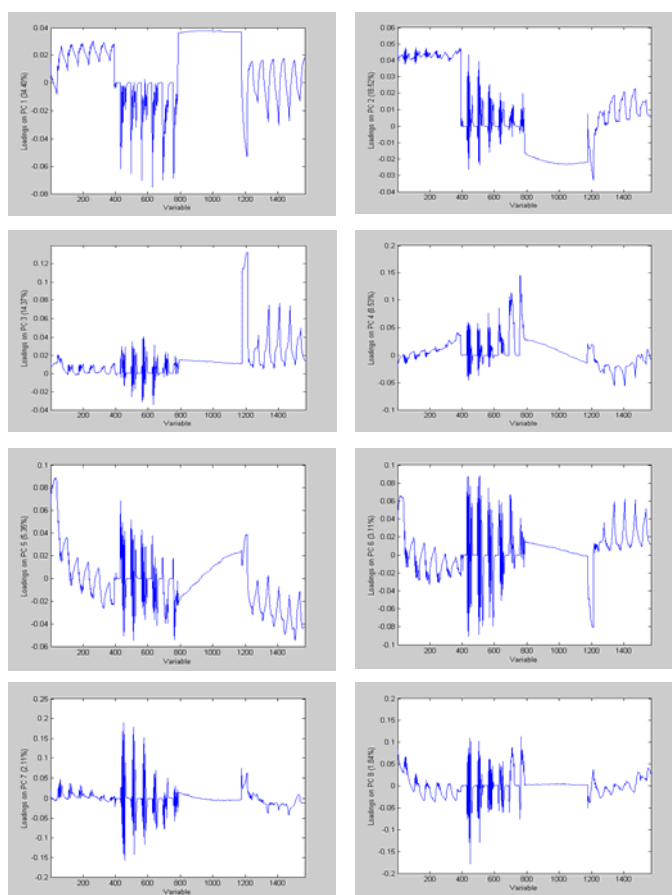


Figura 15. Gràfiques dels loading vectors de les vuit primeres components

En aquestes gràfiques s'observen quatre zones ben diferenciades, cadascuna d'elles corresponents al conjunt de dades referents a les quatre variables de procés. Segons aquest criteri el número de components principals a escollir és 5, ja que no és fins aquesta component que és retinguda la informació de cadascuna de les variables de procés.

En qualsevol criteri establert per a l'elecció del número de components és aconsellable el recolzament de la informació aportada per la taula de variàncies. Aquesta taula recull el valor percentual de la variància retinguda per cadascuna de les variables latents. D'aquesta manera es pot establir un llindar mínim de variància capturada per les components triades per formar el model. Aquest llindar es pot situar al voltant del 80% de variància capturada.

4.4 Control estadístic multivariable per a processos per lots

Un procés discontinu o per lots, és un procés de duració limitada i que consta d'una sèrie de fases; alimentació del reactor, operació sota condicions determinades, descàrrega... tal i com s'ha descrit a l'apartat 3.1 de la memòria . Al llarg d'aquestes etapes les variables de procés presenten una certa evolució temporal o trajectòria.

El conjunt de dades d'un procés per lots s'emmagatzema en una matriu tridimensional com la de la Figura 16. Aquesta matriu tridimensional $\underline{X}(I \times J \times K)$ està formada pels $i=1,2,\dots,I$ lots, les $j=1,2,\dots,J$ variables i els $k=1,2,\dots,K$ instants de temps. D'aquesta manera s'observa que els lots han estat ordenats en el costat vertical de la matriu, les variables de procés han estat organitzades en el costat horitzontal i les seves evolucions en el temps ocupen la tercera dimensió.

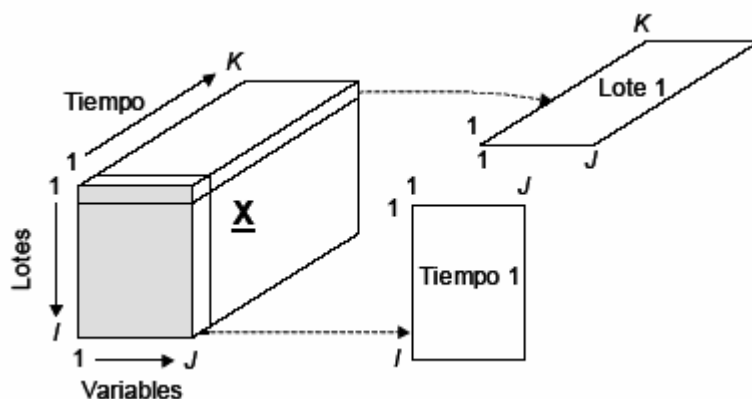


Figura16. Matriu tridimensional (X) de les dades d'un procés per lots

Cada llesca horitzontal ($J \times K$) representa la història en el temps o la trajectòria per a totes les variables d'un lot (i). Cada llesca vertical ($I \times J$) representa els valors de totes les variables per a tots els lots en un interval de temps comú (k).

L'anàlisi de les dades d'un procés per lots emmagatzemades en una estructura com l'anterior, es pot portar a terme mitjançant mètodes que permetin analitzar directament estructures tridimensionals; o bé, es poden desplegar aquestes matrius tridimensionals i convertir-les en matrius bidimensionals. Aquest desdoblament permet l'aplicació de tècniques per a estructures bidimensionals com el PCA, descrit en la secció anterior. Aquesta extensió del control multivariable que permet per tal de permetre l'anàlisi estadístic processos per lots dona lloc al control estadístic multivariable per a processos per lots (a partir d'ara MPCA, de l'anglès Multiway Principal Components Analysis).

4.4.1 Desdoblament de la matriu tridimensional a matriu bidimensional

Per tal d'aplicar tècniques de control estadístic basades en estructures bidimensionals, com el PCA, és necessari desdoblar la matriu tridimensional que conformen les dades obtingudes d'un procés per lots.

Per a una matriu tridimensional hi ha sis possibles maneres de realitzar el desdoblament. Només tres d'aquests desdoblements són independents entre si. Aquests tres desdoblements són els que preserven la direccions de les variables originals: lots, variables i instants de temps. A la Taula 3 es presenten els diferents tipus de desdoblements, i s'indica l'estructura de la matriu resultant així com la direcció que resta inalterada a cada cas.

Tipus	Estructura	Direcció inalterada
A	IK x J	Variables
B	JI x K	Temps
C	IJ x K	Temps
D	I x KJ	Lots
E	I x JK	Lots
F	J x IK	Variables

Taula 3. Tipus de desdoblament de la matriu tridimensional

Alhora d'analitzar i monitoritzar processos per lots, el desdoblament més adient per a portar a terme és el que preserva la direcció original del lot (tipus E). En aquest desdoblament la matriu inicial de dades X es reordena a partir de les seves llesques verticals. Aquestes llesques verticals corresponents a un instant de temps determinat, s'acoblen de manera de

manera successiva i en el sentit horitzontal formant una matriu bidimensional similar a la que s'observa a la Figura 17.

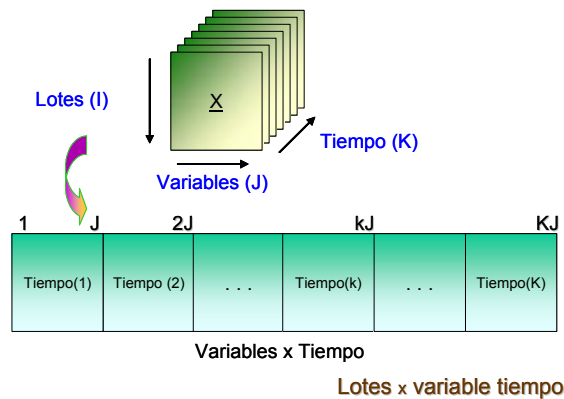


Figura 17.Desdoblament de la matriu X en la direcció del lot

En la composició de la matriu bidimensional mostrada a la Figura 17 existeix l'inconvenient que els elements són una combinació de les variables i el temps. Per tal de facilitar la monitorització i l'estudi de les falles es planteja una petita variació en la reorganització de l'estructura de la matriu X en la qual cadascuna de les variables es manté a través de la duració total del procés, tal i com s'observa a la Figura 18.

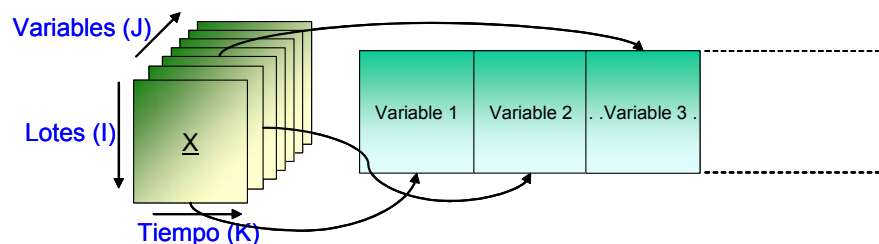


Figura 18.Reorganització de la matriu X mantenint les variables

Amb aquest tipus de desdoblament el lot sencer, és a dir, totes les mesures de les variables durant tots els instants de temps al llarg del procés són considerades com un sol objecte. En aquest desdoblament cada lot és considerat com una mostra amb J (variables de procés) \times K (instants de temps) variables. Aquest fet, permet conèixer la bondat d'un lot, en la seva totalitat, respecte un conjunt de lots de referència. Per altra banda, aquest tipus de reordenació de la matriu tridimensional permet extreure la component no lineal present a les dades, centrant la matriu bidimensional obtinguda (veure apartat 4.4.2).

El desdoblament de la matriu tridimensional que preserva la direcció original de la variable (tipus A), és menys adient alhora de portar a terme la monitorització d'un procés per lots. En aquest tipus de desdoblament la matriu X es reordena a partir de les seves llesques horitzontals acoblant-les de manera successiva en sentit vertical, tal i com es mostra a la Figura 19

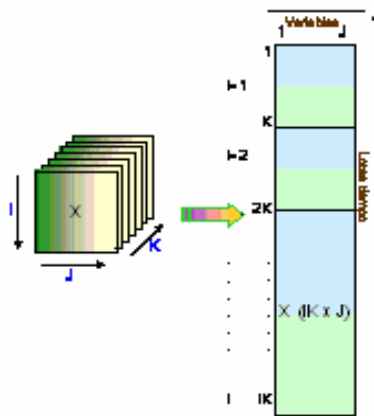


Figura 19 . Desdoblament en el sentit de la variable (tipus A)

En aquest tipus de desdoblament cada instant de temps, o cada punt de la trajectòria de cada lot, es considera com un sol objecte. Cadascuna de les mostres que formen aquesta nova matriu bidimensional està formada per les mesures de les J variables de procés en un instant de temps determinat. Un dels inconvenients que presenta aquest desdoblament és que la no-linealitat del procés resta a la matriu tot i portar a terme un preprocessat de centrat sobre les dades. El fet que aquest tipus de reordenació de la matriu tridimensional no permeti considerar el lot com un sol objecte, comporta que no sigui el mètode més adient per a la monitorització de processos per lots, però diversos autors han suggerit aquest desdoblament com a complement dels inconvenients que presenta el desdoblament en el sentit de lot.

4.4.2 Preprocessat de les dades

En qualsevol tipus d'anàlisi estadístic on es consideri més d'una variable de procés és necessari portar a terme un processat previ a les dades que es disposen del procés. Aquest processat previ o preprocessat es realitza per tal d'extreure d'una manera efectiva la informació rellevant del procés present al conjunt de dades mesurades.

Aquests preprocessats es realitzen per tal d'evitar que certes variables de processos dominin sobre la resta, desvirtuant i fins i tot inutilitzant l'anàlisi estadístic que es porti a terme.

En el cas de la planta pilot i la instal·lació experimental de SBR (veure apartat 3.2 i 3.3), el fet de disposar de quatre variables de procés obliga portar a terme un processat previ del conjunt de dades. En cas de no portar a terme cap tipus de tractament previ, les dades referents a la temperatura tindrien molta menys importància alhora de realitzar l'anàlisi estadístic que les dades referents a la variable Redox. Això és degut a la diferència de variància experimentada per cada variable en ser de diferent naturalesa. Per exemple els valors de la temperatura en un cas determinat poden estar en un rang comprès entre 21°C i 23°C i les dades corresponents al Redox poden estar en un rang entre -65mV i 45mV.

En mètodes estadístics basats en tècniques de reducció de dimensions, com el PCA i el MPCA aquesta necessitat de processat previ de les dades esdevé indispensable.

El preprocessat bàsic de les tècniques PCA i MPCA, així com en el SPC, és el centrat de les dades (en anglès mean center). Aquest preprocessat acostuma a formar part de preprocessats més complexos. En el cas del MPCA i concretament a partir la matriu bidimensional en direcció del lot (tipus E) es calcula la mitjana de cada variable (cada columna) i s'extreu aquest valor al conjunt de dades de la matriu.

En el cas de tenir ordenades les dades del procés en una matriu del tipus E aquest tipus de tractament ens assegura extreure la component no lineal del procés en el moment d'aplicar un centrat previ a les dades. En cas que aquesta matriu estigui ordenada preservant el sentit de la variable (tipus A), tot i aplicar un centrat a les dades, la component no lineal roman present a les dades. Aquest fet es visualitza a la Figura 20, on es realitzen les gràfiques corresponents a la distribució normal de la primera component del vector scores obtinguts a partir de la matriu desdoblada en el sentit de la variable (esquerra) i a partir de la matriu desdoblada en el sentit del lot (dreta). En ambdós casos s'ha portat a terme un centrat previ de les dades.

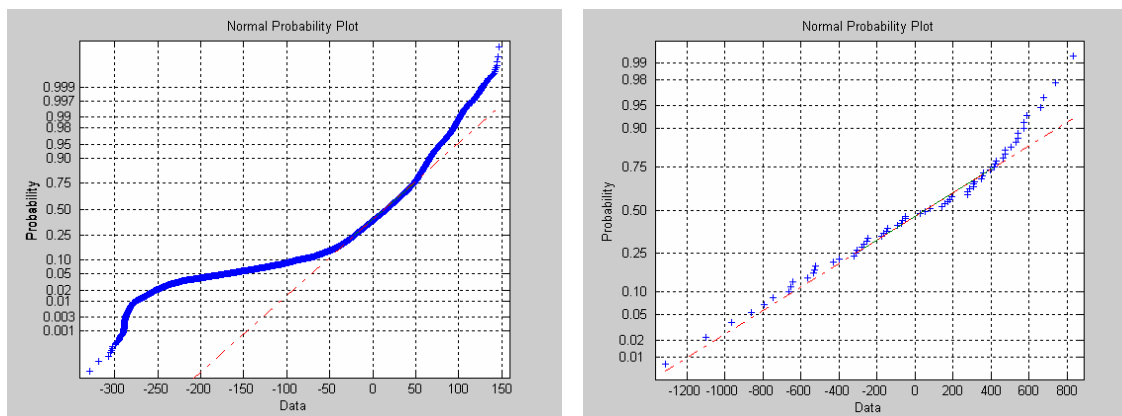


Figura 20. Gràfiques de la distribució normal del vector scores (PC1)

En el cas de tècniques estadístiques multivariables com el PCA i MPCA on es disposen de diverses variables de procés és necessari portar a terme preprocessats més complexes com és el *centrat i escalat* de les dades (a partir d'ara *Autoscale*). Aquest processat previ estandarditza les variables de procés de tal manera que assegura que siguin considerades amb la mateixa importància alhora de realitzar l'anàlisi estadístic. Aquest tractament consta de dos passos simples, un centrat i un escalat. En el cas del MPCA i concretament a partir la matriu bidimensional en direcció del lot (tipus E) en un primer pas es realitza un centrat de les dades tal i com s'ha descrit anteriorment. El segon pas, consisteix en escalar les dades, ja centrades, per la desviació estàndard de la variable (columna) a què pertanyen. D'aquesta manera s'escalen les variables assegurant que les que tenen major variància no dominin per sobre les altres, adquirint major importància.

Hi ha autors que consideren que l'*Autoscale* no assegura una bona estandardització de les variables, és a dir, que no s'aconsegueix que totes les variables són considerades amb la mateixa importància alhora de realitzar l'anàlisi estadístic.

Per aquesta raó s'ha definit un processat de les dades només aplicable a les matriu tridimensionals desdoblades en el sentit del lot (tipus E). Aquest processat es denomina *centrat i escalat per grups* (a partir d'ara *Group Scaling*) i consisteix en un centrat de les dades, tal i com s'ha descrit anteriorment i en escalar a una variància unitària el conjunt de dades referents a una mateixa variable de procés. El valor necessari per realitzar aquest escalat s'obté a partir de l'Equació 11.

$$\sigma_v = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (\text{Eq. 11})$$

On σ_v és la desviació estàndard global utilitzada per realitzar l'escalat i σ_i la desviació estàndard de cada variable (columna) de la matriu bidimensional de tipus E.

S'han definit múltiples mètodes per realitzar el tractament previ de les dades. El preprocessats *Autoscale* i *Group Scaling* són els més utilitzats alhora d'aplicar tècniques estadístiques multivariables com el MPCA.

S'han pres el conjunt de dades de les quatre variables de procés d'un lot sencer durant el seu processat en un SBR i s'han reordenat en un vector de manera anàloga al desdoblament anteriorment esmentat en direcció del lot. A partir d'aquestes dades s'han realitzat les gràfiques de la Figura 21. Aquestes gràfiques corresponen a les dades sense processar (a dalt esquerra) , a les dades processades amb un tractament de centrat (a dalt dreta) , a les dades processades amb un tractament *Autoscale* (a baix esquerra) i a les dades processades amb un tractament *Group Scaling* (a baix dreta). A partir d'aquestes gràfiques s'observa de manera visual l'èstricta necessitat de portar a terme processats complexos de les dades abans de realitzar qualsevol tipus d'anàlisi estadístic. En aquestes gràfiques també s'observen les lleugeres diferències entre l'*Autoscale* i el *Group Scaling*.

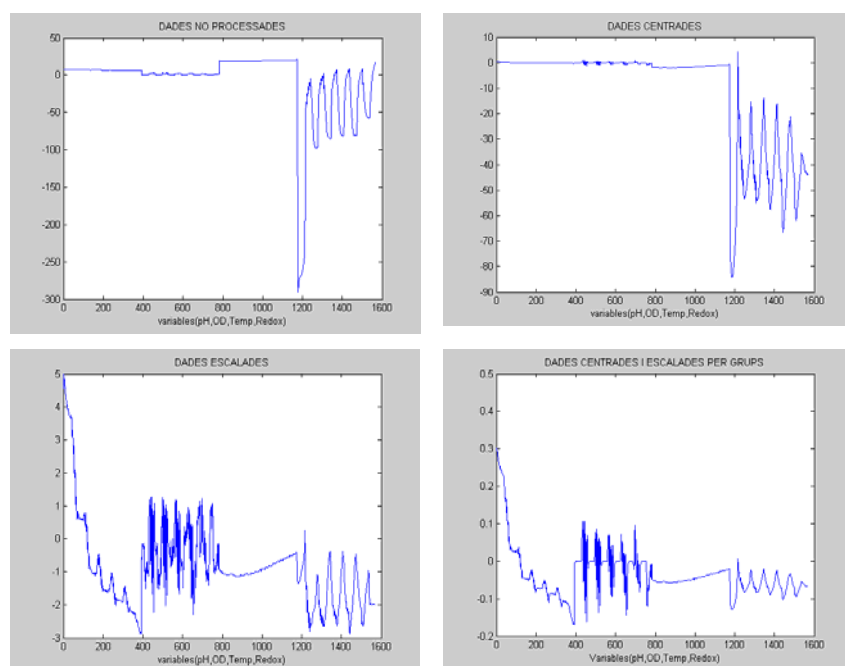


Figura 21. Gràfiques de dades preprocessades

4.5 Anàlisi estadístic multivariable a nivell de lot

L'anàlisi estadístic a nivell de lot és l'anàlisi més adient per a l'anàlisi de processos per lots. Aquest tipus d'anàlisi va ser suggerit per Nomikos i MacGregor (1994). Aquest tipus d'anàlisi permet considerar el lot complet com un sol objecte i conèixer la seva bondat respecte un model de referència creat.

En aquest tipus d'anàlisi es creen uns models a partir de les dades d'uns lots que es consideren de referència i que s'agrupen en funció d'alguna característica comuna. Les dades referents al conjunt de lots de referència són agrupades en una matriu tridimensional i es realitza un desdoblament del tipus E per tal de poder aplicar la tècnica PCA, basada en estructures bidimensionals i crear el model. La matriu formada pels *vectors loading* del model creat determinaran el pla sobre el qual es projectaran les dades dels nous lots.

Aquest tipus d'anàlisi comporta una sèrie d'avantatges i inconvenients. El fet de portar a terme un desdoblament en el sentit del lot permet considerar cadascun dels lots sencers (totes les variables durant tot el procés) com un sol objecte. Aquest fet facilita la comparació d'un lot nou amb els conjunts de lots de referència escollits per crear el model.

Per altra banda, la monitorització a nivell de lot comporta un inconvenient important en el cas que es presenti un lot nou amb manca de dades. Una de les característiques del SBR (veure capítol 3) és la capacitat de poder variar la durada de qualsevol de les etapes que conformen el procés segons les característiques inicials de l'aigua a tractar, segons les especificacions del producte de sortida. En cas que el nou lot no disposi del mateixa estructura i número de dades que el model de referència creat, impossibilita portar a terme de manera directa la monitorització d'aquest nou lot.

Diversos autors han suggerit solucions per tal de poder portar a terme la monitorització en aquestes casos. Es considera com a solució més correcta però alhora inviable pel seu alt cost computacional la creació de múltiples models de referència corresponents a cada instant de temps. Per altra banda s'han plantejat solucions possibles que es basen en la idea d'omplir les dades que manquen del lot nou fins a aconseguir una mateixa estructura de dades que els lots del model de referència. Aquest afegiment de dades es realitza establint diversos supòsits, com que la desviació durant les dades que manquen es manté constant o que la desviació serà nul·la.

Per tal de monitoritzar un nou lot cal escollir en primera instància el model de referència amb el qual es vol comparar. El conjunt de dades que formen el nou lot s'ha de reordenar de la mateixa manera que la matriu tridimensional inicial amb què s'ha creat el model. En el cas de la monitorització a nivell de lot aquest desdoblament és del tipus E. Un cop obtingut el vector de dades reordenats de J (variables) x K (instants de temps) columnes s'aplica el mateix processat previ que s'ha realitzat alhora de crear el model. Aquestes dades es projecten sobre la matriu de *loadings* característica del model i s'obtenen les variables latents que corresponen a les noves dades. A partir d'aquestes variables latents es poden calcular els paràmetres Q i T^2 (veure apartat 4.3).

Un cop obtinguts aquests valors s'introdueixen dins les corresponents caixes de control de les variables latents i dels paràmetres Q i T^2 amb els seus corresponents límits de control.

En aquest tipus d'anàlisi el paràmetre Q ens indica la bondat del nou lot alhora d'ajustar-se a la matriu de variàncies del model de referència. En cas que el nou lot no s'ajusti a aquesta matriu, l'estadístic Q adquireix un valor elevat indicatiu que les dades del lot actual no són de la mateixa naturalesa amb les que s'ha creat el model PCA.

La visualització del valor obtingut del paràmetre T^2 dins les caixes de control corresponent ens indica si la projecció del nou lot sobre l'hiperplà definit per les n components principals del model es troba dins un límits determinats. En cas que el nou lot es situï per sota d'aquests límits es pot considerar que el nou lot és de les mateixes característiques que els lots amb què s'ha creat el model de referència. En el cas que el paràmetre Q d'un nou lot tingui un valor elevat, el paràmetre T^2 no és indicatiu de la bondat del nou lot respecte el model PCA ja que la matriu de variàncies del nou lot és diferent al model i per tant la projecció avaluada pel paràmetre T^2 no és sobre l'hiperplà definit pel model.

En cas que aquests paràmetres es situïn fora de control es pot portar a terme un anàlisi de contribucions (veure apartat 4.3) que permeti passar d'un espai de variables latents no observables a un espai format per les variables de procés. A partir d'aquest anàlisi és possible determinar quines variables de procés i en quin instant de temps són les responsables que el lot es trobi fora dels límits de control.

4.6 Anàlisi estadístic multivariable a nivell d'observació

Com ja s'ha comentat a l'apartat anterior, el principal problema dels mètodes de control per a processos per lots, és la desconeixença de la bonança d'un lot fins la finalització del mateix. En processos on els lots consten d'una llarga durada, com en el cas del SBR, aquest problema s'agreuja.

Per aquesta raó, diversos autors han proposat estratègies de modelització per lots que consten de dos nivells de monitorització del procés, el nivell d'observació i el nivell del lot. D'aquesta manera és possible obtenir informació del lot al llarg del seu procés abans de poder realitzar l'anàlisi a nivell de lot un cop finalitzat.

En el nivell d'observació, en un primer pas es desdobra la matriu tridimensional de les dades en el sentit del lot i es realitza un preprocessat *Autoscale* i d'aquesta manera eliminar la component no lineal present a les dades. En un segon pas, es reorienta la matriu bidimensional preservant el sentit de les variables, desdoblament del tipus A. En aquest tipus de desdoblament cada instant de temps d'un lot constitueix una observació. Un cop realitzat un segon preprocessat de les dades es crea el model en el sentit de la variable. En el cas del nostre SBR, que consta de quatre variables de procés, s'escullen el mateix número de variables latents. D'aquesta manera el valor de la matriu residual és nul i no cal prendre en consideració el valor de l'estadístic Q.

Per tal de crear els gràfics de control que permetin monitoritzar els nous lots, és necessari dividir les quatre variables latents obtingudes en segments corresponents a cada lot. Cadascun d'aquest segments és reorientat de tal manera que per a cada lot s'obtingui un vector $1 \times k$ (on k són els instants de temps de del procés) i agrupat amb els vectors resultants de la resta de lots per a cadascuna de les components principals. Aquesta operació es descriu gràficament a la Figura 22.

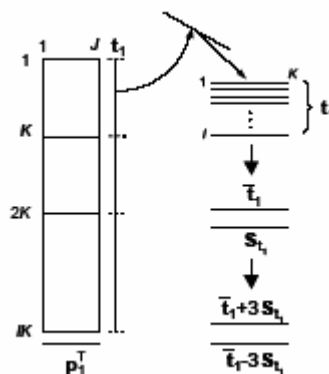


Figura 22. Reordenació de la matriu de scores

Un cop obtingudes aquestes matrius, es calculen la mitjana i la desviació estàndard per cadascun dels instants de temps coincidents amb les columnes de les matrius. Aplicant les equacions corresponents als límits de control per a SPC (veure apartat 4.1) s'obtenen els límits dels gràfics de control emprats per a la monitorització en el nivell d'observació.

Els límits de control corresponents a les variables latents permeten el càlcul dels límits de control corresponents al paràmetre T^2 . Aquests límits es calculen a partir de l'Equació 6.

Per tal d'obtenir les variables latents corresponents a cada nova observació és necessari aplicar sobre aquestes dades els mateixos valors utilitzats per a la creació dels models. Un cop es disposa de les dades tractades es poden projectar sobre la matriu de *loadings* del model realitzat en última instància en sentit de la variable obtenint d'aquesta manera el vector *d'scores* corresponent.

En cas que alguna de les variables latents corresponents a les noves dades es situï fora dels límits de control es pot portar a terme un anàlisi de contribucions (veure apartat 4.1). Aquest anàlisi de contribucions permetrà detectar quines són les variables de procés responsables.

4.7 Programari PCA i MPCA

En el mercat existeixen diversos programes que permeten l'anàlisi estadístic multivariàble de les dades. Aquests programes estan orientats a científics, investigadors, enginyers i a aquells que disposen d'immenses quantitats de dades obtingudes de processos diversos. Aquests programes permeten la transformació d'aquestes dades vastes en informació i d'aquesta manera faciliten la presa ràpida de decisions i amb confiança.

4.7.1 SIMCA-P

Una de les empreses més importants en el desenvolupament d'aquest tipus de programari dedicat a l'anàlisi estadístic multivariàble és la casa UMETRICS. Aquesta empresa disposa d'una família de programes dins aquest camp anomenada *SIMCA*, *The Standard in Multivariate Data Analysis*. Aquesta família de programes està formada pel programa SIMCA-P, SIMCA-P+ i SIMCA-Batch On-Line. A la Figura 23 es mostren les icones corresponents a aquestes programes.

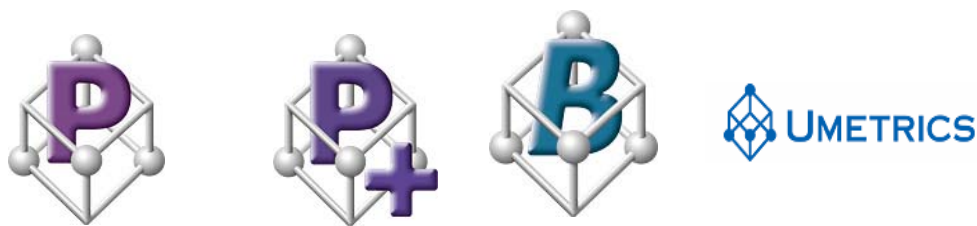


Figura 23. Icones dels programes SIMCA-P, SIMCA-P+ i SBOL de la casa Umetrics

El SIMCA-P és un programa destinat a l'anàlisi estadístic multivariable de processos continus. Permet la importació de dades des de múltiples orígens. Permet la creació de models PCA i PLS en el sentit de la variable. Un cop realitzats els models es pot portar a terme la visualització del mateix o es pot portar a terme la projecció de noves dades. Aquests valors són mostrats en un conjunt de gràfics de control que fan referència a les variables latents creades i als paràmetres estadístics Q i T^2 . També es porten a terme anàlisi de contribucions.

El SIMCA-P+ és una extensió del SIMCA-P que permet realitzar l'anàlisi estadístic multivariable per processos per lots. Permet la creació de models PCA i PLS en el sentit del lot. En aquest programa no s'utilitzen tècniques MPCA, per tant, les dades importades han d'haver estat reordenades en una matriu bidimensional en el sentit del lot. Un cop introduïts els lots es pot especificar les fases que els formen lot per tal realitzar un anàlisi més detallat. El conjunt de gràfiques i resultats és ampli i fan referència a les caixes de control dels vector scores, paràmetres estadístics T^2 i Q . També es porta a terme un detallat anàlisi de contribucions on l'usuari pot especificar l'interval de temps que s'ha de tenir en compte. Aquesta extensió per processos per lots també permet la creació de models a nivell d'observació i la validació d'un conjunt de lots finalitzats en aquest nivell. A la Figura 24 es mostra una la interfície del programa Simca-P+ en un exemple de funcionament.

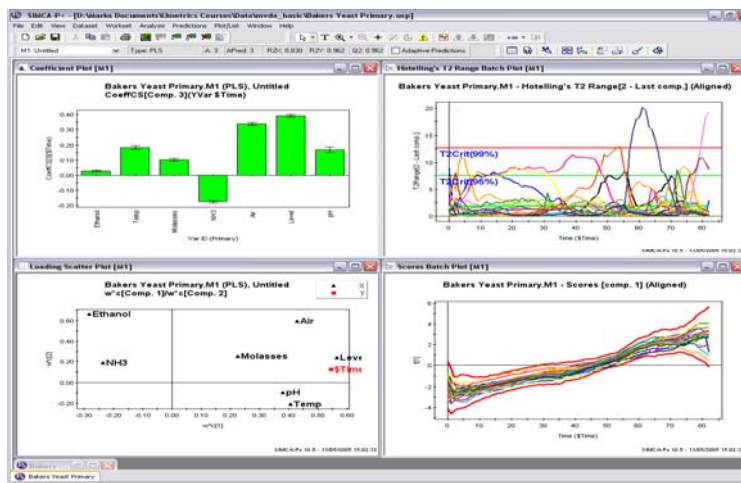


Figura 24 . Interfície del programa SIMCA-P+

El programa SIMCA-Batch On-Line és l'aplicació més recent de la casa UMETRICS. És un programa destinat a realitzar la monitorització en línia d'un procés per lots. Aquest programa permet la connexió a la base de dades on s'actualitzen les dades del procés. Cada instant que s'introdueixen noves dades realitza l'anàlisi estadístic per aquell instant de temps. L'anàlisi estadístic realitzat és en el sentit del lot a nivell d'observació. Aquest programa presenta un conjunt de gràfiques que permeten conèixer l'evolució del lot cada instant de temps. Aquestes gràfiques fan referència a les caixes de control de les variables latents, dels paràmetres T^2 i Q . També es realitza un anàlisi de contribucions per a instants de temps fora de control i es mostren les evolucions temporals de cadascuna de les variables de procés.

La família de programes SIMCA es caracteritza per estar sempre al corrent de les noves idees i desenvolupaments referents a les tècniques multivariables que són actualitzats de manera contínua en els seus programes i per la creació d'unes interfícies d'usuari que faciliten la utilització dels mateixos i la visualització dels resultats.

4.7.2 PLS_Toolbox 3.x

En el mercat existeixen altres aplicacions que permeten portar a terme l'anàlisi estadístic multivariable com la PLS_Toolbox. Aquesta aplicació ha estat implementada per la casa EIGENVECTOR per ser utilitzada com a *toolbox* de MATLAB.

Aquesta *toolbox* recull un conjunt de rutines que fan referència a tècniques basades en l'anàlisi estadístic multivariable que permeten l'exploració de grans quantitats de dades i l'obtenció d'informació rellevant de les mateixes. Aquesta *toolbox* permet construir models que posteriorment seran presos com a referència per a l'anàlisi de noves dades.

Aquesta aplicació va ser creada l'any 1991. En aquell any només va ser implementada per a la realització d'anàlisi estadístic basat en la tècnica multivariable PLS (de l'anglès Partial Least Square), tècnica de la qual rep el nom. Al llarg dels anys es van anar implementant i afegint altres tècniques multivariables com el PCA i el PCR (Principal Components Regression). El nom de l'aplicació es va mantenir.

Aquesta toolbox és utilitzada en la nostra aplicació de LABVIEW com a suport alhora de crear els models PCA i obtenir la informació corresponent dels mateixos. El fet que el conjunt de càlculs necessaris per crear un model PCA sigui complex i l'entorn LABVIEW no sigui molt eficient alhora de realitzar certs càlculs matricials han propiciat la l'establiment d'una comunicació entre LABVIEW i MATLAB i la utilització d'aquesta toolbox. Aquesta aplicació ja s'utilitzava en el grup eXIT (veure apartat 4.8). La versió utilitzada va ser la PLS_Toolbox 3.1. En l'aplicació creada en LABVIEW s'utilitza la versió 3.5 de la toolbox. A la Figura 25 es mostra l'icona corresponent a aquesta aplicació.

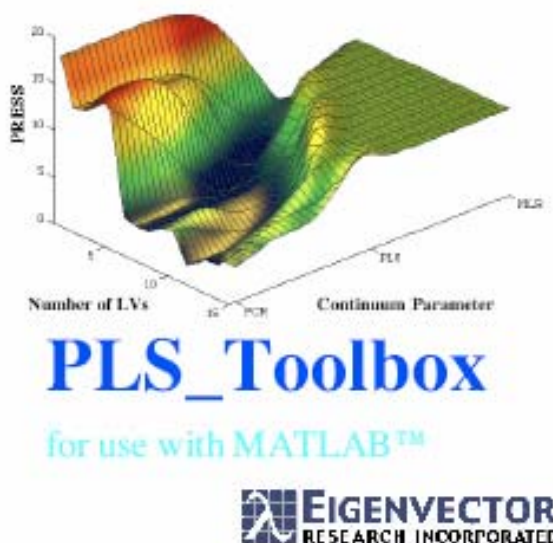


Figura 25. Icona de la PLS_Toolbox de la casa Eigenvector

Pel que respecte a la tècnica multivariable PCA disposa d'una interfície, com la de la Figura 26, a partir de la qual l'usuari de MATLAB pot crear els models PCA i visualitzar els resultats obtinguts. Les dades es poden importar des del *workspace* de MATLAB o des d'algun arxiu que contingui les dades suportat per MATLAB. Les dades importades han de ser en forma de matriu bidimensional. Aquesta matriu ha de ser del tipus E o del tipus A segons si es vol crear un model en el sentit del lot o de la variable.

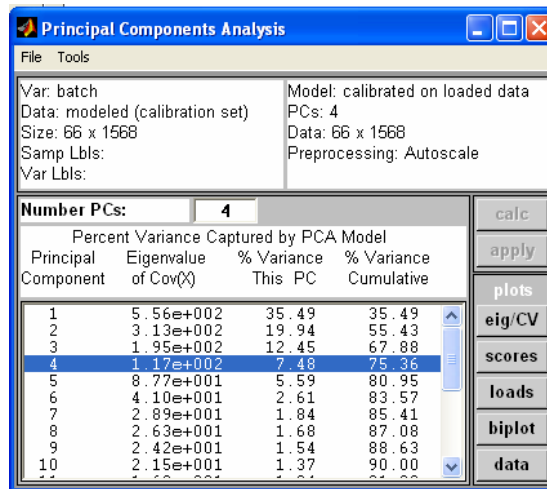


Figura 26. Interfície de la tècnica PCA de la PLS_Toolbox

Un cop importades les dades es poden realitzar diversos processats previs per tal de preparar les dades abans de realitzar l'anàlisi estadístic. Per defecte, la toolbox realitza un processat d' *Autoscale* però aquesta aplicació disposa d'un ventall ampli de processats que es poden realitzar tal i com es mostra a la Figura 27.

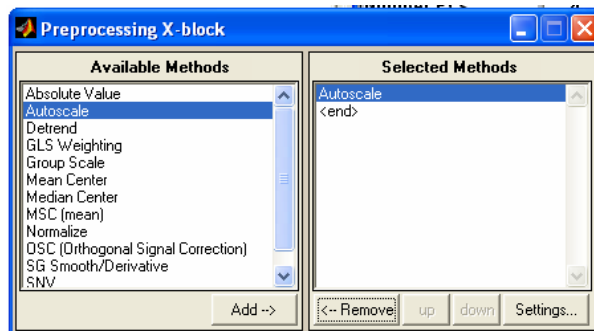


Figura 27. Llistat de possibles preprocessats de les dades

Quan es crea un nou model, la *toolbox* crea un model previ d'un número de components principals elevat i presenta una taula formada pel percentatge de variància retingut per cada component i el seu *eigenvalue* corresponent. D'aquesta manera l'usuari pot decidir quin és el número adient de components per crear el model i conseqüentment crear el model definitiu. Aquest model pot ser guardat en un arxiu d'extensió MATLAB o ser exportat al *workspace* per treballar amb les seves dades.

També és possible treballar amb l'aplicació PLS_Toolbox directament des de la *command window* de MATLAB. En aquest cas, les dades a partir de les quals es vol crear el model han d'estar presents al *workspace* i s'ha d'executar una comanda similar a la de la Figura 28.

I/O: model = pca(x,ncomp,options); %identifies model (calibration step)

Figura 28. Comandes per a crear models PCA mitjançant PLS_Toolbox

En aquesta comanda cal definir com a variable *x* la matriu bidimensional ordenada segons el model que es vulgui crear, la variable *ncomp* corresponent al número de components amb què es vol crear el model i la variable *options* on cal indicar el processat previ que es vol aplicar a les dades.

Aquesta *toolbox* també permet la creació de models a partir de la tècnica multivariable MPCA. Aquesta tècnica és característica dels processos per lots, on les dades són emmagatzemades en una matriu tridimensional, La tècnica MPCA desdobra aquesta matriu creant una matriu bidimensional on pot aplicar-se la tècnica PCA. La *toolbox* no disposa d'interfície gràfica per aquesta tècnica i cal treballar directament des de la *Command Window* utilitzant la mateixa comanda que en la Figura 28 , canviant *pca* per *mpca*.

Utilitzant aquesta comanda la matriu tridimensional es desdobra sempre en una matriu bidimensional del tipus E, en el sentit del lot. Aquest és el sentit més utilitzat i adient alhora de realitzar anàlisis de processos per lots i la *toolbox* no permet realitzar models PCA en el sentit de la variable a partir de la comanda MPCA.

Tant si es crea el model a partir de la comanda PCA com de la comanda MPCA, la *toolbox* crea una variable on s'hi emmagatzema tota la informació relacionada amb el model. Aquesta variable és del tipus *struct array* que conté la informació ordenada en els camps que es presenten a continuació:

Model.

modeltype: 'MPCA' 'PCA',

datasource: s'inclou una matriu amb informació de les dades (autor, data, descripció...)

date: data de creació

time: hora de creació

info: informació addicional del model

loads: cell array amb la matriu de scores(1) i la matriu de loadings(2)

tsqs: cell array amb els valors Hotelling- T^2 per a cada lot

ssqresiduals: cell array amb els valors Q-Residuals per a cada lot

detail: subestructura amb informació addicional del model.

reslim: valor del límit per Q-Residual al 95%

tsqlim: valor del límit per Hotelling T^2 al 95%

label: cell array amb el nom dels lots que formen el model

*preprocessing: estructura informació relacionada amb el preprocessat
description: tipus de processat realitzat*

out: cell array amb les vectors de preprocessat (mitjanes i desviacions estàndard, en cas que correspongui)

userdata: número de grups (en cas de realitzar el preprocessat Group Scaling)

Aquest tipus d'estructura resultant és una tipologia de dades que no es pot transferir directament de MATLAB a LABVIEW ja que només es poden importar un número reduït de tipus de dades. Per aquesta raó des de l'aplicació de LABVIEW s'ha d'escollir a cada moment quina és la informació necessària del model i importar els valors o matrius corresponents.

Per a la visualització de les dades del model i els seus gràfics de control corresponents és necessària la importació de les dades referents a la matriu de vectors *loading*, la matriu de vectors *scores* i els paràmetres Q i T^2 de cadascun dels lots que formen el model. Per a la construcció de les caixes de control també és necessari importar els valors dels límits per a Q i T^2 per a diferents nivells de confiança.

Per a la projecció de dades de nous lots sobre un model de referència és necessària la importació de la matriu de vectors *loadings*, la taula de variàncies on es detalla el valors dels *eigenvalues* i el valors dels límits del paràmetre T^2 per a diferents nivells de confiança.

Per defecte, la *toolbox* només emmagatzema el valor dels límits dels paràmetre Q i T^2 per a un nivell de confiança del 95%. Per tal de poder obtenir el valor dels límits en un rang de

confiança entre el 90% i 99% s'ha afegit part de codi complementari al que disposava implementat la toolbox.

Des d'aquesta aplicació també es permet validar nous lots, és a dir, projectar un conjunt de nous lots sobre un model de referència determinat i comparar la bondat d'aquests lots respecte el model. Aquesta funció es porta a terme mitjançant la comanda de la Figura 29.

>>pred=mpca(mw,model,options)

Figura 29. Comanda per a la validació de nous lots mitjançant la PLS_Toolbox

En aquesta comanda *mw* és una matriu tridimensional formada per les dades del nou lot, *model* és el model de referència i *options* el preprocessat que s'ha realitzat alhora de crear el model.

L'aplicació de LABVIEW no requereix per al seu funcionament d'aquesta comanda de la toolbox. L'aplicació de LABVIEW calcula internament la validació dels lots. D'aquesta manera es millora l'eficiència en no tenir la necessitat d'importar les dades dels nous lots per validar, s'evita la transferència dels resultats obtinguts i la pertinent connexió entre MATLAB i LABVIEW.

Tant pel que respecte a la creació de models com la validació de nous lots la *toolbox* proporciona un conjunt de gràfiques que permeten la visualització dels resultats obtinguts. Aquestes gràfiques fan referència als vectors *scores* com es pot veure a la Figura 30.

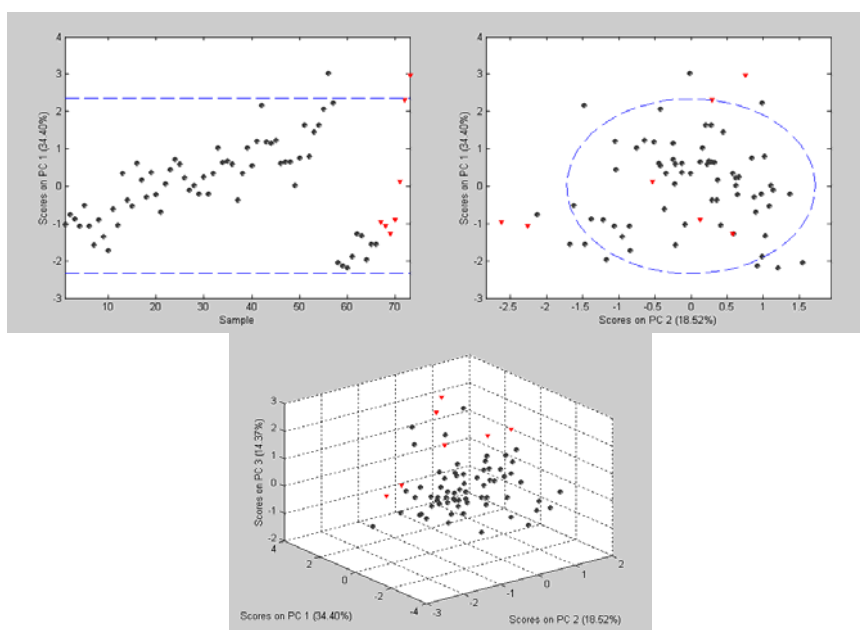


Figura 30. Gràfiques de la PLS_Toolbox referents als vectors scores

Aquests vectors es poden projectar en les tres dimensions. En el cas de les tres dimensions no es poden visualitzar els límits de confiança corresponents.

La *toolbox* també proporciona un conjunt de gràfiques referents als vectors loading, característics del model i a partir dels qual s'obté la projecció dels nous lots sobre el model. Aquestes gràfiques tal i com s'observa en la Figura 31, poden ser projectades en les tres dimensions.

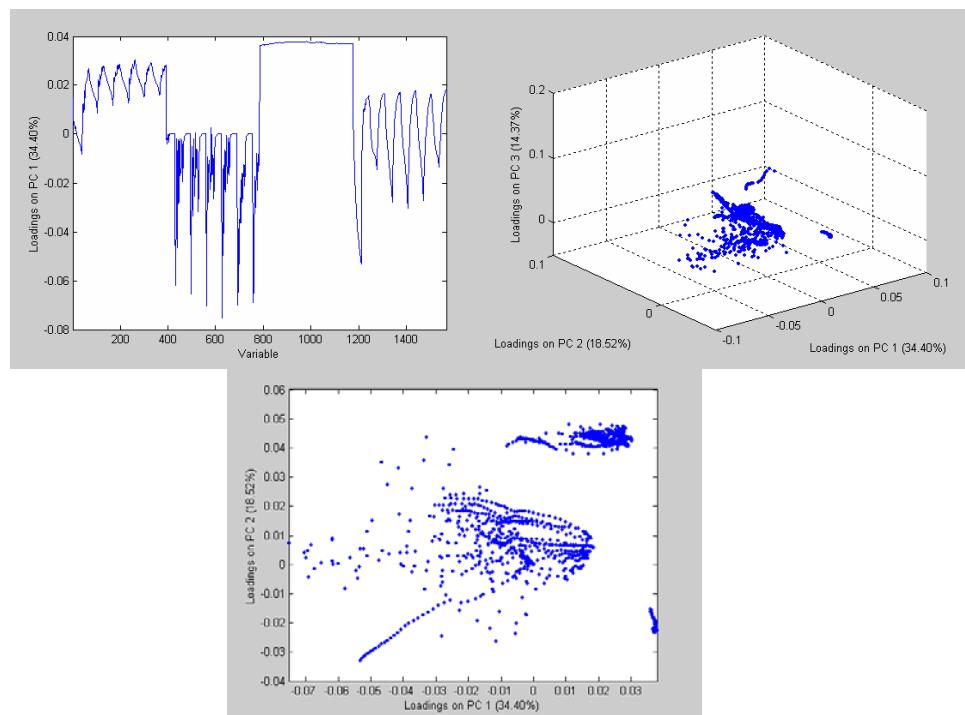


Figura 31. Gràfiques de la PLS_Toolbox referents als vectors loadings

Els paràmetres estadístics multivariables Q i T^2 també són mostrats gràficament dins les corresponents caixes de control com es mostra a la Figura 32. El valor de confiança dels límits presents en aquestes gràfiques es pot modificar, ajustant el valor al nivell de confiança amb que es vulgui treballar (per defecte, 95%).

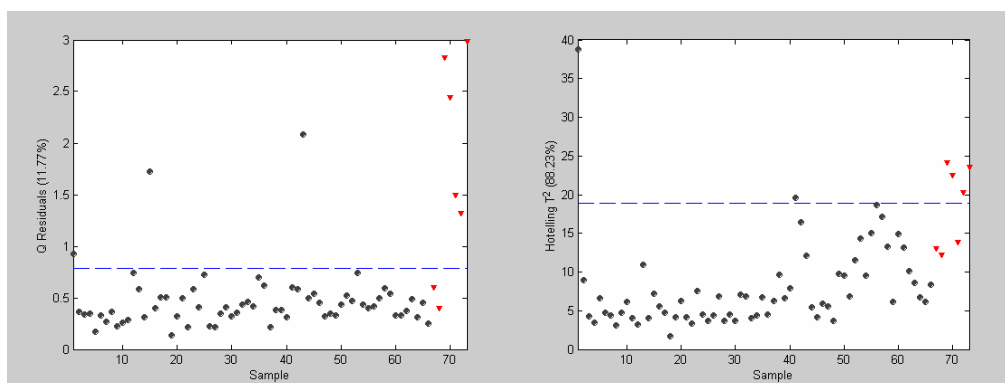


Figura 32. Gràfiques de les caixes de control pels paràmetres Q i T^2

La toolbox ens proporciona també les gràfiques corresponents a la contribució al paràmetre Q i T^2 de cadascuna de les variables per a cada instant de temps, com es pot veure a la Figura 33. Aquest valor només es pot consultar de manera visual i no es pot accedir al seu valor numèric. En el càlcul de contribucions portat a terme per la *toolbox* no es distingeix entre les components responsables que un lot determinat estigui fora de control. El càlcul de contribucions es realitza sumant la contribució de totes les components que formen el model, siguin o no responsables que el lot estigui fora de control. Això pot provocar errors alhora de conèixer quines variables i en quins instants de temps són les vertaderes responsables que el lot no hagi tingut un comportament correcte.

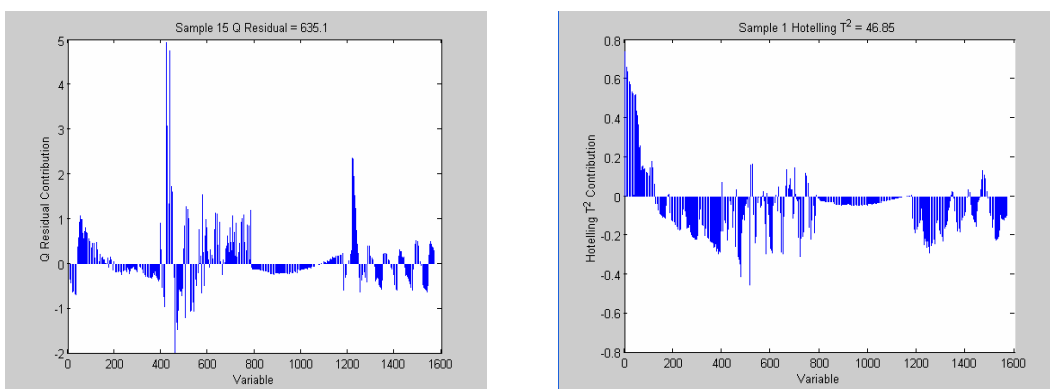


Figura 33. Gràfiques de contribució d'un lot als paràmetres Q i T^2

4.8 Models MPCA realitzats en el grup eXIT

En el grup eXIT s'està treballant en l'estudi de tècniques de control estadístic multivariable pel control de plantes SBR per a la depuració d'aigües residuals.

En aquest sentit es van crear un conjunt de models PCA per tal de poder treballar en aquest camp. Aquests models van ser creats a partir de dades obtingudes de la planta pilot de què disposa el grup eXIT i Lequia (veure apartat 3.2). En el moment en què es van obtenir les dades la planta pilot estava situada en l'EDAR de Cassà de la Selva.

El conjunt de dades utilitzat feia referència a un total de 179 lots, que es corresponien amb un espai temporal de dos mesos. Cadascun d'aquests lots es va emmagatzemar en arxiu d'extensió *txt*. En aquests arxius s'hi inclouïa el valor de les mesures preses referents a les quatre variables de procés (pH,OD,Temp,Redox), mesures que havien estat preses cada cinc segons al llarg del processat del lot.

Primerament, es va portar a terme una classificació dels lots per tal de poder crear una sèries de models, cadascun d'aquests models estaria format per lots amb característiques similars. Aquesta classificació dels lots es va realitzar amb la col·laboració d'investigadors del grup Lequia i va ser realitzada basant-se en paràmetres químics a partir dels corresponents anàlisis realitzats als lots un cop processats. La classificació va permetre dividir el conjunt de lots de què es disposava, en lots excel·lents, lots bons, lots normals, lots formats per aigua de pluja, lots que en ser processats es va produir una fallada de l'ordinador o una fallada de tensió,...

Un cop classificats els lots, es van crear una sèrie d'algoritmes programats en MATLAB que permeten la importació de les dades dels fitxers *txt*, la creació dels models PCA i la validació de lots posteriors.

Un d'aquests algoritmes permet la importació de les dades emmagatzemades en els arxius *txt*. Els arxius corresponents a una mateixa classificació han estat emmagatzemats en una mateixa carpeta. A partir d'aquestes dades es realitza una selecció de les mateixes, de tal manera que es disposi d'una mostra per a cada minut del processat del lot. Aquesta selecció es realitza perquè el procés que es porta a terme a la planta pilot és un procés sobregregat, es prenen mesures en intervals molt reduïts de temps per a un procés lent com són els processos biològics. Un cop realitzada aquesta selecció s'ordenen les dades dels lots per tal d'obtenir una matriu tridimensional formada pels I lots, J variables i K instants de temps. Aquesta matriu s'emmagatzema en una estructura de dades típica de MATLAB anomenada *dataset* on s'hi pot incloure informació addicional com pot ser l'autor de la creació de l'estructura, la data i hora de creació o una breu descripció de les dades.

Per a cadascuna de les classificacions anteriorment descrites es crea una matriu tridimensional.

Un altre dels algorismes realitzats permet crear models PCA a partir de les matrius tridimensionals creades corresponents a cadascuna de les classificacions. Aquests models es creen mitjançant la *toolbox* de control estadístic multivariable PLS_Toolbox utilitzant les comandes referents al MPCA. En aquests algorismes no s'inclou cap eina que permeti el coneixement de certa informació prèvia per tal de decidir el número de components adient alhora de crear el model. Els models que es creen es realitzen en direcció del lot. Per defecte, es realitza un tractament previ *Group Scaling* i el número de components amb què es crea el model són vuit.

L'algorisme restant permet la validació d'un conjunt de lots a partir dels models de referència creats. Aquesta validació també es porta a terme mitjançant la *toolbox* PLS_Toolbox. Aquesta *toolbox* proporciona un conjunt de gràfics i dades que permeten decidir la bondat dels lots escollit respecte el model de referència.

Aquests algorismes realitzats en el grup eXIT han estat una pauta alhora de realitzar l'aplicació de control estadístic en l'entorn LABVIEW. Les característiques de l'aplicació en línia no permeten la utilització "automàtica" d'aquests algorismes a través dels comunicacions que s'estableixen entre LABVIEW i MATLAB.

La importació de les dades es realitza directament des de LABVIEW. Aquestes dades, a diferència dels algorismes anteriors, s'obtenen d'una base de dades MySQL on estan emmagatzemats els valors corresponents a cada lot.

La creació de models també presenta canvis respecte la utilitzada en els algorismes anteriors. Tot i que la creació estricta dels models també es realitza a partir de la *toolbox* PLS_Toolbox, s'han hagut de complementar de manera important alhora de crear els models per tal de portar a terme l'anàlisi a nivell d'observació. La *toolbox* PLS_Toolbox no permet la creació d'aquests models i s'ha hagut d'implementar codi MATLAB que permetés la creació de les caixes de control corresponents.

Tots els càlculs necessaris per a la validació dels lots es realitzen des de la pròpia aplicació. S'ha implementat tot el càlcul matricial en codi LABVIEW. D'aquesta manera no és necessària la utilització de la *toolbox* PLS_Toolbox ni la corresponent connexió entre LABVIEW i MATLAB.

Els models realitzats en el grup eXIT juntament amb el conjunt de lots amb els quals es van formar han estat utilitzats en la realització de l'aplicació de LabVIEW per al control estadístic. A partir d'aquestes dades s'han anat realitzant les comprovacions pertinents de les diferents implementacions dels càlculs estadístics.

5. ENTORN PER A L'ANÀLISI ESTADÍSTIC EN LÍNIA D'UN PROCÉS SBR

5.1 Introducció

L'aplicació realitzada és una eina que permet realitzar l'anàlisi estadístic multivariable d'un procés SBR. Aquest anàlisi estadístic està basat en les tècniques multivariables PCA i MPCA.

Per a la implementació en l'entorn LabVIEW del conjunt d'operacions necessàries per a l'aplicació d'aquestes tècniques s'han estudiat els seus fonaments, càlculs i metodologies utilitzades més recents, tal i com es recull al capítol 4.

L'entorn ha estat implementat mitjançant el programa LabVIEW de la casa National Instruments. L'elecció d'aquest programa per a la creació de l'entorn ha estat condicionada pel fet que l'actualització del mòdul de monitorització de la planta pilot SBR s'està construint amb aquest mateix programa. D'aquesta manera s'intenta evitar possibles incompatibilitats entre programaris i homogeneïtzar connexions externes entre el propi LabVIEW i altres programes.

Un dels propòsits inicials en la realització d'aquest entorn ha estat dotar-lo de la màxima autonomia possible pel que respecte al procés de càlcul, vertader motor de la nostra aplicació. El fet que la família de programes LABVIEW 7.x estigui mancada d'eines potents per a la realització de càlcul intern ha provocat que una sèrie de càlculs s'efectuïn mitjançant un programa extern. El programa utilitzat per a la realització d'aquests càlculs més complexos i laboriosos és Matlab de la casa MathWorks recolzat per la toolbox de control estadístic multivariable PLS_Toolbox 3.5 de la casa Eigenvector (veure apartat 4.7.2).

Aquest entorn està connectat a la base de dades MySQL on s'emmagatzemen les dades referents als lots que es processen en la planta SBR. Gràcies a aquesta connexió es pot conèixer l'estat actual del lot en procés i importar les dades necessàries per tal de portar a terme l'anàlisi estadístic del nivell corresponent.

Aquest entorn és una eina completa per a portar a terme un anàlisi estadístic multivariable complet. Aquesta eina permet realitzar aquest anàlisi en línia amb el procés, coneixent l'evolució a cada instant de temps del lot que s'està processant i coneixent el comportament de l'últim lot finalitzat. Aquesta eina també permet realitzar l'anàlisi estadístic fora de línia, portant a terme l'anàlisi estadístic d'un conjunt de lots ja finalitzats.

Aquesta eina està formada per tres mòduls, cadascun dels quals destinats a una funció. Aquesta eina està formada per un mòdul de monitorització en línia, el mòdul de creació de models i el mòdul de validació de lots finalitzats. Aquest conjunt de mòduls permeten la creació de models de referència a partir dels quals es pot portar a terme l'anàlisi estadístic multivariable de l'últim lot acabat i d'un conjunt escollit de lots ja finalitzats. Aquest anàlisi es realitza a nivell del lot en cas que hagi finalitzat i a nivell d'observació en cas que el lot s'estigui processant.

5.2 Mòdul per a l'anàlisi estadístic en línia

5.2.1 Funció

La funció d'aquest mòdul és realitzar de manera automàtica l'anàlisi estadístic en línia de la planta SBR.

L'anàlisi estadístic portat a terme en aquest mòdul es realitza a nivell d'observació i a nivell del lot. L'anàlisi a nivell d'observació es realitza sobre el lot actual en procés per cada nou instant de temps. L'anàlisi a nivell de lot es realitza un cop finalitzat el lot i els valors obtinguts es visualitzen juntament amb els darrers lots finalitzats.

Aquest mòdul mostra de manera ràpida, senzilla i actualitzada tant la informació del lot en curs com la informació dels darrers lots processats a la planta SBR.

5.2.2 Funcionament

Aquest mòdul treballa de manera automàtica. Aquest mòdul calcula i presenta l'anàlisi estadístic d'un lot determinat al nivell que correspongui en l'instant en què disposa de nova informació rellevant per tal de realitzar el mateix.

Aquest mòdul té establerta una comunicació periòdica amb una freqüència d'un minut amb la base de dades del SBR. En aquesta base de dades s'hi emmagatzemen els valors obtinguts de les sondes connectades al procés. Aquesta connexió permet l'obtenció de certs paràmetres de la base de dades que informen de l'estat del lot actual i permeten en cas de modificació d'algun d'ells donar pas a un anàlisi estadístic al nivell que correspongui a cada cas.

Els paràmetres obtinguts referents a la base de dades a cada connexió són tres: *el número de registres introduïts vàlids per l'anàlisi estadístic, l'etapa en què es troba el procés del lot actual, el número de taules existents a la base de dades.*

A la Figura 34 es detalla el diagrama on es mostra de manera esquemàtica el funcionament d'aquest mòdul. En aquest diagrama es mostra la seqüència d'operacions que regeix el funcionament del mòdul. També s'indiquen les connexions que es realitzen a cada moment amb programes externs, per tal d'importar la informació necessària.

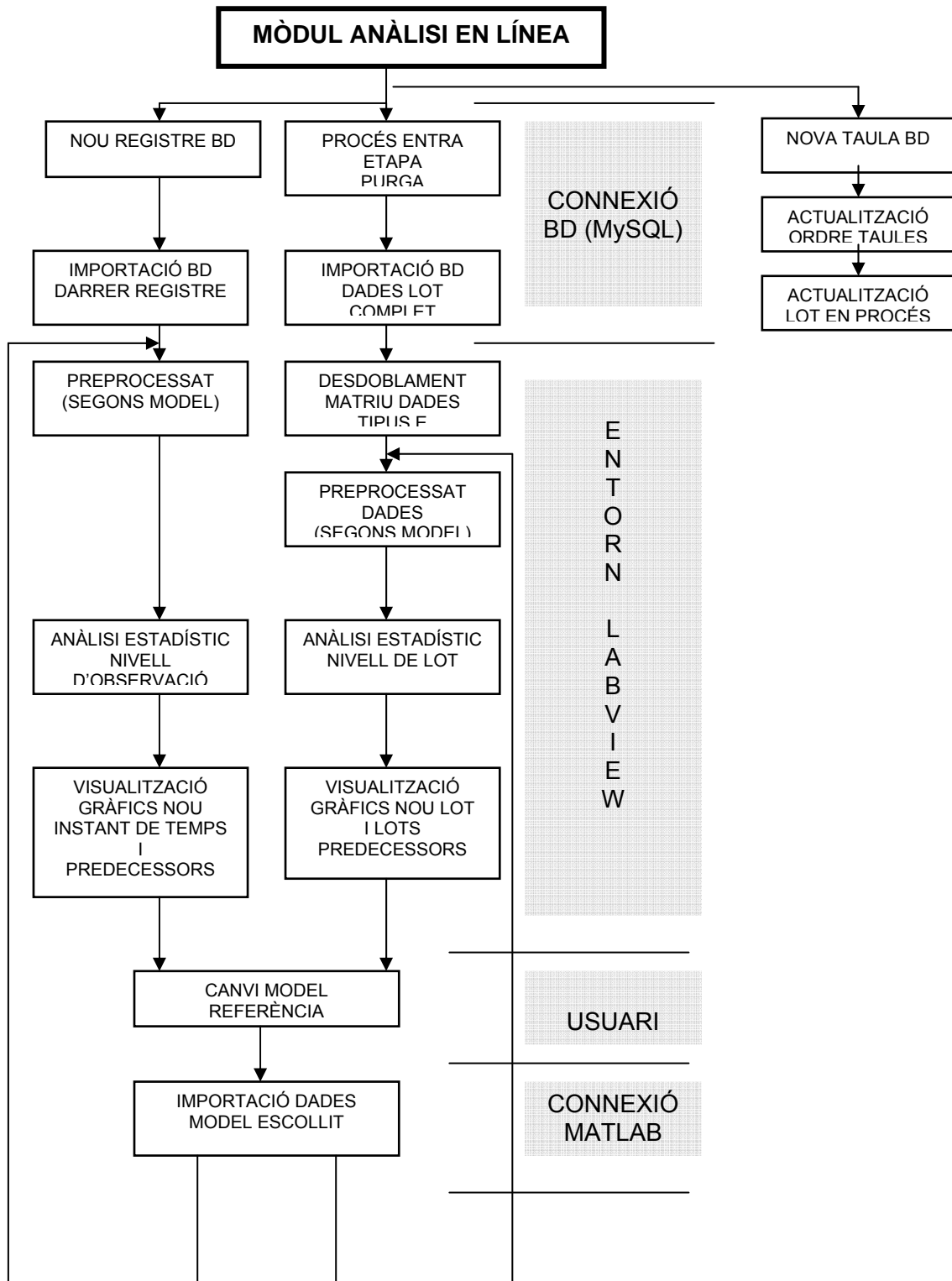


Figura 34. Diagrama de funcionament del mòdul per a l'anàlisi en línia del procés

El fet que *el número de registres introduïts vàlids per l'anàlisi estadístic* (veure apartat 3.5) s'hagi modificat, permet donar pas a la importació de les dades corresponents a aquest darrer instant de temps. Un cop s'han importat aquestes dades referents a les variables de procés i juntament amb les dades que ja es disposen del model de referència indicat es porta a terme la realització de l'anàlisi estadístic a nivell d'observació.

En qualsevol instant durant el processat del lot, l'usuari pot decidir canviar el model de referència amb el qual s'enfronten els lots. En aquest cas, es porta a terme la importació de dades referents al nou model de referència escollit mitjançant la connexió puntual amb MATLAB. També és necessari la importació dels registres vàlids per l'anàlisi estadístic introduïts fins al moment del lot en procés. Seguidament es realitza l'anàlisi estadístic a nivell d'observació fins a l'instant actual.

A partir del càlcul estadístic a nivell d'observació s'obté el vector d' *scores* corresponents a les quatre components principals i el valor de l'estadístic T^2 . Tots aquests valors són visualitzats dins les corresponents caixes de control i d'aquesta manera l'usuari és sabedor de la bonança de lot actual per un instant determinat respecte el model de referència especificat.

En el moment que en una de les connexions puntuals a la base de dades es detecta que el procés del lot actual ha entrat en el procés de purga es pot donar pas a l'anàlisi estadístic del lot actual a nivell de lot. El fet que no siguin necessàries les dades obtingudes del procés referents a les etapes de purga, sedimentació i buidat permet que l'anàlisi del lot s'efectuï aproximadament una hora i mitja abans de la seva finalització.

En aquest instant es pot donar pas a la importació del conjunt de dades referents a les dades del procés i etiquetades com a necessàries del lot actual. Com ja s'ha comentat anteriorment, ja es disposen de les dades del model de referència i per tant es pot portar a terme el càlcul de l'anàlisi estadístic a nivell del lot.

En qualsevol moment l'usuari pot decidir canviar el model de referència per fer l'anàlisi estadístic. En aquest moment s'importen les dades referents al nou model escollit mitjançant la connexió amb MATLAB. En aquest moment ja es pot portar a terme el càlcul de l'anàlisi

estadístic ja que l'aplicació emmagatzema les dades *postprocessades* dels últims tres lots finalitzats.

Els resultats obtinguts del càlcul estadístic a nivell de lot són el vector d'scores amb tantes components principals com tingui el model de referència i el valor dels paràmetres estadístics T^2 i Q . També s'obtenen el percentatge de contribució de cada component al paràmetre T^2 i la contribució per cada variable de procés i cada instant de temps de les components responsables que el paràmetre T^2 es trobi fora de control (en cas que realment es trobi). Tots aquests resultats són visualitzats juntament amb els corresponents als dos (per defecte) lots finalitzats anteriorment.

En el moment que en una de les connexions puntuals es detecti que una nova taula ha estat inicialitzada a la base de dades l'aplicació reordena tot el conjunt per tal de conèixer quina és la taula en què s'emmagatzemen les dades del lot actual en procés. Tot i que de la manera en què estan nomenades les taules, la base de dades les ordena per defecte en ordre cronològic, de més antiga a més recent, és aconsellable portar a terme una ordenació dins l'aplicació.

En aquest moment s'informa que s'ha iniciat un el processat d'un nou lot, indicant el seu nom, la data d'inici del mateix i l'hora provisional que es portarà a terme l'anàlisi estadístic a nivell del lot que difereix de l'hora de finalització del processat complet del lot.

5.3 Mòdul per a la creació de models

5.3.1 Funció

Aquest mòdul permet a l'usuari la creació de nous models de referència. Aquests models seran emmagatzemats i es podran utilitzar en els altres mòduls de l'entorn com a referència alhora de validar altres lots.

Aquests models es poden crear a partir dels lots emmagatzemats a la base de dades de la planta. També es poden crear a partir de dades ja importades anteriorment per tal de crear un model i que ja estiguin ordenades en una matriu tridimensional. Aquest mòdul també permet la visualització de models ja creats per conèixer en detall les seves característiques. En tots els casos, un cop creat el model aquest pot ser editat eliminant qualssevol dels lots que el formen i d'aquesta manera creant un nou model.

5.3.2 Funcionament

Aquest mòdul està destinat a la creació de models per tal de ser utilitzats com a referència alhora de projectar conjunts de lots i portar a terme la seva validació. Aquests models poden ser creats a partir de dades de diversa procedència. L'usuari pot escollir quin tipus de dades requereix per crear un nou model. A la Figura 35 es detalla un diagrama on s'observa la seqüència d'operacions que es porten a terme en aquest mòdul, segons la procedència de les dades requerides. En aquest diagrama també hi figuren les connexions puntuals necessàries a programes externs que permeten la transferència de dades o la realització de certs càlculs.

Aquest mòdul disposa d'una connexió periòdica a la base de dades que actualitza el llistat de lots emmagatzemats. Aquests lots poden ser escollits per l'usuari per tal crear un determinat model. L'elecció d'aquests lots pot realitzar-se de manera unitària es poden seleccionar de manera automàtica tots aquells lots corresponents al mes d'un any determinat.

Cada vegada que es crea un nou model s'emmagatzemen també les dades dels lots a partir de les quals ha estat creat. Aquestes dades estan ordenades de tal manera que formen una matriu tridimensional on cada llesca horitzontal correspon al conjunt de dades que formen un lot. En aquesta matriu tridimensional s'hi emmagatzemen les dades originals de lots sense haver-hi aplicat cap tractament previ.

Aquestes matrius són emmagatzemades cadascuna d'elles en un arxius d'extensió MATLAB. L'usuari pot seleccionar una d'aquestes matrius com a base per a la creació d'un nou model.

Un cop seleccionats els lots o dades a partir de les quals es crearà el model és necessària certa informació, prèvia a la creació del mateix. Aquesta informació fa referència al tractament previ que es vol aplicar a les dades dels lots seleccionats, el nom amb què es vol guardar el model i una descripció (opcional) de les dades, on es pot detallar certa informació característica dels lots escollits.

Un cop introduïda aquesta informació es dona pas a la importació dels lots seleccionats emmagatzemats a la base de dades. En cas d'haver seleccionat un conjunt de lots referents a un mes determinat, el mòdul realitza aquesta selecció a partir de la nomenclatura dels lots

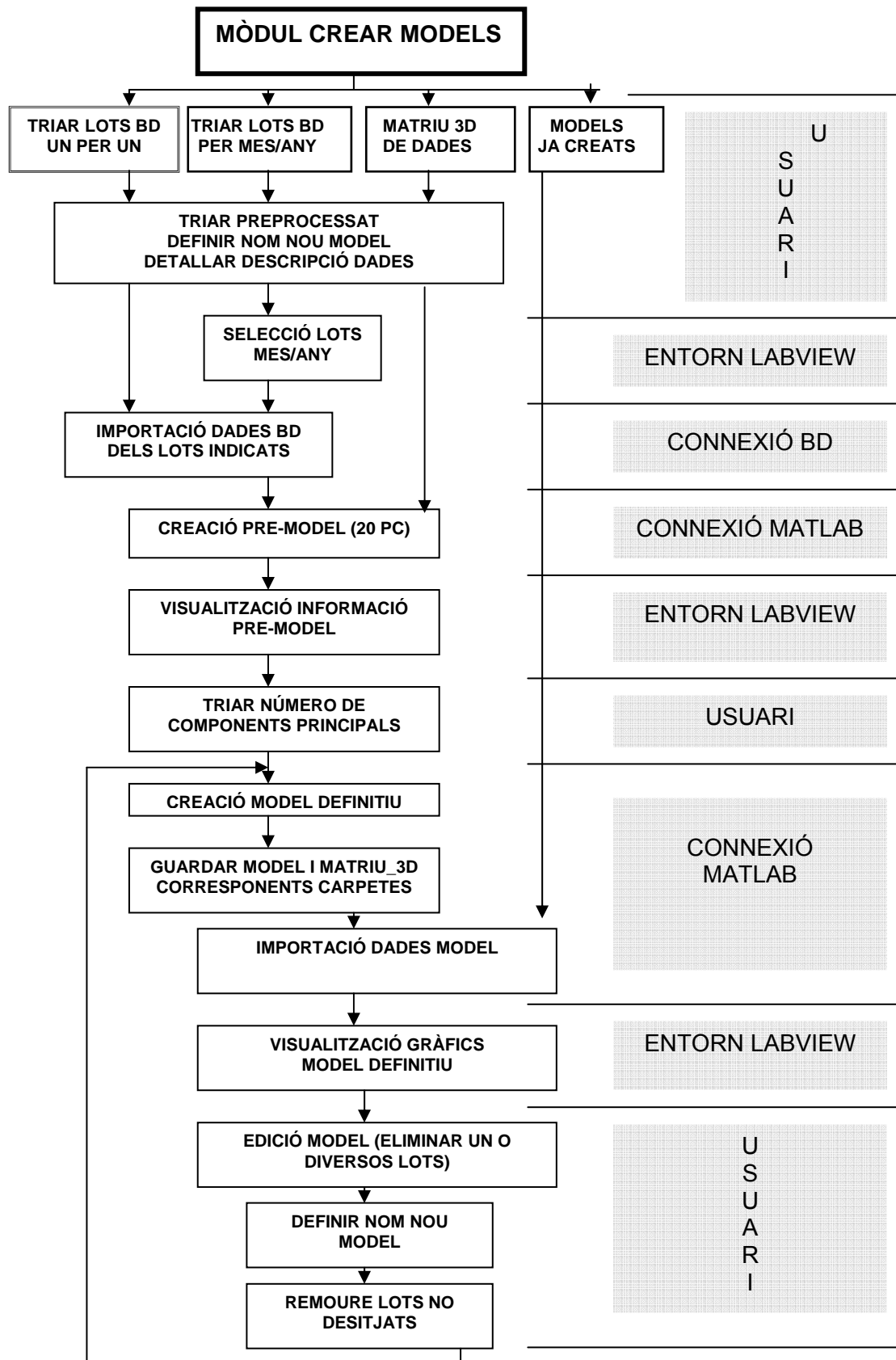


Figura 35. Diagrama de funcionament del mòdul per a la creació de models

on s'especifica la data i hora de l' inici del seu processat. El conjunt de dades que conformen un lot és molt extens, per a la realització de l'anàlisi estadístic es requereix un volum de dades molt menor per a un procés com el que ens ocupa . D'aquesta manera s'importa una selecció de les dades de cada lot que agilitza aquesta operació. Aquest conjunt de dades importades s'ordena en una matriu bidimensional [I (lots) K (instants de temps) x J (variables)] per tal de poder ser transferida a MATLAB ja que el bloc de LABVIEW que estableix la connexió entre els dos entorns no permet la exportació ni importació d'estructures tridimensionals.

En el cas en què les dades escollides per a realitzar el model ja formen part d'una matriu tridimensional no és necessària la connexió amb la base de dades ja que aquestes dades estan emmagatzemades en arxiu MATLAB.

Un cop finalitza la importació de les dades provinents de la base de dades, en cas que sigui necessària, ja es pot donar pas a la creació del model. Aquest model serà creat mitjançant una eina més potent de càlcul de matricial com és MATLAB complementat amb una *toolbox* específica de control estadístic multivariable que s'anomena PLS_Toolbox (veure apartat 4.7.2). Aquesta connexió o comunicació entre els dos entorns LABVIEW i MATLAB es realitza mitjançant el bloc de LABVIEW *Matlab Script* (veure apartat 5.5.5).

Un cop establerta la comunicació amb MATLAB, es crea una matriu tridimensional (I x J x K) a partir de la matriu bidimensional on s'han emmagatzemat les dades importades de la base dades. En cas que s'hagi seleccionat un conjunt de dades ja agrupat s'obté aquesta matriu de l'arxiu on ha estat emmagatzemada.

En tenir agrupades les dades en una estructura tridimensional és possibles utilitzar les comandes de la *PLS_Toolbox* que permeten l'aplicació de tècniques de control estadístic multivariable com MPCA.

Mitjançant aquestes comandes i establint el tractament previ a les dades especificat per l'usuari es crea un model provisional format per vint components. Aquests model es crea per tal de donar certa informació a l'usuari sobre l'estructura del model format per aquest conjunt de dades, com per exemple el tant per cent de variància que captura cadascuna de les direccions que formen el model. D'aquesta manera es pot escollir el número de components principals que formaran el model definitiu (veure apartat 4.3.1).

Aquesta informació referent al model provisional, matriu formada pels vectors *loadings* i taula de variàncies és exportada a l'entorn LabVIEW i presentada en diverses gràfiques i taules a l'usuari. En aquest moment l'usuari ha de decidir el número de components amb què vol crear el model definitiu.

Un cop especificat el número de components escollit, es dona pas a la creació del model definitiu. Aquest model és creat mitjançant els algorismes MPCA de la PLS_Toolbox a partir de la matriu de dades tridimensional anteriorment importada i establint les especificacions de preprocessat i número de components establerts per l'usuari. Aquest model és creat en el sentit del lot per tal de poder portar a terme validacions de conjunts de lots a nivell lot, considerant cada lot com un sol objecte. També es crea un model a partir de les mateixes dades per tal de portar a terme la validació de lots a nivell d'observació, considerant cada instant de temps com un objecte. Aquests dos models són emmagatzemats en un mateix arxiu amb el nom especificat per l'usuari i emmagatzemats en la carpeta de models. La matriu de dades també s'emmagatzema en un arxiu MATLAB.

En aquest punt, s'importen a l'entorn LabVIEW el conjunt de dades que caracteritzen el model per ser visualitzades mitjançant gràfiques i taules en el mòdul.

En cas que l'opció triada per l'usuari sigui visualitzar un model ja creat, s'estableix d'igual manera la connexió entre LabVIEW i MATLAB, s'obté el model emmagatzemat en l'arxiu MATLAB corresponent i s'importen les dades de manera anàloga al cas que el model sigui creat de nou.

En tots els casos anteriorment descrits i un cop visualitzades les dades d'un model determinat l'usuari pot decidir editar un model remonent algun dels lots que el formen. Aquests lots poden ser triats de manera unitària o es pot seleccionar un conjunt de lots que es troben fora dels límits de control dels paràmetres Q i T^2 . Un cop seleccionats aquests lots s'especifica el nom amb què es vol anomenar el nou model. En aquest moment s'estableix una comunicació amb MATLAB on s'obté la matriu de dades amb què s'ha creat el model es vol editar. Els lots no desitjats són extrets d'aquesta matriu creant una nova matriu tridimensional. A partir d'aquesta matriu i mitjançant la *toolbox* de control estadístic es crea un model amb el mateix preprocessat i número de components que el model original. De manera anàloga als altres casos es crea un model per tal de portar a terme la monitorització a nivell de lot i un altre per tal de portar-la a terme a nivell d'observació. Aquests models i la nova matriu tridimensional de dades són emmagatzemades amb el nou nom indicat en les seves corresponents carpetes.

Les dades referents a aquest nou model són importades a l'entorn LABVIEW per ser visualitzades.

En aquest moment l'usuari pot tornar a editar aquest model o crear un nou model a partir de lots emmagatzemats a la base de dades o a partir de lots ja agrupats amb anterioritat.

5.4 Mòdul per a la validació de lots

5.4.1 Funció

L'entorn per a l'anàlisi estadístic multivariable està format per un últim mòdul destinat a la validació de lots. En aquest mòdul l'usuari pot escollir un conjunt de lots emmagatzemats a la base de dades i projectar-los sobre un model de referència. L'usuari pot escollir entre diversos models emmagatzemats en memòria creats a partir del mòdul específic de l'entorn o creats en programes externs com MATLAB. A partir d'aquest mòdul l'usuari pot conèixer la bondat de cadascun dels lots escollits respecte el model escollit. En aquest mòdul l'anàlisi estadístic de cadascun dels lots es realitza exclusivament a nivell de lot. L'anàlisi a nivell d'observació no ens aporta informació rellevant en el cas que els lots hagin finalitzat.

5.4.2 Funcionament

Aquest mòdul disposa d'una connexió periòdica a la base de dades, d'aquesta manera manté actualitzat el llistat de lots emmagatzemats i disponibles per a la seva validació. Aquests lots poden ser projectats en qualsevol dels models que s'hagin creat tant amb el mòdul de creació de model de l'entorn o directament des de MATLAB, en ambdós casos mitjançant la *toolbox* PLS_Toolbox. Aquests models estan guardats en la carpeta corresponent de models.

La seqüència d'operacions que regeixen el funcionament d'aquest mòdul es mostra al diagrama de la Figura 36.

En aquest mòdul l'usuari tria un conjunt de lots emmagatzemats a la base de dades. Aquest conjunt de lots pot estar format per un lot o per n lots. L'usuari ha de triar quin és el model sobre el qual es projectaran als lots, el model que actuarà de referència alhora de validar els lots.

Un cop realitzades aquestes seleccions s'estableix la comunicació amb MATLAB mitjançant les connexions ja esmentades. A partir d'aquesta comunicació s'obté el model emmagatzemat a l'arxiu corresponent de MATLAB i s'importa a l'entorn LABVIEW el conjunt de dades necessàries per tal de validar els lots.

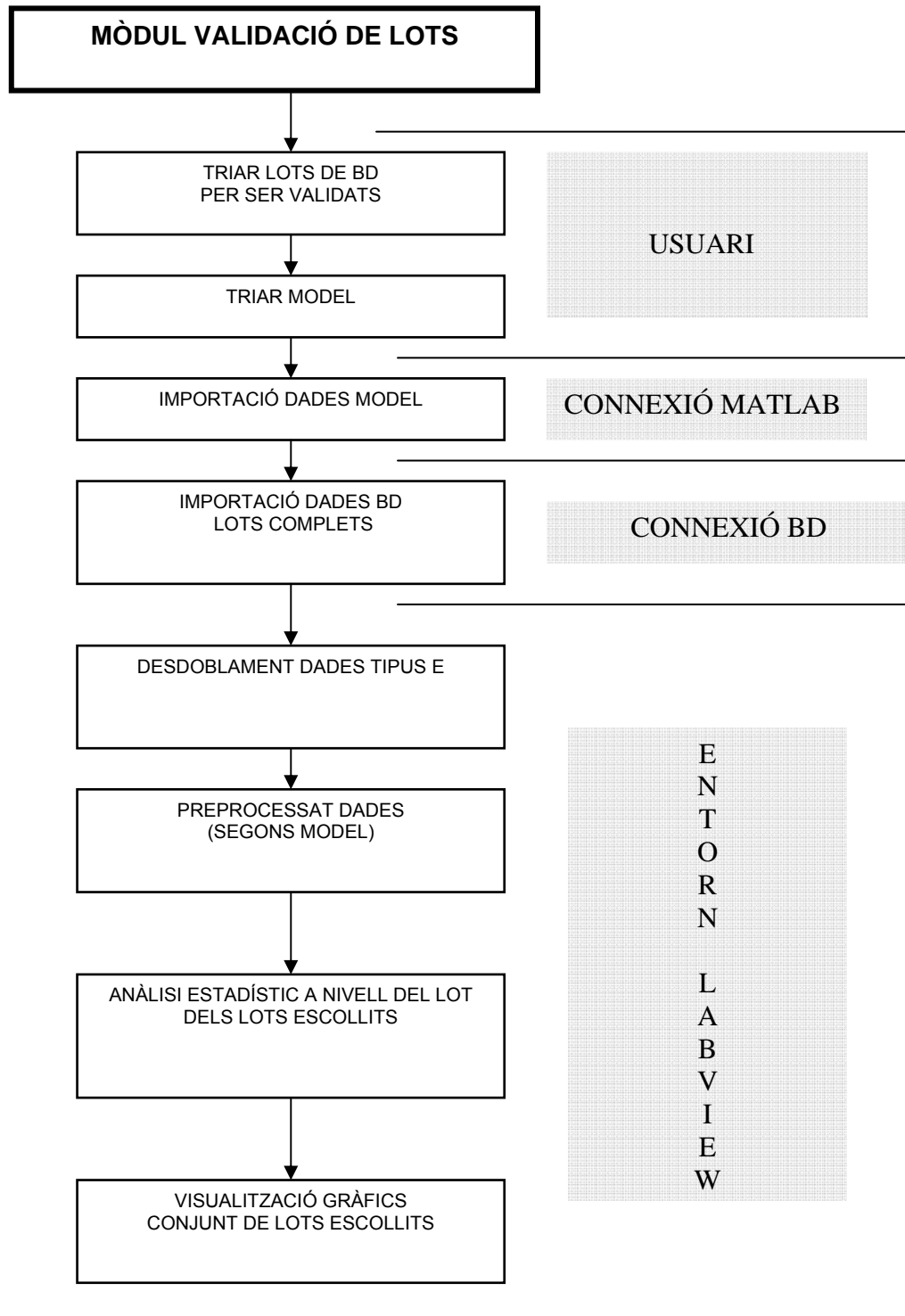


Figura 36. Diagrama funcionament per a la validació de models

Els càlculs corresponents per tal de portar a terme la validació dels lots es realitzen internament des de l'entorn LabVIEW. Tot el codi corresponent a aquests càlculs s'ha implementat en l'entorn LabVIEW per tal de dotar l'entorn de màxima autonomia de càlcul sense dependre de programes externs.

Un cop importades les matrius de dades anteriorment esmentades es porta a terme la importació del conjunt de lots seleccionats. Com s'ha comentat anteriorment no s'importa la totalitat de dades que conformen el processat del lot, només aquelles necessàries per portar a terme un anàlisi estadístic correcte per a aquest tipus de procés.

Els lots s'importen de manera unitària. Les dades corresponents a cada lot arriben a l'entorn LABVIEW en forma de matriu bidimensional (K (instants de temps) x J (Variables)). Per tal de portar a terme l'anàlisi estadístic a nivell de lot, es reordena cadascuna d'aquestes matrius fins a obtenir un vector ($1 \times (K \times J)$). El conjunt de vectors resultants corresponents a cada lot escollit formen una matriu bidimensional del tipus E , matriu adequada per portar a terme un anàlisi a nivell del lot.

Un cop formada aquesta matriu bidimensional el següent pas que es porta a terme és la realització del tractament previ de les dades. Aquest tractament és necessari per homogeneïtzar la importància de les variables de procés presents. El preprocessat a aplicar a aquest conjunt de dades ha de ser el mateix que s'ha aplicat alhora de crear el model i utilitzant els mateixos valors per a realitzar el centrat i escalat corresponents.

Quan s'ha realitzat el preprocessat de les dades corresponents es pot portar a terme l'anàlisi estadístic dels lots a nivell de lot, és a dir, es projecten les dades dels nous lots sobre el model que es pren com a referència.

Els valors obtinguts a partir de l'anàlisi estadístic són visualitzats en les corresponents gràfiques i taules. Aquests valors fan referència al vector *scores*, projecció dels nous lots sobre l'hiperplà del model, així com pels paràmetres Q i T^2 . Els lots que es troben fora dels límits de control del paràmetre $T^2_{90\%}$, se'ls realitza un anàlisi de contribucions. Aquest anàlisi consisteix en determinar quina de les components són responsables que el lot es trobi fora de control. Un cop seleccionades aquestes components, es calcula la contribució d'aquestes al paràmetre T^2 per a cada instant de temps i per a cada variable.

5.5 Funcions específiques

En aquest apartat s'amplien i es detallen un conjunt d'operacions importants que es porten a terme a l'aplicació. Aquestes funcions específiques formen part d'un o més mòduls i són una part més de l'engranatge que fa funcionar l'aplicació. En els apartats anteriors s'ha detallat en quins instants porten a terme la seva funció.

5.5.1 Preprocessat de les dades

El preprocessat és un pas necessari alhora de realitzar qualsevol anàlisi estadístic. Aquesta necessitat s'agreuja quan es porta a terme un control estadístic multivariable, on es disposa de dades de diferents variables. Abans de realitzar l'anàlisi estadístic cal homogeneïtzar la importància de les diverses variables que hi intervenen.

El preprocessat, doncs, com a fase prèvia de l'anàlisi estadístic, és present a tots els mòduls de l'entorn. Aquesta operació es realitza des del propi entorn LabVIEW.

Quan es selecciona un model a qualsevol dels mòduls s'importa un conjunt de dades corresponents a aquest model mitjançant la connexió amb MATLAB. Entre aquestes dades s'inclouen les referents al processat que s'ha portat a terme alhora de crear el model. Aquestes dades es troben emmagatzemades dins l'estructura que crea la *toolbox* PLS_Toolbox quan crea un nou model (veure apartat 4.7.2).

Les dades importades que fan referència al preprocessat són *el tipus de preprocessat* que s'ha realitzat i *els vectors de les mitjanes de les desviacions estàndard* de cada variable (de cada columna). Aquestes dades són suficients en cas que el preprocessat realitzat sigui centrat o *Autoscale*.

En el cas que el preprocessat que s'hagi portat a terme sigui del tipus Group Scaling (només matrius desdoblades en el sentit del lot), aquesta informació no és suficient per realitzar el preprocessat. És necessari el vector de dades utilitzat per escalar les dades. Aquesta informació no està inclosa en l'estructura de dades que forma el model. Per aquesta raó, s'han implementat unes línies codi complementari que permeten l'obtenció d'aquest vector.

El preprocessat de les dades sempre es realitza en l'entorn LABVIEW excepte en el moment de crear un nou model. Quan es crea un nou model la *toolbox* PLS_Toolbox realitza el preprocessat que s'indica a la comanda corresponent i que la *toolbox* ja té implementat.

5.5.2 Creació de models

Un dels propòsits inicials alhora de realitzar aquesta interfície, com ja s'ha comentat anteriorment, era dotar-la de la màxima autonomia de càlcul possible. D'aquesta manera, tot el càlcul corresponent a l'anàlisi estadístic es realitza a l'entorn a excepció de la creació de nous models. La creació de models és conseqüència d'un càlcul complex que es realitza mitjançant la *toolbox* PLS_Toolbox. Aquesta *toolbox* té implementats un conjunt d'algoritmes que permeten la creació de models a partir de tècniques estadístiques multivariables com són el PCA i el MPCA (veure apartat 4.3 i 4.4). Aquesta *toolbox* funciona sota l'entorn MATLAB. La connexió entre LABVIEW i MATLAB es realitza mitjançant el bloc anomenat *Matlab Script* que es descriu a l'apartat .

Com s'ha comentat en l'apartat corresponent al mòdul per a la creació de models permet la creació de models a partir de dades emmagatzemades a la base de dades o dades ja importades que estan guardades en arxius MATLAB.

Pel que respecte a la base de dades, les dades que s'importen ho fan en forma de matriu bidimensional $K(\text{instants de temps}) \times J(\text{variables})$. Els lots s'importen de manera unitària i es concatenen les corresponents matrius de dades fins a formar una matriu bidimensional $(I(\text{lots}) \times K(\text{instants de temps})) \times J(\text{variables})$. Aquesta matriu es transfereix a MATLAB mitjançant el bloc *Matlab Script*. A partir d'aquesta connexió amb MATLAB es reordena aquesta matriu creant una matriu tridimensional $I \times J \times K$. Un cop obtinguda la matriu d'aquest ordre es pot aplicar les comandes MPCA implementades a la *toolbox* PLS_Toolbox, destinades a l'anàlisi estadístic de processos per lots.

També es pot crear un nou model a partir de dades que ja han estat agrupades i emmagatzemades en un arxiu MATLAB. Aquestes dades poden haver estat importades amb anterioritat des de la base de dades per a la realització d'un model, o poden ser dades que no provenen de la base de dades i que han estat ordenades i emmagatzemades directament en un arxiu MATLAB. En ambdós casos, els conjunts de dades ja estan ordenats en una matriu tridimensional que permet l'aplicació de la tècnica MPCA.

A partir de la tècnica MPCA es creen models que serveixen de referència en l'anàlisi estadístic a nivell de lot. En aquest nivell, cada lot és considerat com un sol objecte i d'aquesta manera es podrà conèixer la bondat d'un lot un cop finalitzat respecte el model de referència establert. En crear un model la *toolbox* PLS_Toolbox emmagatzema el conjunt de dades referents a aquest model en una estructura de dades característica de MATLAB

(veure apartat 4.7.2). Aquest tipus d'estructura no pot transferir mitjançant el bloc *Matlab Script* i només s'importen les matrius de dades referents als vector *loadings* i *eigenvalues*, necessaris per a la posterior validació de nous lots i les matrius de dades referents als vectors *scores*, als paràmetres Q i T^2 i al nom de cadascun dels models que formen el lot per a la seva visualització. També s'importen els valors dels límits corresponents als paràmetres Q i T^2 per a la construcció de les corresponents caixes de control i el càlcul des l'entorn LabVIEW dels límits de les diferents variables latents.

Un dels problemes de la monitorització de processos per lots és no poder realitzar l'anàlisi estadístic a nivell de lot fins a la seva finalització. Per tal de fer front a aquest inconvenient és possible realitzar l'anàlisi estadístic a nivell d'observació com a complement a l'anàlisi a nivell de lot i per a conèixer amb anterioritat a la seva finalització el comportament del lot en procés. D'aquesta manera i un cop creat el model a nivell de lot es crea un model que permeti la monitorització a nivell d'observació. La *toolbox* PLS_Toolbox no disposa d'algoritmes que permetin la creació d'aquest tipus de models. Per aquesta raó s'ha implementat un codi MATLAB que en combinació amb la *toolbox* PLS_Toolbox permet obtenir els paràmetres necessaris.

Aquests algoritmes parteixen de la matriu tridimensional ($I \times J \times K$) de dades a partir de la qual s'ha creat el model per a l'anàlisi a nivell de lot. Aquesta matriu es desdobla en el sentit del lot obtenint una matriu bidimensional del tipus $E (I \times JK)$ i s'hi aplica un preprocessat *autoscale*. Aquest processat es realitza per extreure la component no lineal present a les dades del procés. Un cop s'ha realitzat aquest tractament es reordena la matriu fins a obtenir una matriu bidimensional del tipus $A (IK \times J)$ i es crea mitjançant la *toolbox* PLS_toolbox un model a partir de la tècnica PCA per a estructures bidimensionals amb un processat *autoscale*. Un cop creat aquest model s'obté la matriu formada pels vectors *scores*, aquestes dades són reordenades i agrupades per components per tal de crear unes caixes de control que permeten la monitorització a nivell d'observació (veure apartat 4.6).

Les importades a l'entorn LABVIEW per a la monitorització a nivell d'observació són el conjunt de vectors que determinen els límits de control al 95% i 99% per cadascuna de les components principals i per al paràmetre T^2 . En aquest cas el valor del paràmetre Q és nul ja que es crea un model amb tantes components principals com variables de procés i per tant no hi resta informació a la matriu residual. Per la monitorització posterior és necessari importar els vectors corresponents al processat *Autoscale* realitzat tant a nivell de lot com a nivell de variable, així com la matriu formada pels vectors *loadings* i la matriu d'*eigenvalues*.

Els models creats a nivell del lot i a nivell d'observació són guardats en un mateix arxiu Matlab i emmagatzemats a la carpeta designada de models. Les dades utilitzades per a la seva creació són guardades en forma de matriu tridimensional en un arxiu Matlab dins la carpeta designada de dades.

5.5.3 Projeció de nous lots

Com ja s'ha comentat en apartats anteriors l'aplicació realitza tot el càlcul estadístic a l'entorn LABVIEW a excepció del càlcul de nous models. Aquesta autonomia de càlcul permet major eficiència ja que l'aplicació requereix menys recurs de màquina en necessitar menys comunicacions amb programes externs.

D'aquesta manera, tots els càlculs necessaris per a la monitorització tant a nivell de lot com a nivell d'observació han estat implementats en codi LabVIEW.

Els valors necessaris per a la monitorització d'un nou lot a nivell de lot són el seu vector d'*scores* i els paràmetres estadístics multivariables Q i T^2 . Aquests valors ens permeten conèixer si el nou lot s'ajusta a l'hiperplà creat pel model de referència i si la seva projecció es troba dins dels límits establerts. En cas que aquesta projecció es trobi fora d'aquests límits també és necessari portar a terme un anàlisi de contribucions del lot, on s'observa la contribució al paràmetre T^2 per a cada instant de temps i per a cada variable.

Per al càlcul de tots aquests valors són necessàries les matrius que formen els vectors loadings i la matriu d'*eigenvalues* del model que s'estableixi de referència.

Tot el conjunt d'operacions necessàries per a l'obtenció dels valors corresponents a la projecció d'un lot ha seguit una mateixa metodologia. Un cop estudiades aquestes operacions han estat implementades en codi Matlab per comprovar la seva correcció. Un cop realitzat aquest pas s'han implementat aquests algorismes de càlcul en codi LabVIEW on seguidament s'han comparat amb els valors obtinguts amb els algorismes Matlab. El conjunt d'operacions necessàries per l'obtenció de la projecció d'un lot es detallen a l'apartat 4.3.

5.5.4 Connexió amb la base de dades MySQL

Per a la realització de pràcticament la totalitat de les accions de la nostra aplicació és necessària la importació de les dades de les variables de procés que descriuen el funcionament de la planta SBR.

El control de monitorització de la planta té com una de les seves funcions bàsiques l'emmagatzemament en una base de dades de les mesures realitzades a les variables de procés. En el cas del nostre SBR són quatre les variables de procés: pH, Temperatura, Redox i Oxigen Dissolt (veure capítol 3). La base de dades on s'emmagatzemen aquestes dades corre sobre el programa MySQL.

Per tal d'establir la comunicació entre LABVIEW i la base de dades MySQL s'utilitza la toolbox *LABVIEW Database Connectivity Toolset*.

Aquesta *toolkit* proporciona un conjunt de blocs de llenguatge d'alt nivell que s'afegeixen als ja existents a la paleta de controls que permeten la connexió i totes aquelles accions necessàries a realitzar entre LABVIEW i una base de dades.

Per tal que aquesta connexió pugui ser establerta cal instal·lar a l'ordinador un controlador compatible amb ODBC (Open Database Connectivity) que permeti actuar de pont entre LabVIEW i la base de dades amb la qual es vol establir la connexió. En el nostre cas s'utilitza el controlador *MySQL ODBC 3.51 Driver*. Un cop instal·lat aquest controlador s'ha de configurar la base de dades relacionant-la amb el controlador abans esmentat. Per tal de configurar la base de dades cal adreçar-se a Origen de dades(ODBC) situat al Panell de Control clicar a la pestanya *DSN d'usuari* i introduint les dades corresponents de la base de dades.

Mitjançant la utilització d'aquesta toolkit disposem de nous blocs que ens faciliten la connexió a la base de dades MySQL a partir de blocs com els de la Figura 37 que efectuen la connexió i desconnexió entre LABVIEW i la base de dades corresponent tot introduint tres paràmetres necessaris: el nom de la base de dades, el nom de l'usuari i la contrasenya corresponent.

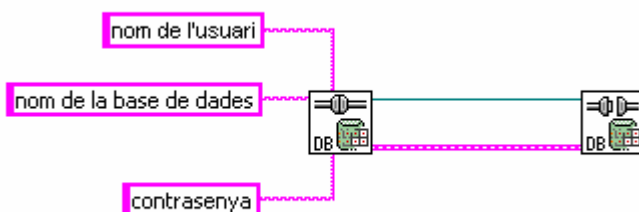


Figura 37. Blocs que permeten la connexió i desconnexió amb una base de dades

El procés de la nostra planta, és un procés sobrerregistrat. El mòdul de monitorització emmagatzema les dades corresponents a les quatre variables de procés cada cinc segons al llarg de tot el cicle, cicle que té una duració de vuit hores. Les dades necessàries per portar a terme l'anàlisi estadístic es troben sobre el 10% del total de dades emmagatzemades a la base de dades.

Per tal d'importar un número de dades s'estableix un filtre referenciat a la columna que actua com senyal anomenada *flag* (columna formada per 0's i 1's) que permet importar només una dada per minut de cadascunes de les variables de procés al llarg del cicle.

La nostra aplicació requereix d'informació complementària per tal de donar pas a l'execució de diverses accions que s'han de realitzar de manera automàtica.

Algunes d'aquestes informacions fan referència a conèixer el moment en el qual el procés ha entrat en l'etapa de purga a partir de la qual es pot aplicar l'anàlisi estadístic a nivell del lot. Per tal d'obtenir aquesta informació s'utilitza combinació de blocs de que disposa la *toolkit* específica de base de dades de LabVIEW . D'aquesta manera es poden importar les dades de la columna anomenada ID_Etapa on les cel·les de la qual fan referència a l'etapa de Purga. En el moment en què la matriu receptora d'aquesta importació sigui inicialitzada, és a dir, rebí el primer valor, es pot donar llum verda a la importació de les dades de les variables i el seu conseqüent anàlisi estadístic. Altres informacions sobre l'estat de la base de dades és el coneixement del número de files que han estat registrades i que compleixen la condició de ser importades (és a dir, columna *flag* igual a 1). En el moment en què aquest número variï es dona pas automàticament a l'anàlisi estadístic a nivell d'observació per a un nou instant de temps. Per tal d'obtenir aquesta informació és necessari la utilització de blocs de nivell avançat que ens permeten executar comandes SQL des de LABVIEW.

El fet que els cicles que es van succeint al llarg del funcionament de la planta SBR estiguin regits per un mateix esquema temporal i que l'emmagatzemament de dades segueixi un mateix criteri, permet que una connexió puntual entre LABVIEW i la BD cada 60 segons sigui suficient per actualitzar les dades i estadístics de la nostra aplicació. Tot i que seguint aquest criteri, estrictament es podria considerar que es realitza un control fora de línia, aquest fet ens allibera d'una connexió permanent que consumiria molt recurs de màquina.

5.5.5 Connexió amb MATLAB

Tal i com s'ha comentat anteriorment la nostra aplicació requereix d'un programa extern per tal de realitzar certes accions que comporten un procés de càlcul laboriós com són la creació

dels models basats en mètodes estadístics multivariables. Aquest programa extern és la PLS_Toolbox 3.5 (veure apartat 4.7.2) que corre sota el marc del programa de càlcul matricial MATLAB.

El programa LABVIEW disposa d'un bloc anomenat *Matlab Script*, com el que s'observa a la Figura 38 que permet establir una comunicació directa amb el programa Matlab a través del *Matlab Script Server*. És necessari tenir instal·lat el Matlab amb les toolbox corresponents a l'ordinador.

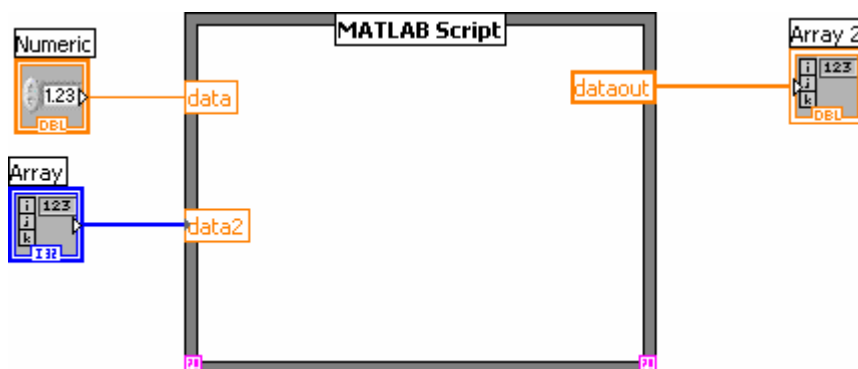


Figura 38 .MATLAB Script, bloc de LABVIEW que permet la connexió amb MATLAB

Mitjançant aquest bloc es poden realitzar les mateixes operacions que des de la finestra de la *Command Window*; des d'accedir a qualsevol tipus de dades emmagatzemades en un format suportat per Matlab fins a executar una sèrie de funcions descrites en codi Matlab.

Des d'aquest bloc es poden enviar dades des de LABVIEW fins a Matlab i a la inversa. Totes aquestes dades han de ser etiquetades segons la seva tipologia: un número real, un vector de números reals, una matriu de números reals, un *string*. Aquest és un requeriment de LABVIEW, un programa que es caracteritza per la seva exigència en la tipificació de les dades que s'utilitzen en la totalitat de qualsevol aplicació. En cas que en el moment de l'execució d'aquest bloc algunes de les dades no es corresponguin amb la tipologia etiquetada amb anterioritat, el bloc provocarà un error i el codi en el seu conjunt no podrà ser executat.

El fet que LABVIEW només permeti la transferència de vuit tipus de dades pot ocasionar complicacions alhora de treballar amb un tipus no permès. En la nostra aplicació és necessària la importació d'una matriu de caràcters que es corresponen amb els noms dels lots amb els quals es forma un determinat model, per tal d'importar aquests noms és necessària una conversió de la matriu de caràcters a números reals i un cop a LABVIEW, realitzar el pas invers convertint aquesta matriu en una matriu de caràcters altra vegada.

Aquest bloc tampoc permet la transferència d'estructures complexes com són les matrius tridimensionals, característiques dels processos per lots i utilitzades en la tècnica multivariable MPCA. Per a transferir aquest tipus de dades s'ha de formar una matriu bidimensional i un cop a l'entorn MATLAB reordenar-la fins a crear la matriu tridimensional corresponent.

Una de les característiques i avantatges de programar mitjançant un arxiu d'extensió *m* (característic de MATLAB) és la seva senzilla i ràpida depuració d'errors. El bloc MATLAB Script de LABVIEW, però, no permet aquesta depuració i aquest fet esdevé un dels inconvenients majors alhora d'utilitzar aquest bloc ja que s'ha d'exportar el codi a MATLAB, convertir-lo en un arxiu amb extensió *m* i executar-lo un cop inicialitzades les variables necessàries.

5.6 Interfícies d'usuari

Aquest entorn per a l'anàlisi estadístic multivariable està format per tres mòduls. La funció de cadascun d'aquests mòduls dins aquesta eina així com el seu funcionament intern han estat descrits en seccions anteriors.

A continuació es mostren les interfícies d'usuari corresponents a cadascun dels mòduls. Cada mòdul li pertany una interfície pròpia a partir de la qual l'usuari pot interactuar i obtenir les informacions necessàries referents a qualsevol aspecte de l'anàlisi estadístic de la planta.

La interfície principal és la que correspon al mòdul d'anàlisi en línia. Aquesta interfície és sempre visible ja que el mòdul corresponent es troba en funcionament permanent. A partir d'aquesta interfície és possible accedir als altres dos mòduls que formen l'entorn, el destinat a la creació de nous models i el destinat a la validació de lots ja finalitzats. Les interfícies d'aquests dos mòduls només són visibles quan són requerides per l'usuari.

5.6.1 Interfície del mòdul per a l'anàlisi en línia

El mòdul per a l'anàlisi en línia es troba en funcionament permanent. Mitjançant una connexió periòdica amb la base de dades detecta l'estat en què es troba el processat del lot actual. Aquest mòdul realitza l'anàlisi estadístic multivariable a nivell d'observació per a cada instant de temps del lot en procés. En el moment que finalitza el procés del lot, el mòdul realitza de manera automàtica l'anàlisi estadístic a nivell del lot. A la Figura 39 es mostra la interfície del mòdul destinat a l'anàlisi en línia de la planta SBR.

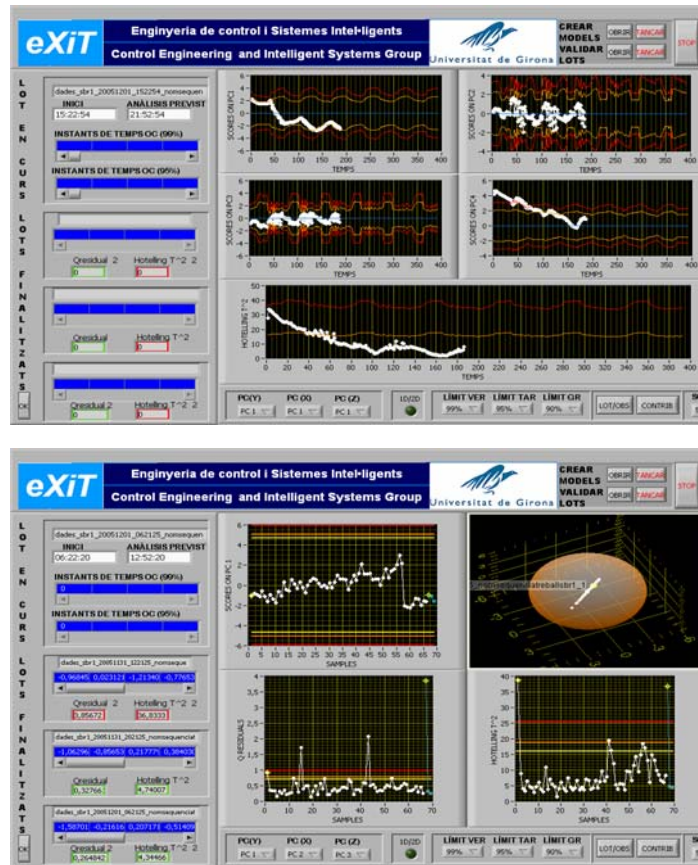
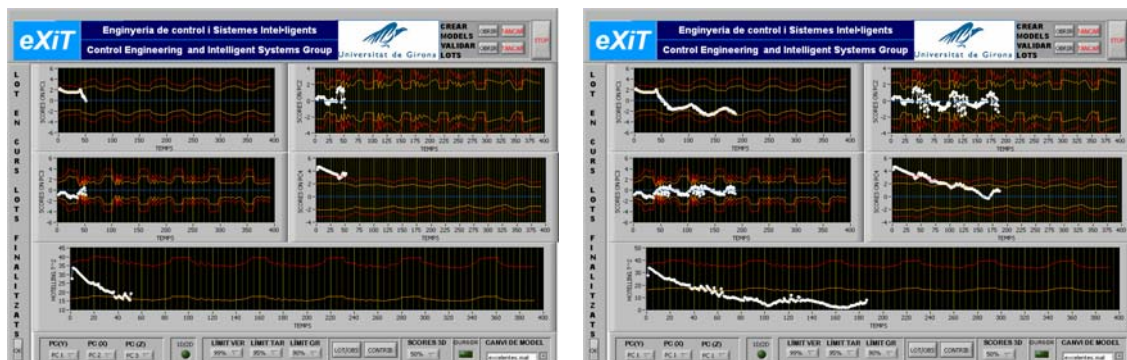


Figura39. Interfície del mòdul per a l'anàlisi en línia

En aquest mòdul es realitza l'anàlisi estadístic dels lots a dos nivells: a nivell d'observació i a nivell de lot.

Pel que respecte a l'anàlisi estadístic a nivell d'observació, el mòdul realitza de manera automàtica aquest anàlisi per a cada instant de temps del procés. Un cop realitzat aquests càlculs l'interfície mostra aquests valors dins les corresponents caixes de control. A la Figura 40 s'observa l'anàlisi d'un lot a nivell d'observació per a diversos instants al llarg del procés.



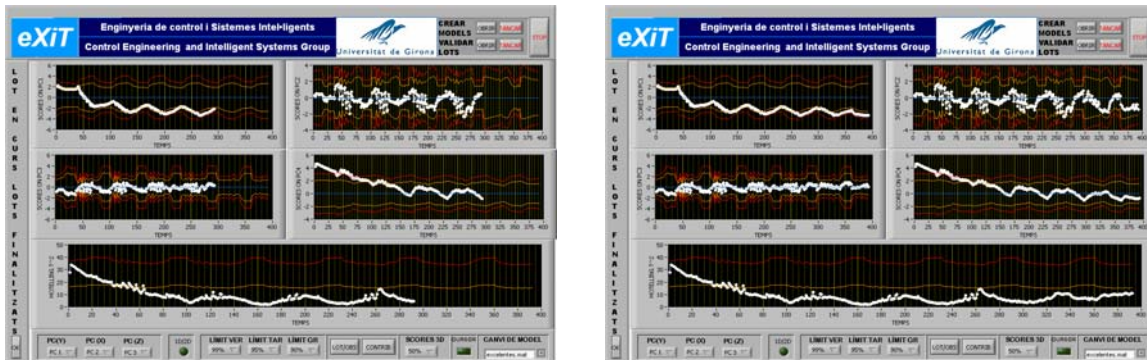


Figura 40. Interfície en diversos instants en l'anàlisi a nivell d'observació

Aquestes gràfiques mostren les caixes de control corresponents a les components principals que formen el model i la caixa de control corresponent del paràmetre estadístic T^2 . Aquestes caixes de control s'obtenen a partir de la creació d'un model de referència en el sentit d'observació. Aquest model es crea amb tantes variables components principals com variables de procés es disposen del model. D'aquesta manera no és necessari monitoritzar el paràmetre Q ja que el seu valor és nul. Per a cada nou instant de temps es calculen els valors corresponents a la projecció sobre el model de referència i són visualitzats en les caixes de control. A la Figura 41 es mostra una ampliació de la gràfica corresponent a la primera component i al paràmetre T^2 . A partir d'aquestes gràfiques és fàcilment observable si el comportament del lot es troba dins o fora de control per a cada instant de temps.

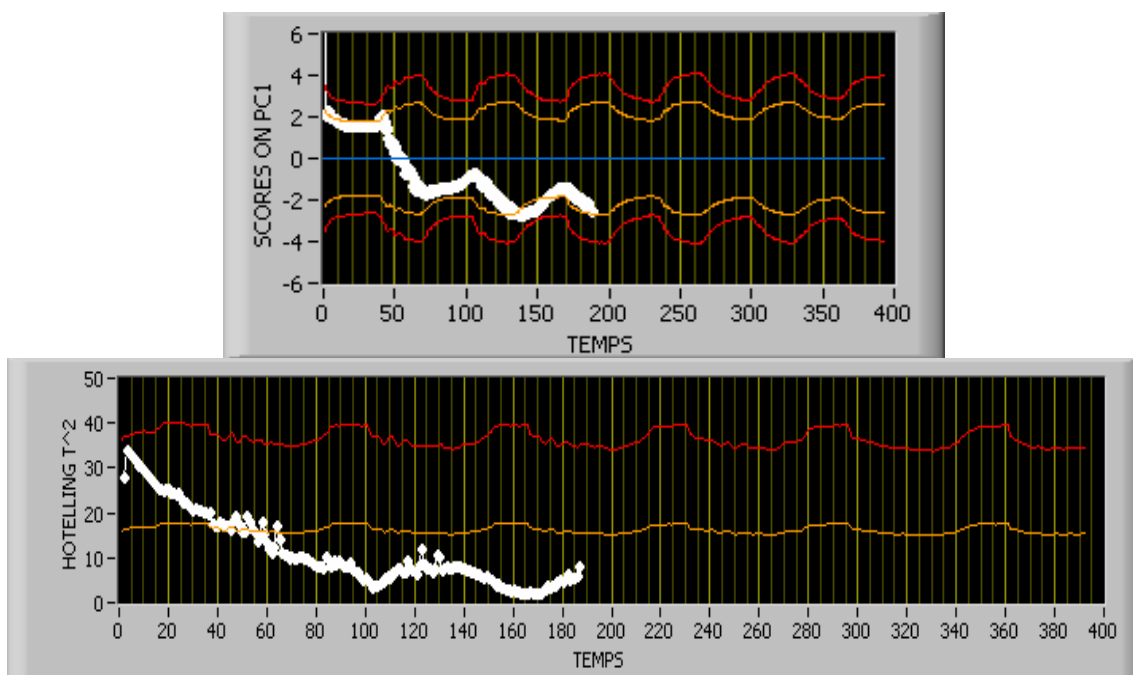


Figura 41 . Gràfiques de dues de les caixes de control a nivell d'observació

Quan un punt o instant de temps es situa fora dels límits de control és necessari portar a terme un anàlisi de contribucions. L'anàlisi de contribucions permet passar d'un espai amb presència de variables latents (components principals) a un espai amb presència de les variables de procés.

D'aquesta manera és possible conèixer quines variables de procés són les responsables que un punt corresponent a un instant de temps del procés es situï fora de control. Per a cada instant de temps l'usuari pot conèixer la contribució de les variables de procés al paràmetre T^2 a partir de la contribució de cada variable de procés a cadascuna de les quatre components principals del model.

Aquest anàlisi de contribucions es realitza per a tots els instants de temps, es trobin dins o fora dels límits de control. Les gràfiques s'escalen d'igual manera, d'aquesta manera l'usuari pot distingir de manera clara quina és la component responsable que en aquell instant de temps el procés es trobi fora de control. A la Figura 42, s'observen les gràfiques de barres que mostren la contribució percentual de cada variable de procés en el càlcul de cada component principal o variable latent.

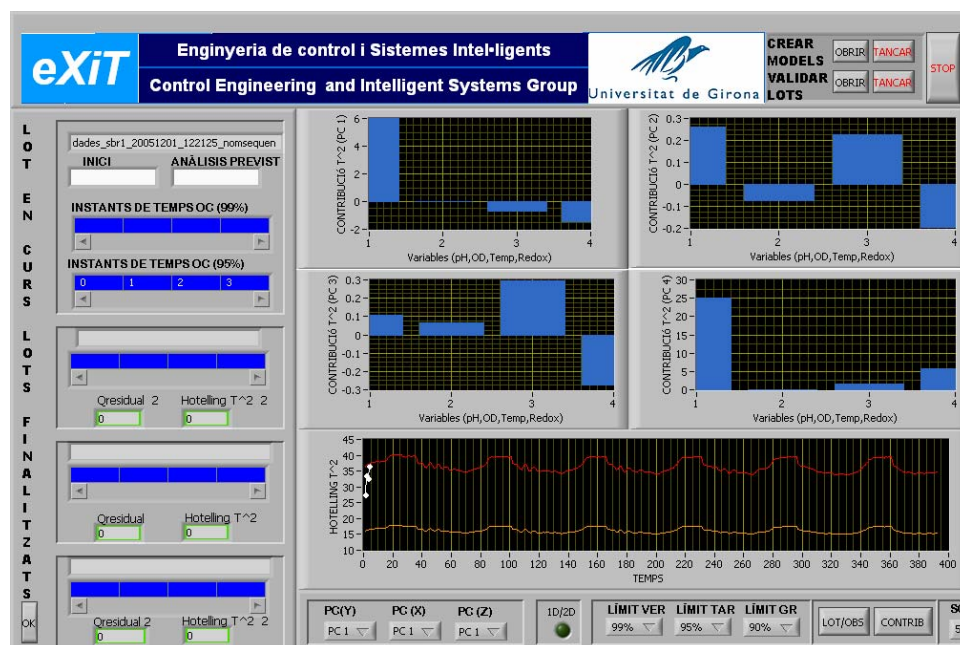


Figura 42. Gràfiques de contribucions a nivell d'observació

L'anàlisi a nivell d'observació es realitza mentre el lot encara no ha finalitzat. En el moment en què el lot arriba a l'etapa de Purga es pot portar a terme l'anàlisi del lot a nivell de lot.

Aquest anàlisi es realitza un cop el lot arriba a l'etapa de Purga perquè els models de referència es creen sense tenir en compte les dades referents a les etapes de Purga, Sedimentació i Buidat. D'aquesta manera, els nous lots no requereixen de les dades corresponents a aquestes etapes per a la seva validació.

Un cop el lot arriba a l'etapa de Purga i es realitza l'anàlisi a nivell de lot la interfície presenta les valors corresponents dins les caixes de control dels scores, Q i T^2 . Els valors de l'últim lot finalitzat són presentats de manera conjunta amb els dos últims lots finalitzats. D'aquesta manera es pot observar l'evolució del processat de la planta al llarg d'un dia. A la Figura s'observa la interfície mostrant les gràfiques corresponents a l'anàlisi a nivell de lot dels últims tres lots processats a la planta.

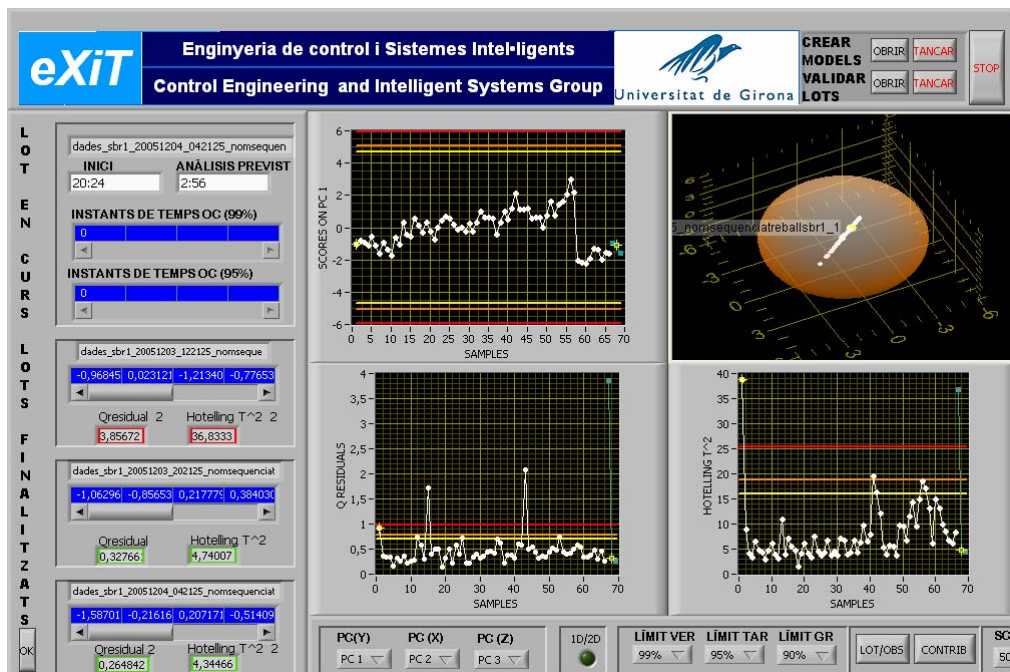


Figura 43 . Visualització de l'anàlisi a nivell de lot dels tres últims lots processats

En aquestes gràfiques els punts blancs fan referència als lots que formen el model de referència. Els lots nous projectats sobre aquest model de referència corresponen als quadrats verds. Aquests lots són els últims tres lots processats a la planta, en el moment que han entrat a l'etapa de *purga* s'ha realitzat automàticament l'anàlisi a nivell de lot de cadascun d'ells.

Totes aquestes gràfiques disposen de dos cursors, un destinat als lots del model i l'altre destinat al últims tres lots projectats. Aquests cursors estan connectats entre ells, d'aquesta manera si es selecciona un lot a qualssevol de les gràfiques de manera automàtica queda seleccionat a la resta. D'aquesta manera es mostra una informació ràpida i completa dels valors estadístics que caracteritzen un lot determinat.

Les gràfiques superiors de la Figura 43 s'observen les caixes de control corresponents als vectors *scores*. Aquests valors són les projeccions dels nous lots sobre les diferents direccions de l' hiperplà format pel model de referència. Mitjançant aquestes caixes de control s'observa en quines direccions o components el lot es troba fora dels límits de control. A partir d'aquestes gràfiques es pot visualitzar la caixa de control d'una sola components, de dues components o de tres components. A la Figura 44 s'observen les caixes de control per a cadascuna de les dimensions.

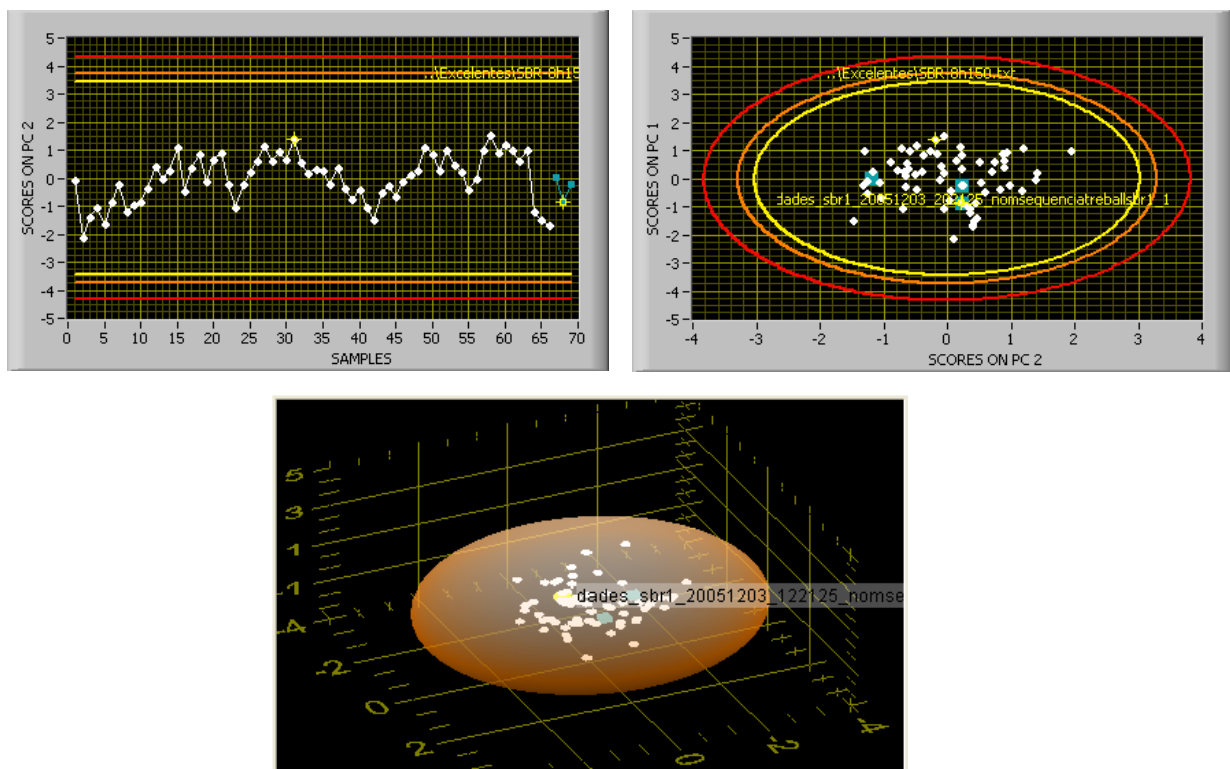


Figura 44. Gràfiques de les caixes de control dels scores

En aquestes gràfiques es pot observar com varia la geometria dels límits a cada dimensió. Quan s'enfronten dues components en una mateixa caixa de control, les regions de control formades són el·líptiques. En el cas d'enfrontar tres components en una mateixa caixa de control, el volum de control és el·lipsoïdal.

Les gràfiques que s'observen a la part inferior de la Figura 43 corresponen a les caixes de control dels paràmetres Q i T^2 . A la Figura 45 es mostra una ampliació de els gràfiques corresponents a les caixes de control d'aquests dos paràmetres estadístics. Quan un lot nou es troba fora de les regions de control del paràmetre Q denota que no és de la mateixa naturalesa que els lots amb què s'ha format el model. La major part de la informació d'aquest nou lot queda retinguda a la matriu residual. Per aquesta raó no es pot prendre en consideració el valor T^2 d'aquest nou lot ja que no representa la projecció sobre el model i no ens dóna informació sobre la bondat d'aquest lot.

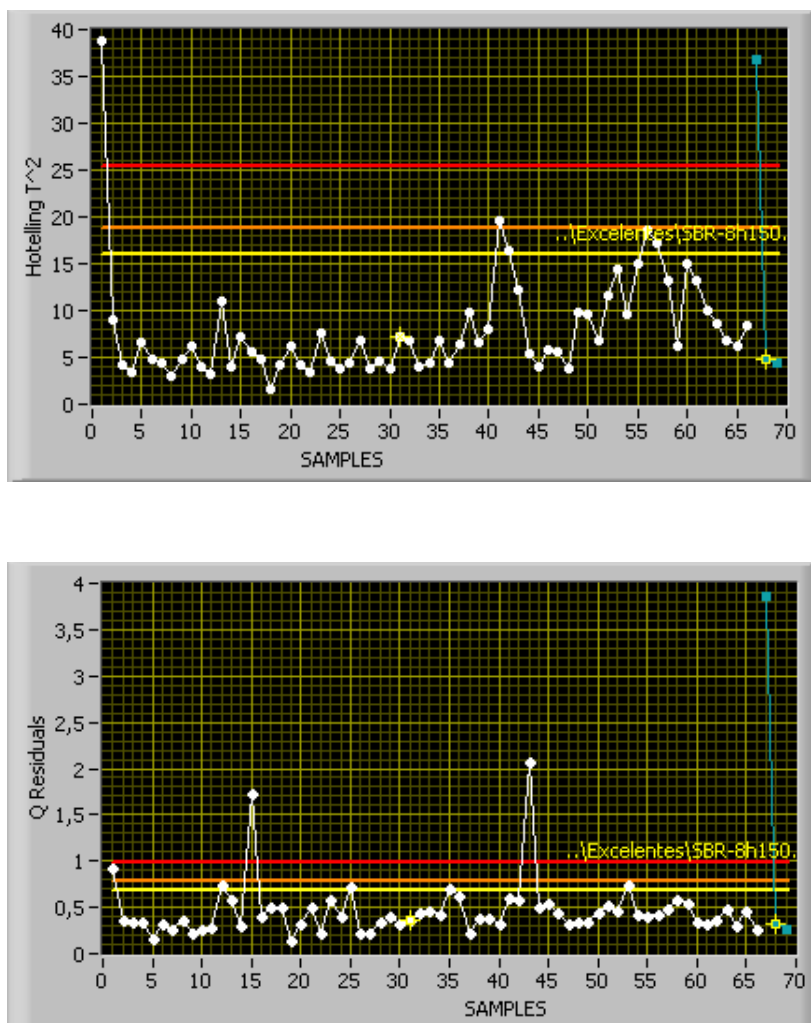


Figura 45. Gràfiques de les caixes de control dels paràmetres Q i T^2

Per altra banda, en cas que un lot es trobi sota control pel que respecte al paràmetre Q i fora de control pel que respecte al paràmetre T^2 es pot portar a terme un anàlisi de contribucions. L'anàlisi de contribucions permet passar d'un espai amb presència de variables latents (components principals) a un espai amb presència de les variables de procés. D'aquesta

manera és possible conèixer quines variables de procés i en quin instant de temps són les responsables que el lot es trobi fora de control. A la Figura 46 es mostra la interfície presentant l'anàlisi de contribucions d'un lot determinat.

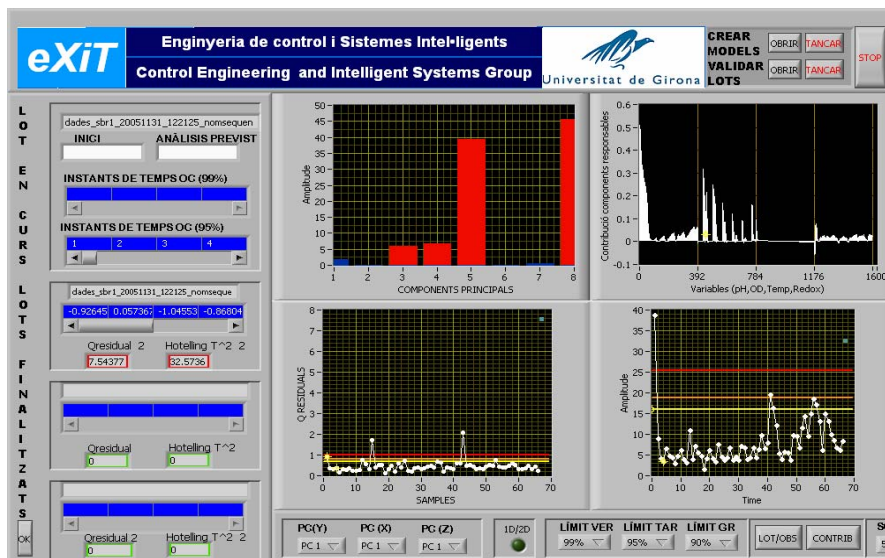


Figura 46. Gràfiques de contribucions d'un lot fora de control

A partir de la gràfica situada al marge superior esquerre s'observen en vermell les components principals responsables que el lot es situï per fora dels límits de control. Aquesta gràfica fa referència a la contribució percentual de cadascuna de les components al paràmetre T^2 . La gràfica situada al marge superior dret mostra la contribució del conjunt de components responsables que el lot estigui fora de control. Aquesta gràfica permet determinar quina variable de procés i en quin instant contribueix a que el comportament de lot no es pugui considerar acceptable. A la Figura 47 es mostra una ampliació de la gràfica de contribucions. En aquesta gràfica com el cursor indica a quina variable de procés i a quin instant de temps correspon cada punt de la gràfica.

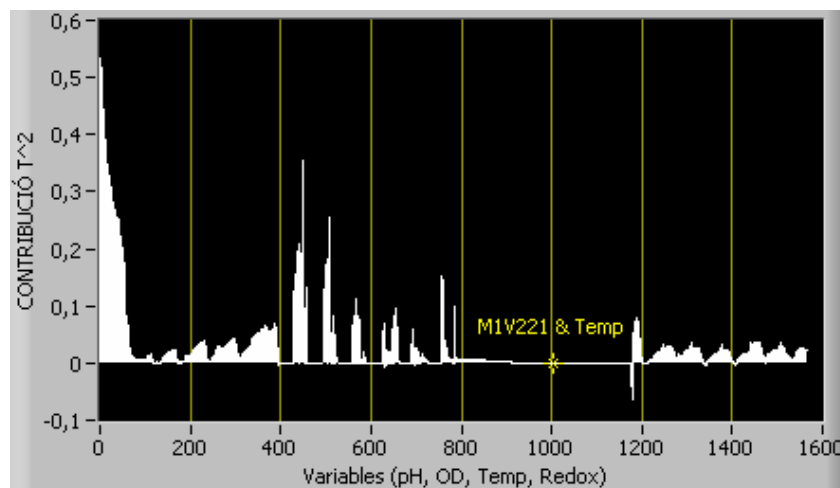


Figura 47. Gràfica de contribucions d'un lots fora de control del paràmetre T^2

En cas d'escollir un nou per a la realització de l'anàlisi de contribucions la interfície mostra les gràfiques del mode en que s'observa la Figura 48. La gràfica corresponent a la contribució percentual de les components no presenta components responsables ja que el lot es troba sota control. En aquest cas l'anàlisi de contribucions no es realitza en la seva totalitat i el gràfic apareix a la interfície deshabilitat.

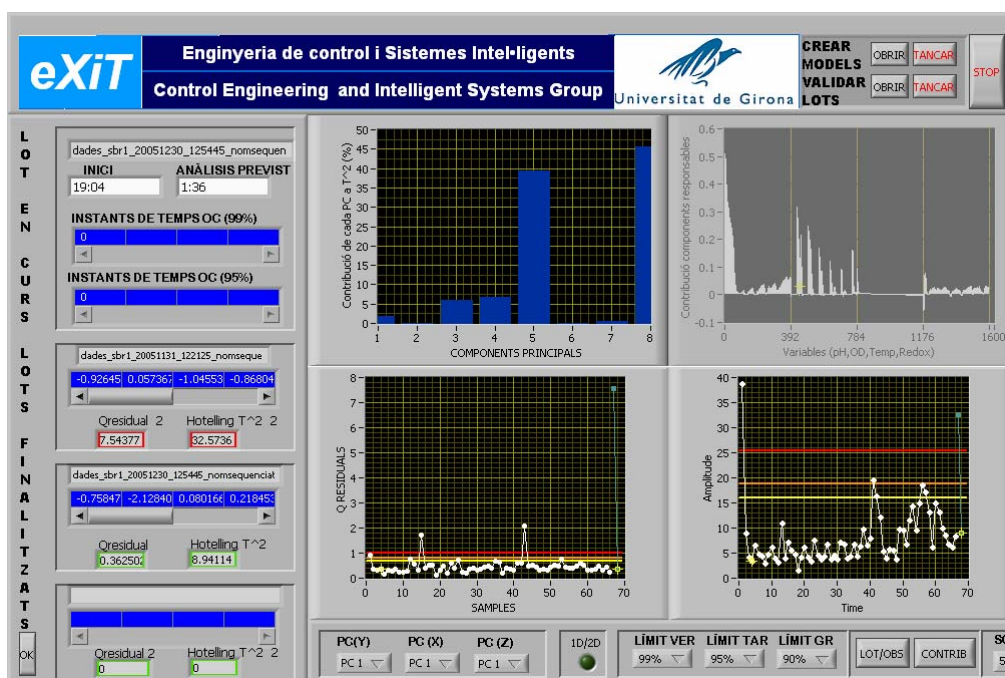


Figura 48. Gràfic de contribucions d'un lot sota control

Aquesta interfície disposa d'una finestra, com la de la Figura 49, on es mostra informació actualitzada sobre el lot que s'està processant en la planta i sobre els tres últims lots finalitzats que han estat processats.

Pel que respecte al lot que s'està processant a la planta s'indica la hora d'inici del seu procés i l' hora prevista en què es realitzarà l'anàlisi. Cal tenir en compte que l' hora de l'anàlisi previst i l' hora de finalització del lot no coincideixen. L'anàlisi es realitza una hora i mitja abans que finalitzi el lot, això és degut a que no són necessàries per a l'anàlisi les tres últimes etapes del procés (purga, sedimentació i buidat).

Aquesta informació també mostra un llistat dels instants de temps en què el lot es situa fora de control al llarg de l'anàlisi a nivell d'observació per als límits de confiança establerts del 95% i 99%.

Pel que respecte als últims tres lots finalitzats de la planta es detallen els valors dels scores i dels paràmetres Q i T². S'especifica en quina regió de control es troben aquestes paràmetres mitjançant el codi de colors establert.

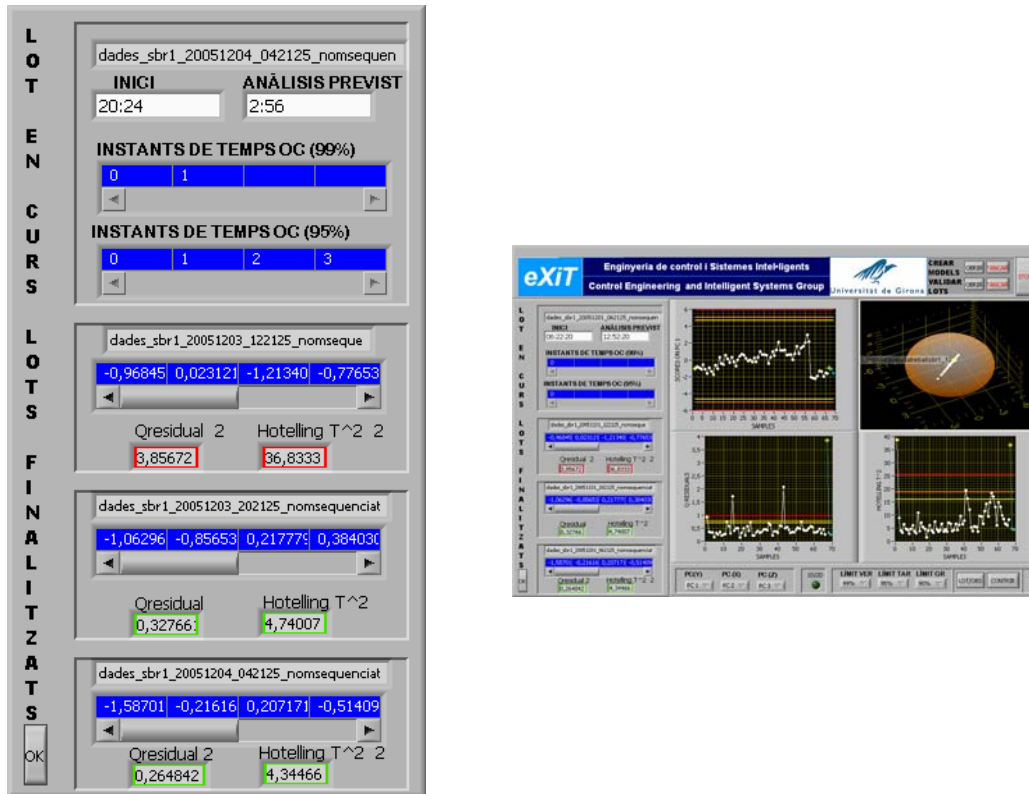


Figura 49. Finestra de la interfície del mòdul per a l'anàlisi en línia

Aquesta interfície correspon al mòdul per a l'anàlisi en línia. Aquest mòdul es troba en funcionament permanent i la seva interfície és sempre visible. D'aquesta manera, aquesta interfície actua com interfície principal de l'entorn per a l'anàlisi estadístic. A partir d'aquesta interfície es pot accedir a les altres dues interfícies corresponents als altres mòduls que conformen l'entorn. A la Figura 50 es mostra una ampliació dels comandaments situats al marge superior dret de la interfície. A partir d'aquests comandaments es poden obrir i tancar les altres dues interfícies. La interfície principal romandrà sempre oberta i el seu mòdul corresponent en funcionament en qualsevol dels casos.



Figura 50. Comandament per a accedir als altres mòduls de l'entorn

Totes les gràfiques de la interfície són controlades per una mateixa barra de comandaments que es situa a la part inferior de la interfície. A la Figura 51 es mostra aquesta barra de comandaments.

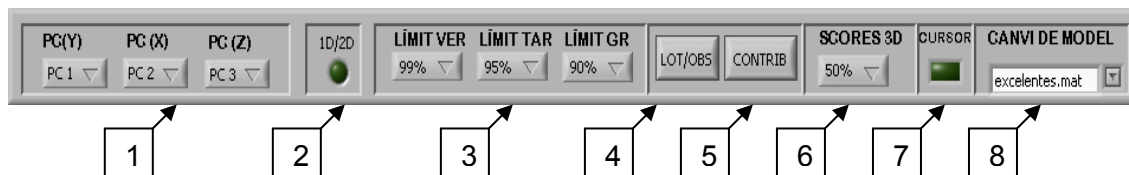


Figura 51. Barra de comandaments de la interfície

Els comandaments [1] permeten escollir les components que escollides per a cada eix gràfiques corresponents als vectors *scores*. El comandament [2] permet visualitzar les gràfiques dels *scores* en una o dues dimensions. Els comandaments [3] permeten modificar els nivells de confiança de les corresponents caixes de control. Aquests nivells poden variar entre el 90% i el 99% per cadascun dels tres límits: vermell, taronja i groc. El comandament [4] permet visualitzar les gràfiques corresponents a l'anàlisi a nivell d'observació o les gràfiques corresponents a l'anàlisi a nivell de lot. El comandament [5] permet visualitzar les gràfiques corresponents de l'anàlisi de contribucions que pertoqui segons si s'estan visualitzant les gràfiques per al nivell d'observació o per al nivell del lot. El comandament [6] permet variar el nivell de transparència de l'el·lipsoide que delimita la regió de control en un espai tridimensional. El comandament [7] permet visualitzar el cursor de les gràfiques corresponent als lots del model, el fet de coincidir en l'espai pot dificultar la visió del cursor dels lots projectats. El comandament [8] permet escollir i canviar el model de referència establert tant a l'anàlisi a nivell de lot com a l'anàlisi a nivell d'observació.

5.6.2 Interfície del mòdul de creació de models

El mòdul de creació de models és el mòdul de l'entorn destinat a la creació de nous models de referència. Aquests models són requerits en els altres mòduls per a projectar nous lots per a la seva validació. La interfície corresponent a aquest mòdul a partir de la qual pot interactuar és la que es mostra a la Figura 52. L'usuari pot accedir a aquesta interfície des la interfície principal.

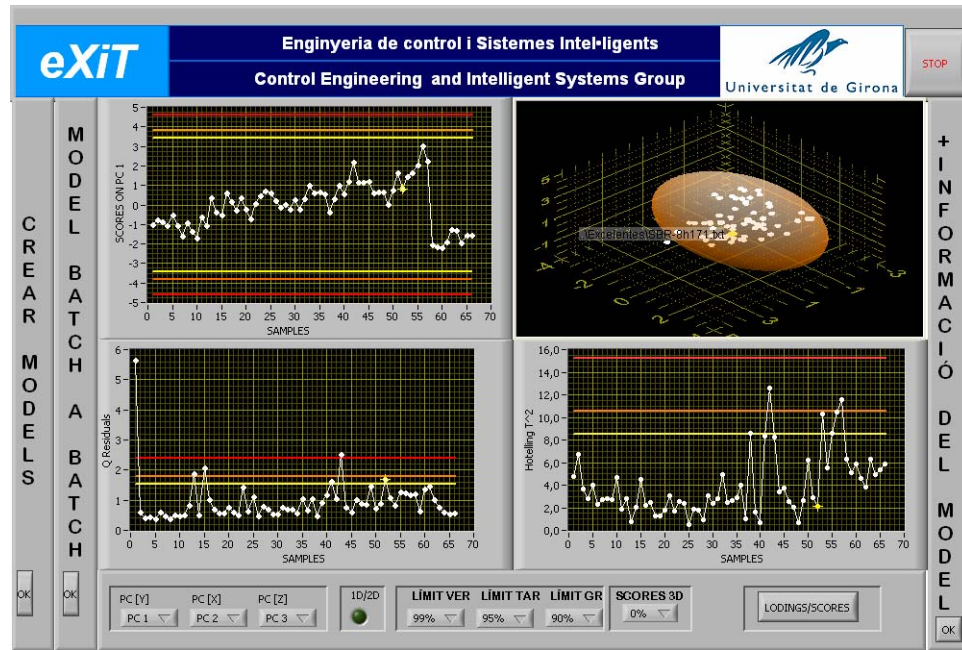


Figura 52. Interfície del mòdul de creació de models

En aquest mòdul l'usuari pot crear nous models. Aquests models es poden crear a partir dels lots emmagatzemats a la base de dades i a partir de dades emmagatzemades en arxius MATLAB. Aquest mòdul també permet la visualització de models ja creats per a la seva revisió. A la Figura 53 es mostra la finestra que permet a l'usuari escollir entre les diverses possibilitats esmentades per a la creació d'un nou model i dels lots emmagatzemats a la base de dades en el cas corresponent.

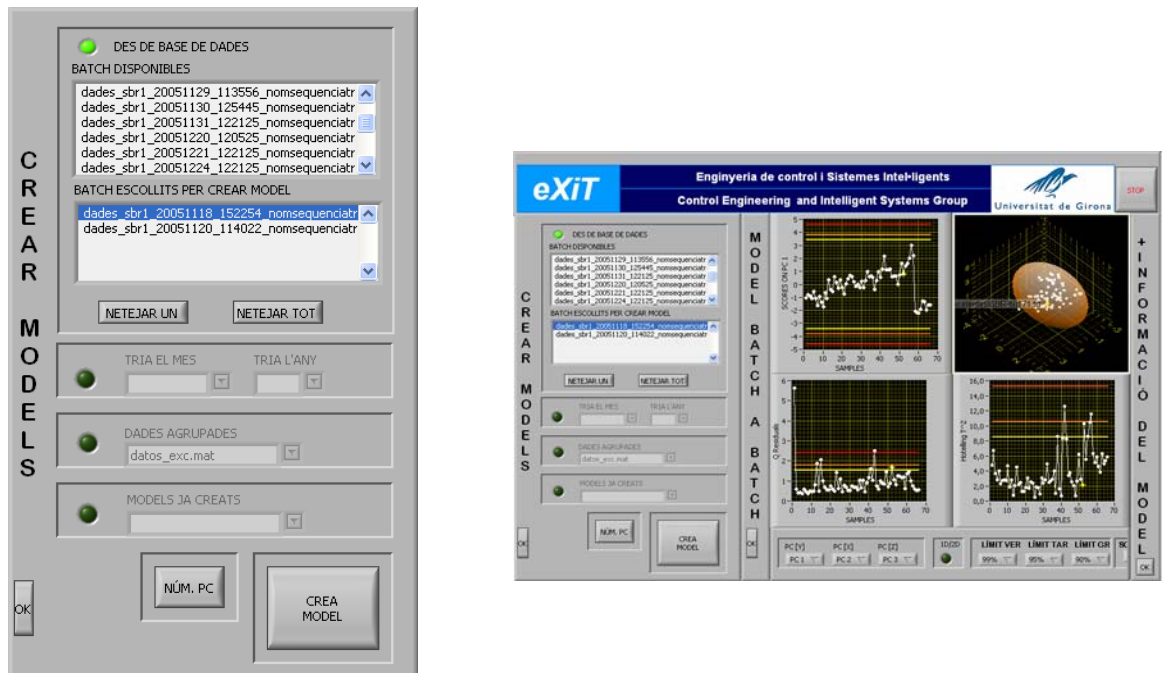


Figura 53. Finestra de la interfície per a l'elecció de les dades

Un cop s'escullen les dades necessàries per a la creació d'un nou model, apareix una finestra flotant on s'introdueix el processat de les dades que es vol portar a terme i el nom amb què es vol guardar el model. Un cop introduïda la informació requerida el mòdul crea un model previ a partir del qual l'usuari pot decidir el número de components amb què vol crear el model definitiu. A la Figura 54 es mostra la interfície en el moment que es presenta la informació corresponent al model previ creat.

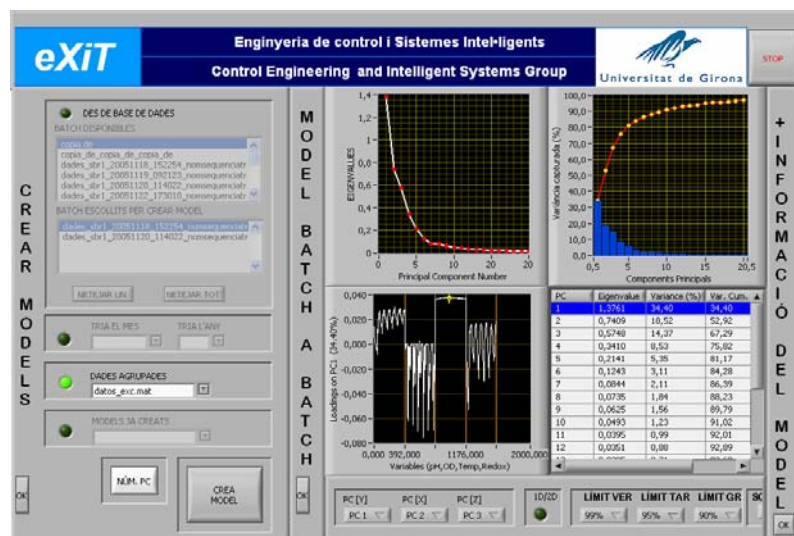


Figura 54. Visualització de les gràfiques del model previ

A partir les gràfiques i la taula presentades corresponents al model previ creat, l'usuari pot decidir el número de components principals adient per a la creació del model definitiu. Els criteris per a l'elecció del número de components han estat descrits a l'apartat 4.3.1. A la Figura 55 es mostren de manera ampliada dues de les gràfiques corresponents al model previ presentades a la part superior de la interfície.

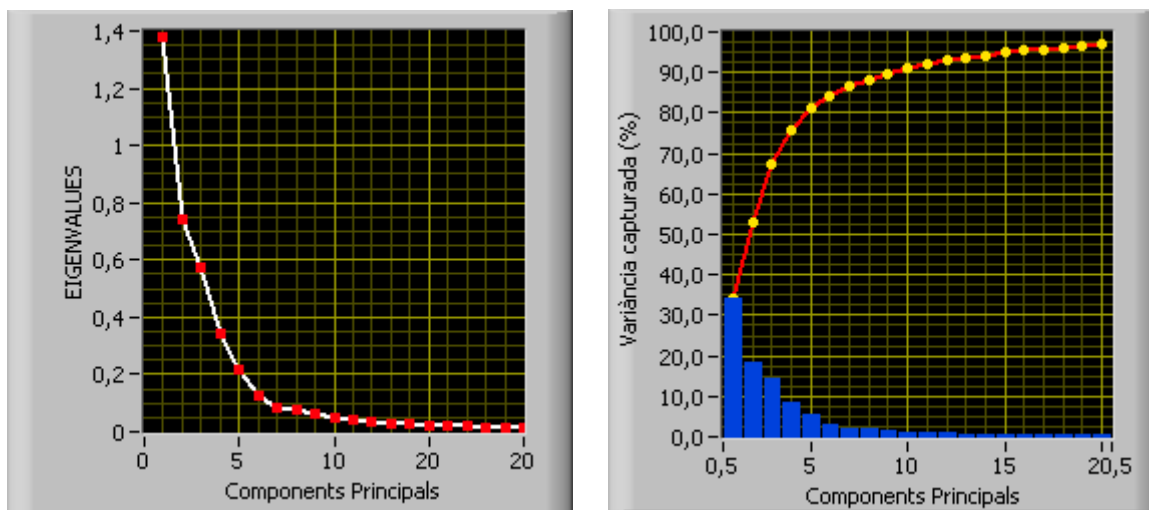


Figura 55 . Gràfiques dels eigenvalues i variància capturada del model previ

A la gràfica de l'esquerra es mostren els *eigenvalues* corresponents a cada component principal. Aquesta gràfica permet observar els pendents resultants de les rectes entre cada component. Un dels criteris per a l'elecció del número de components es basa en l'aparició d'un *genoll* a la gràfica. La gràfica de la dreta mostra el tant per cent de variància retinguda per cadascuna de les components i la variància acumulada corresponent. A partir d'aquesta gràfica es visualitza el tant per cent de variància que reté el model si s'escull un número determinat de components.

La gràfica i la taula presentades a la part inferior de la interfície i que també corresponen al model previ es recullen ampliades a la Figura 56.

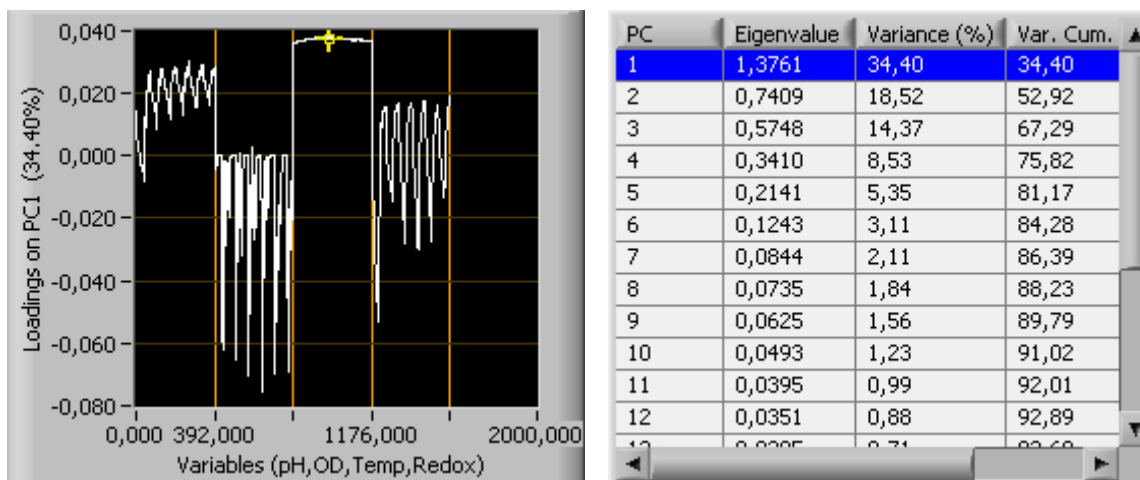


Figura 56 . Gràfica dels vectors *loadings* i taula de variàncies del model previ

La gràfica de l'esquerra mostra els valors corresponents als vectors *loadings* del model previ. En aquest cas es visualitza el vector corresponent a la primera component. L'usuari pot visualitzar qualsevol dels vint vectors *loadings* corresponents a les vint components principals amb què es crea el model previ. Aquesta gràfica permet l'elecció del número de components a partir del criteri basat en els vectors *loadings*.

La taula de la dreta recull de manera numèrica els valors corresponents als *eigenvalues*, la variància capturada i la variància acumulada per a cada component. Aquesta taula es mostra per complementar la informació presentada per la resta de gràfiques i per permetre l'elecció del número de components basant-se en un criteri diferent als esmentats a l'apartat 4.3.1.

Un cop l'usuari determina el número de components adient es crea el model definitiu. A la Figura 57 es mostra la interfície en el moment de presentar les gràfiques corresponents als valors característics del model.

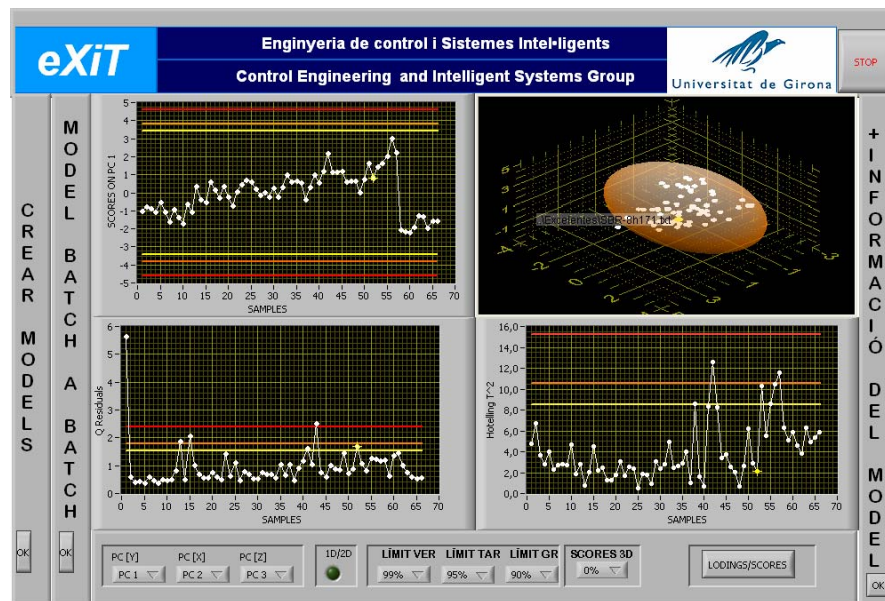


Figura 57. Visualització gràfiques del model definitiu

Les gràfiques presentades per la interfície corresponents al model definitiu són de la mateixa tipologia que les presentades per la interfície del mòdul per a l'anàlisi en línia per a l'anàlisi a nivell de lot. Aquestes gràfiques mostren els valors dels vectors *scores* i els paràmetres estadístics multivariables Q i T^2 per a cadascun dels lots que formen el model. Tots aquests valors són presentats dins les corresponents caixes de control on es marquen els límits de control per a tres nivells de confiança. Els vectors *scores* poden ser visualitzats en una, dos o tres dimensions.

Mitjançant els cursors presents a cadascuna de les caixes de control podem conèixer a quin lot correspon cadascun dels lots. Aquests cursors estan connectats entre ells i per tant es mouen de manera simultània a totes les caixes de control.

Aquesta interfície disposa de més gràfiques que descriuen el model creat. Aquestes gràfiques fan referència als vectors *loadings*. Aquests vectors ens donen informació sobre l'estructura de l'hiperplà creat pel model i permeten portar a terme anàlisi més complexos referents a la naturalesa de les variables de procés que hi intervenen. A la Figura 58 es presenta la interfície mostrant les gràfiques corresponents als vectors *loadings*.

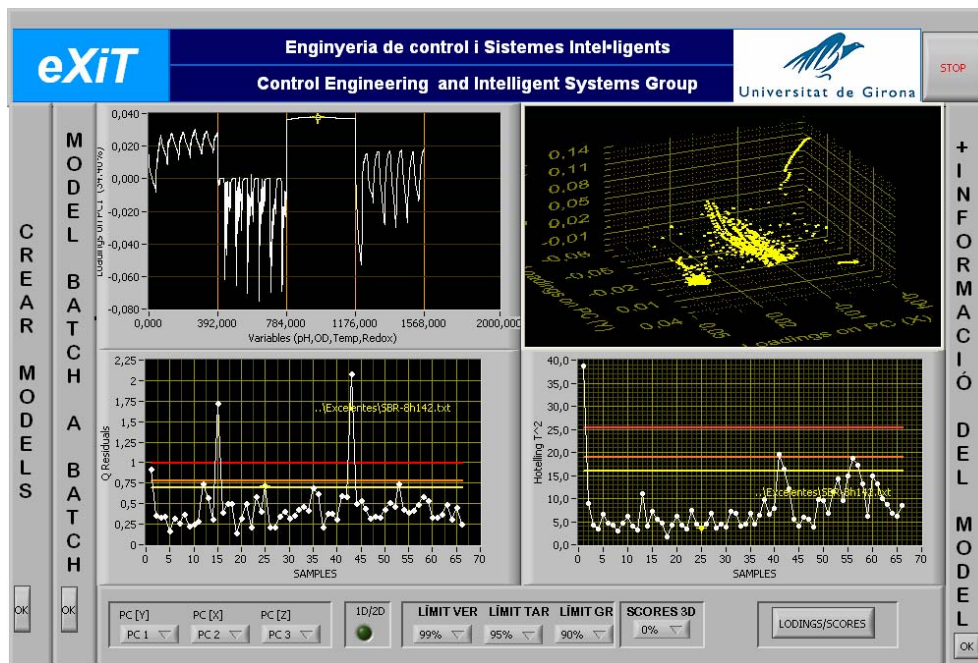


Figura 58. Visualització gràfiques dels vectors loadings del model definitiu

Aquests valors dels vectors loadings poden ser visualitzats de la mateixa manera que els vectors *scores* en una, dues o tres dimensions. A la Figura 59 s'observa les gràfiques resultants dels vectors *loadings* en una, dues o tres dimensions.

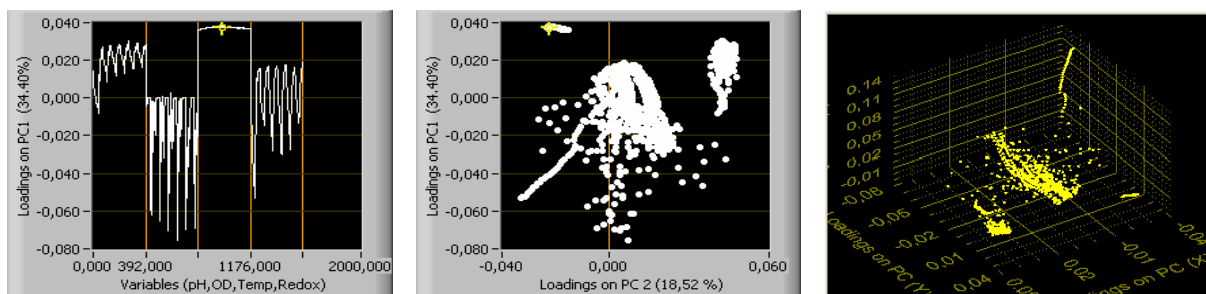


Figura 59 . Gràfiques dels vectors loadings en les tres dimensions

Aquesta interfície també permet conèixer de manera detallada la informació de cadascun dels lots que formen el model. A la Figura 60 es mostra la finestra de la interfície que permet visualitzar aquesta informació addicional. En aquesta finestra es mostra la gràfica dels *scores* corresponents a un lot determinat i els valors dels paràmetres Q i T^2 . També s'especifica en quina regió de control es troben aquests valors mitjançant el codi de colors.

El lot visualitzat a cada instant es correspon amb el lot seleccionat a les gràfiques mitjançant els cursors anteriorment esmentats.

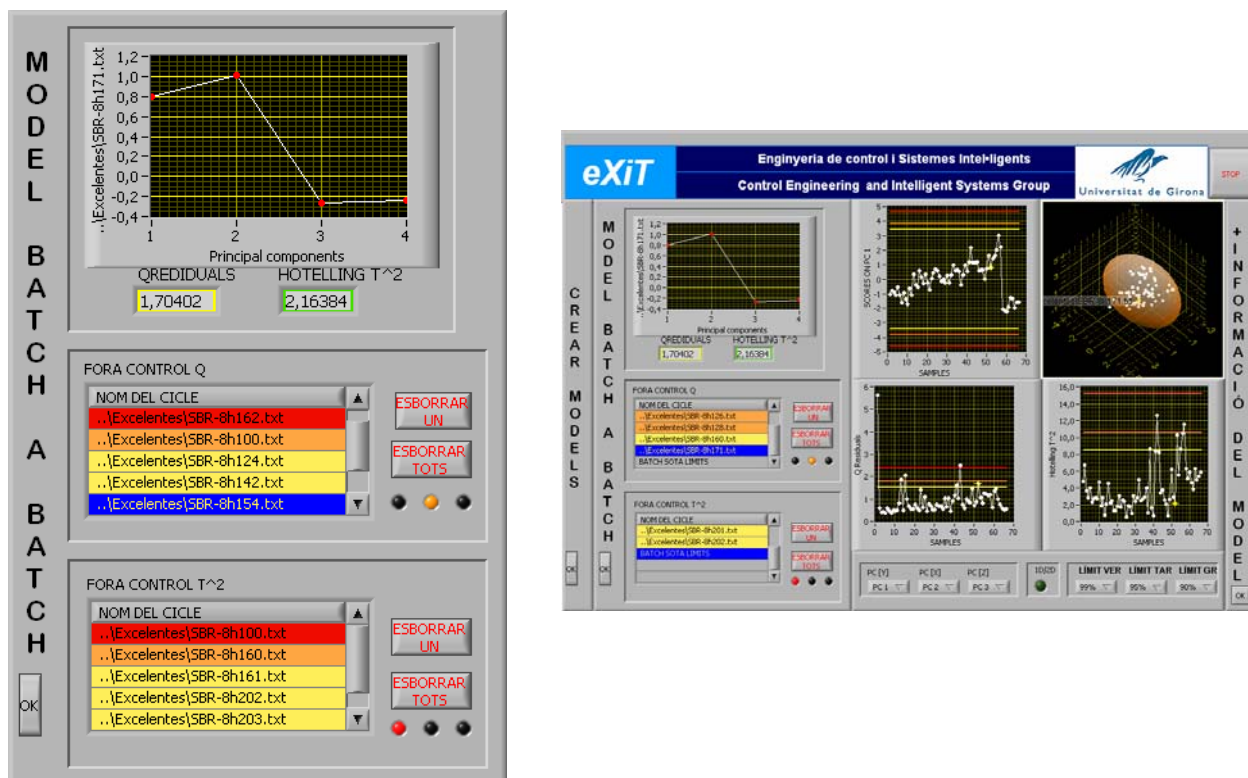


Figura 60 . Finestra de la interfície que mostra informació detallada dels lots

Aquesta finestra també mostra un llistat de tots els lots que formen el model i que es troben fora dels límits de control tant del paràmetre Q com del paràmetre T^2 . Aquests llistats mostren de manera ordenada els lots que es troben a les corresponents regions fora de control. Aquests llistats també es troben connectats als cursors de les gràfiques, i per tant es pot situar un lot a les gràfiques a partir d'aquestes llistes.

Una de les funcions que també presenta aquest mòdul és l'edició de models. Un cop creat un model es pot decidir l'eliminació d'alguns dels lots que el formen. D'aquesta manera es pot aconseguir un model més ajustat i sense la presència de lots que puguin desvirtuar el model. La finestra que es mostra a la Figura permet l'edició de manera senzilla dels models.

L'edició dels models es pot realitzar eliminant un sol lot o es poden escollir de manera directa el conjunt de lots que es troben dins una regió fora de control per a un dels dos paràmetres estadístics.

La interfície disposa d'una última finestra on es presenta informació addicional referent al model creat. En aquesta finestra es mostra la gràfica corresponent a la variància capturada i acumulada per cadascuna de les components. També es mostren les taules amb els valors numèrics corresponents a totes les gràfiques presentades a la interfície. Aquestes taules recullen els vectors loadings i els valors dels scores, Q i T^2 per a cadascun dels lots que formen el model. Aquestes taules també es troben connectades a les gràfiques de la interfície. En seleccionar un lot a qualsevol de les gràfiques es visualitza automàticament el conjunt de valors que li correspon. A la Figura 61 es mostra la finestra de la interfície on es visualitza tota aquesta informació esmentada.

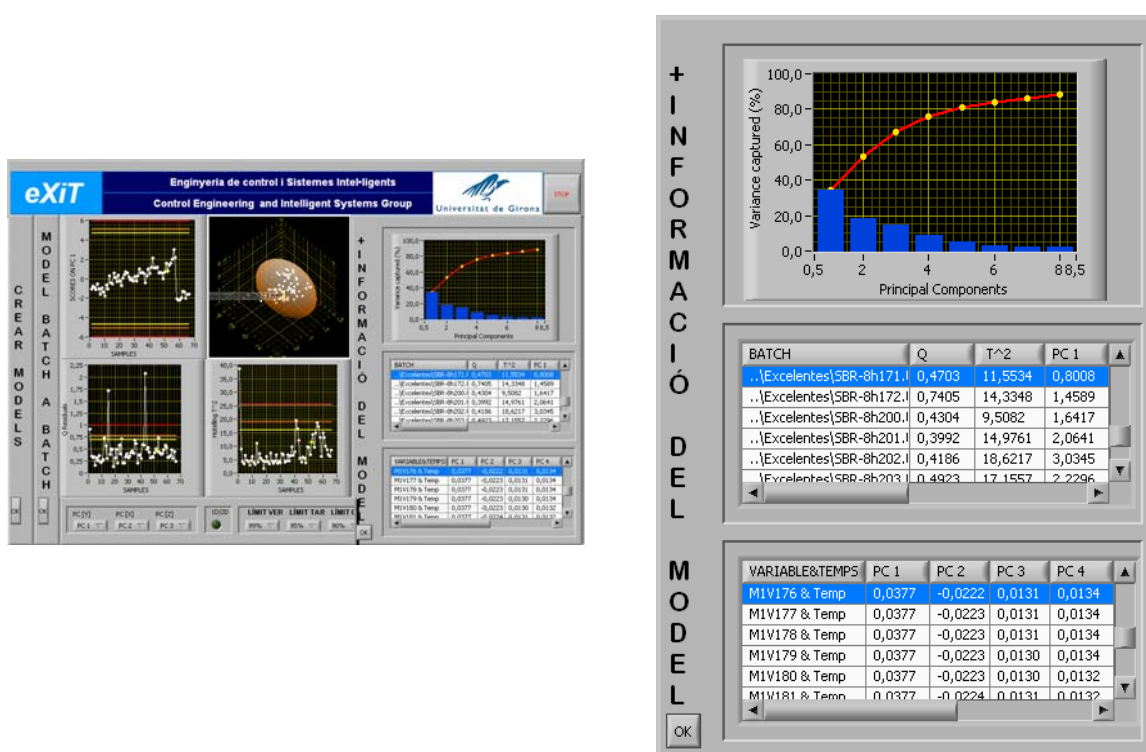


Figura 61 . Finestra de la interfície amb informació addicional del model

Totes les gràfiques de la interfície són controlades per una mateixa barra de comandaments que es situa a la part inferior de la interfície. A la Figura 62 es mostra una ampliació d'aquesta barra de comandaments.

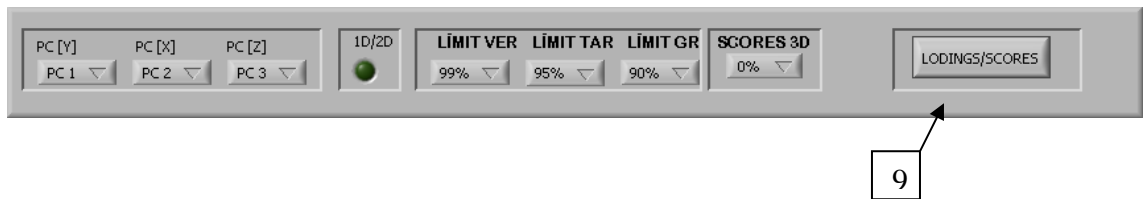
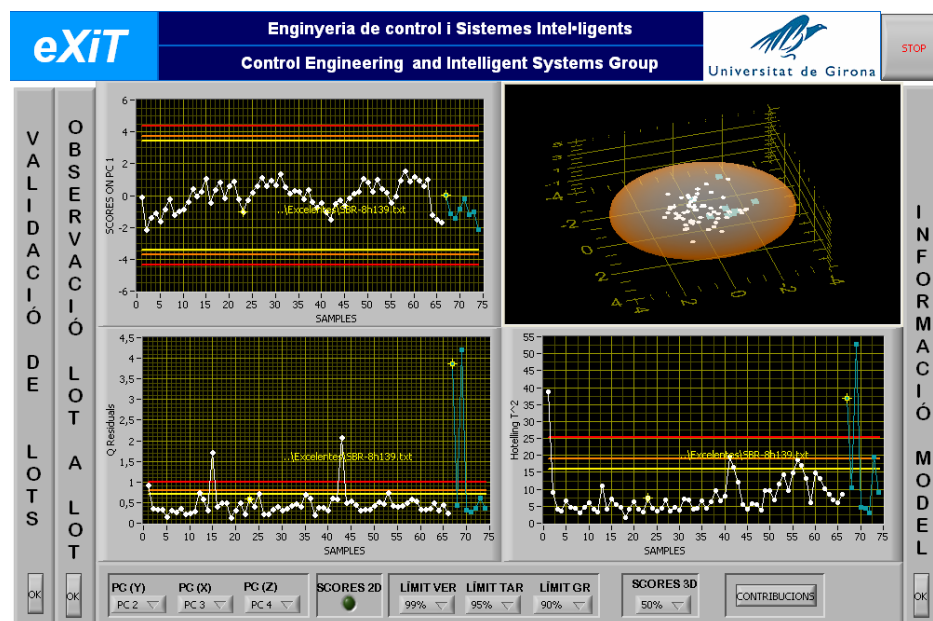


Figura 62 . Barra de comandaments de la interfície

Aquesta barra està formada per un conjunt de comandaments similar al descrit en la interfície del mòdul per a l'anàlisi en línia. En aquesta barra apareix un comandament diferent [9] permet escollir entre la visualització de les gràfiques corresponents als *scores* o les gràfiques corresponents als *loadings*.

5.6.3 Interfície del mòdul validació de lots

El mòdul per a la validació de lots permet projectar un conjunt de lots emmagatzemats a la base de dades sobre un model de referència determinat. Aquesta projecció dels lots sobre un model ens permet conèixer la bondat dels lots en referència amb un conjunt de lots que estan agrupats per unes característiques determinades. L'anàlisi estadístic realitzat en aquest mòdul es porta a terme exclusivament a nivell del lot. Per portar a terme aquest tipus d'anàlisi el procés del lot ha d'haver finalitzat. La interfície corresponent a aquest mòdul es mostra a la Figura 63 . L'usuari pot accedir a aquest mòdul des de la interfície principal.



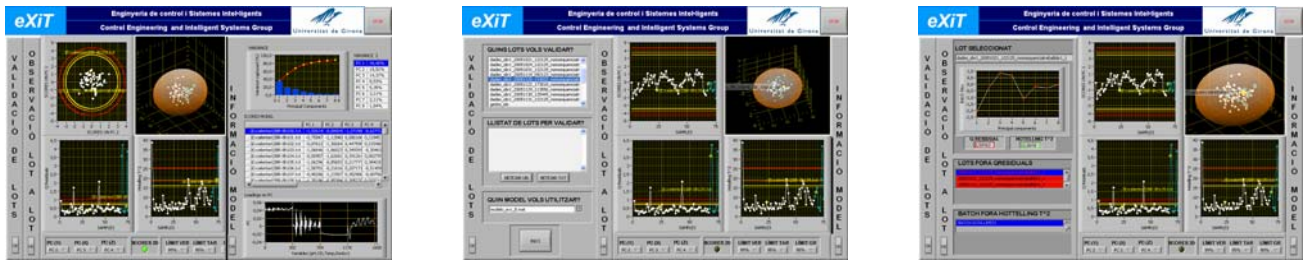


Figura 63. Interfície del mòdul per a la validació de lots

En aquest mòdul l'usuari pot escollir un conjunt de n lots emmagatzemats a la base de dades per ser validats. La interfície disposa d'una finestra, com la de la Figura 64, que mostra el conjunt de lots presents a la base de dades. L'usuari pot escollir un per un els lots desitjat per a ser validats. En aquesta mateixa finestra es mostren els models que poden ser escollits per actuar com a base de la projecció. Aquests models han estat creats pel mòdul de l'entorn corresponent a la creació de models. També poden ser models creat en programes externs i emmagatzemats en arxius MATLAB.

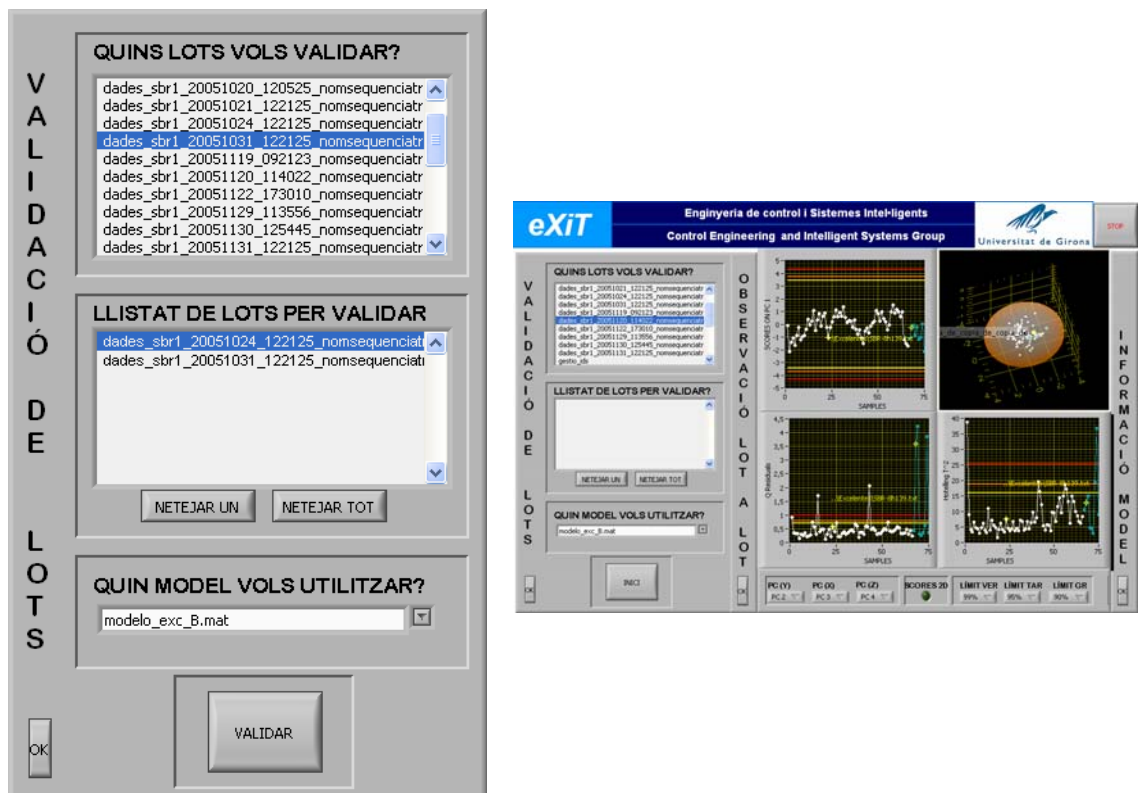


Figura 64. Finestra de la interfície que mostra els lots i model per a validar

Una vegada l'usuari tria els lots per a validar i el model determinat, el mòdul porta a terme les corresponents operacions per a projectar els nous lots sobre el model. Els valors obtinguts per a cada lot són introduïts a les corresponents caixes de control. Les gràfiques d'aquestes caixes de control es visualitzen a la interfície del mòdul com es pot veure a la Figura 65 .

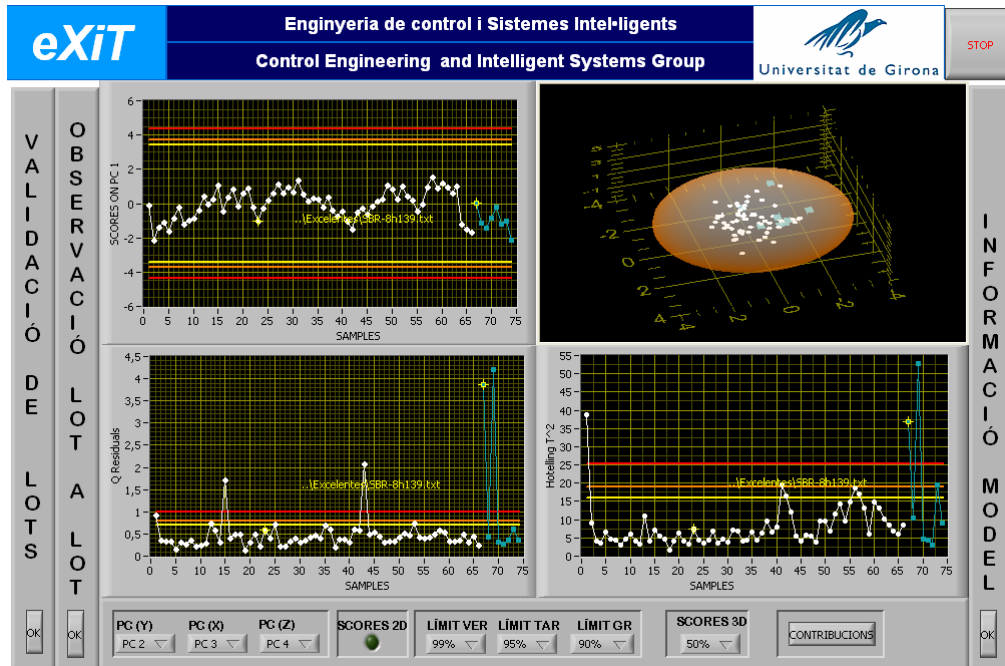


Figura 65. Visualització de les gràfiques de la validació dels lots

Aquestes gràfiques són de la mateixa tipologia que les mostrades en mòduls anteriors referents a l'anàlisi estadístic a nivell de lot. Aquestes gràfiques fan referència a les caixes de control corresponents als vectors *scores* i als paràmetres Q i T^2 . En aquestes gràfiques com en els altres mòduls, els punts blancs representen els lots que formen part del model de referència. Els lots nous, escollits per ser validats, són representats per quadrats verds per a ser fàcilment visualitzats. Aquestes gràfiques estan dotades de doble cursor, un destinat als lots del model i l'altre destinat als lots nous. Aquests cursors es mouen de manera simultània en totes les gràfiques. A la Figura 66 es mostren les gràfiques ampliades corresponents a les caixes de control dels paràmetres Q i T^2 on es diferencien els lots nous i en quina regió de control es troben situats.

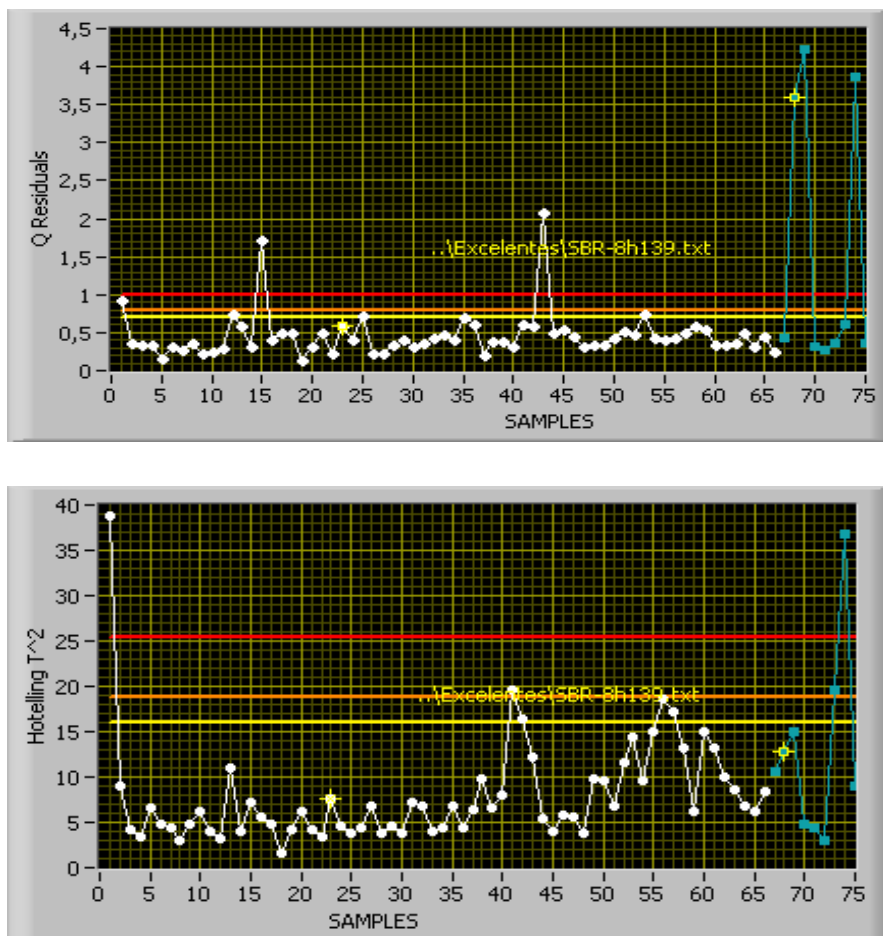


Figura 66. Gràfiques de les caixes de control dels paràmetre Q i T^2

Quan un lot nou es troba fora de les regions de control del paràmetre Q denota que no és de la mateixa naturalesa que els lots amb què s'ha format el model. La major part de la informació d'aquest nou lot queda retinguda a la matriu residual. Per aquesta raó no es pot prendre en consideració el valor T^2 d'aquest nou lot ja que no representa la projecció sobre el model i no ens dóna informació sobre la bondat d'aquest lot.

Per altra banda, en cas que un lot es trobi sota control pel que respecte al paràmetre Q i fora de control pel que respecte al paràmetre T^2 es pot portar a terme un anàlisi de contribucions. L'anàlisi de contribucions permet passar d'un espai amb presència de variables latents (components principals) a un espai amb presència de les variables de procés. D'aquesta manera és possible conèixer quines variables de procés i en quin instant de temps són les responsables que el lot es trobi fora de control. A la Figura 67 es mostra la interfície presentat l'anàlisi de contribucions d'un lot determinat.

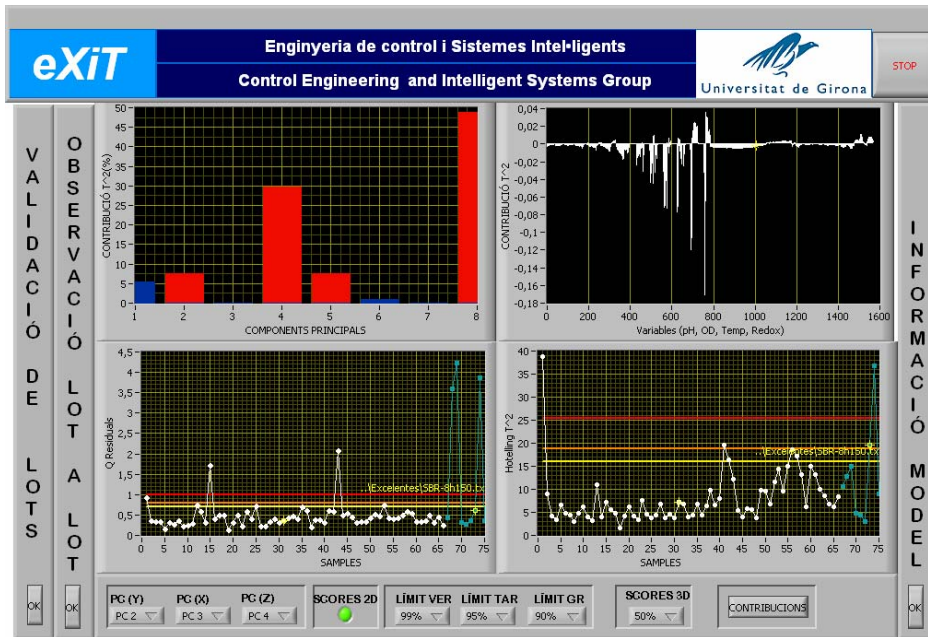


Figura 67. Gràfiques de l'anàlisi de contribucions d'un lot fora de control

A partir de la gràfica situada al marge superior esquerre s'observen en vermell les components principals responsables que el lot es situï per fora dels límits de control. Aquesta gràfica fa referència a la contribució percentual de cadascuna de les components al paràmetre T^2 . La gràfica situada al marge superior dret mostra la contribució del conjunt de components responsables que el lot estigui fora de control. Aquesta gràfica permet determinar quina variable de procés i en quin instant contribueix a que el comportament de lot no es pugui considerar acceptable. A la Figura 68 es mostra una ampliació de la gràfica de contribucions. En aquesta gràfica com el cursor indica a quina variable de procés i a quin instant de temps correspon cada punt de la gràfica.

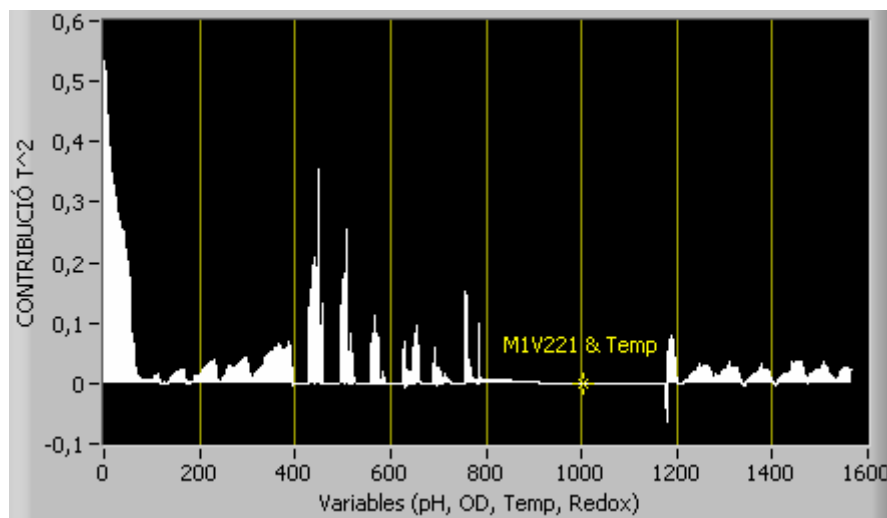


Figura 68. Gràfica de contribucions per als lots fora de control

En cas d'escollir un nou per a la realització de l'anàlisi de contribucions la interfície mostra les gràfiques del mode en que s'observa la Figura 69. La gràfica corresponent a la contribució percentual de les components no presenta components responsables ja que el lot es troba sota control. En aquest cas l'anàlisi de contribucions no es realitza en la seva totalitat i el gràfic apareix a la interfície deshabilitat.

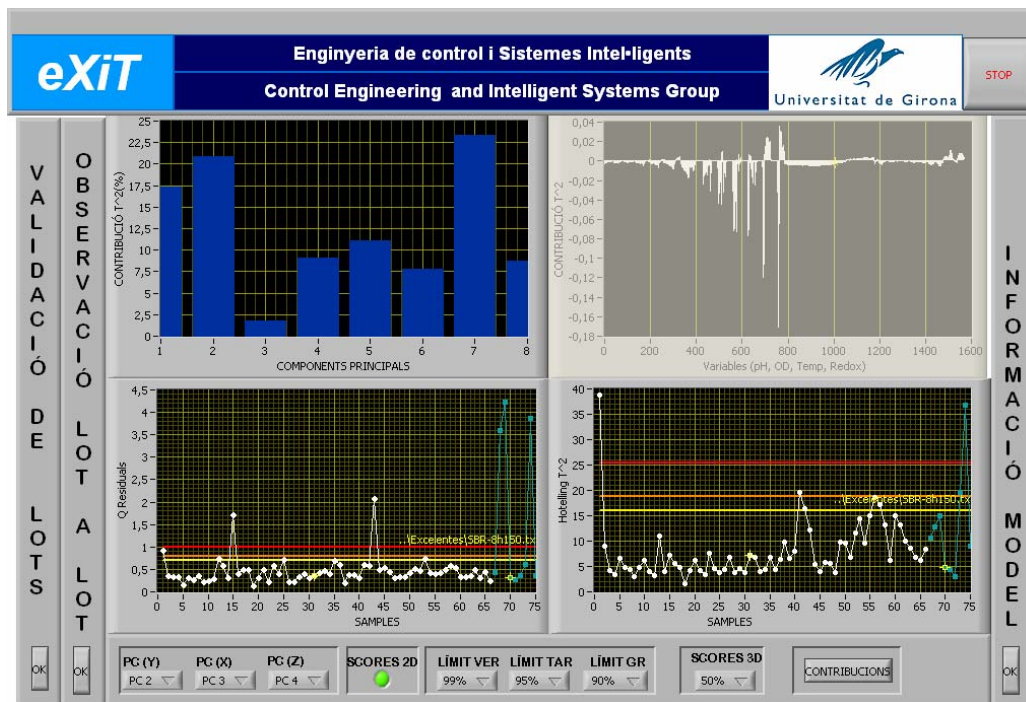


Figura 69. Gràfiques de contribucions per a un lot sota control

Aquesta interfície permet obtenir informació detallada de cadascun dels nous lots escollits per a validar. La finestra de la interfície que s'observa a la Figura 70, mostra una gràfica amb el vector scores del lot nou seleccionat. També es detalla el valor numèric dels paràmetres Q i T^2 per aquest nou lot i la regió de control en la que es troben mitjançant un codi de colors. Aquesta mateixa interfície presenta els llistats dels lots nous fora de control dels paràmetres Q i T^2 , ordenats per regions de control. Aquesta finestra es troba connectada al cursor corresponent als nous lots i mostra automàticament la informació del lot seleccionat en qualsevol de les gràfiques.

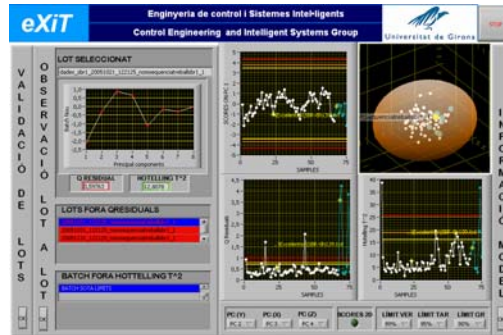
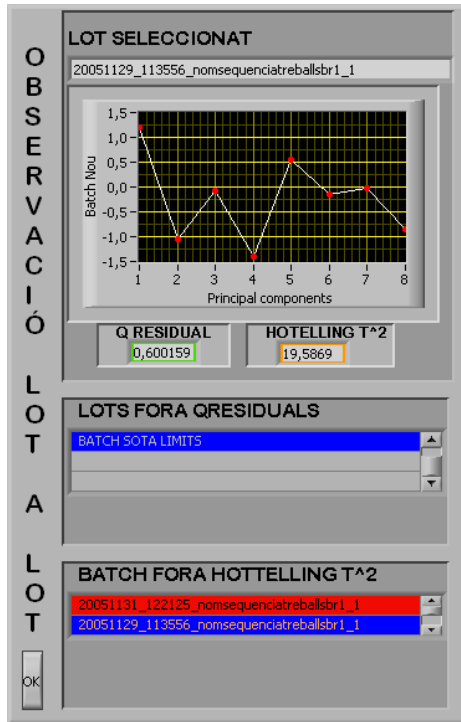


Figura 70. Finestra de la interfície que mostra informació detallada del lot seleccionat

Aquesta interfície també disposa d'una finestra on es detalla més informació sobre el model de referència escollit. En aquesta finestra, com es veu a la Figura 71, es mostra la gràfica de variància capturada i acumulada per a cadascunes de les components i la gràfica corresponents al vectors *loadings* del model. També es mostra una taula amb els valors dels scores i els paràmetres Q i T² per a cadascun dels lots que formen el model.

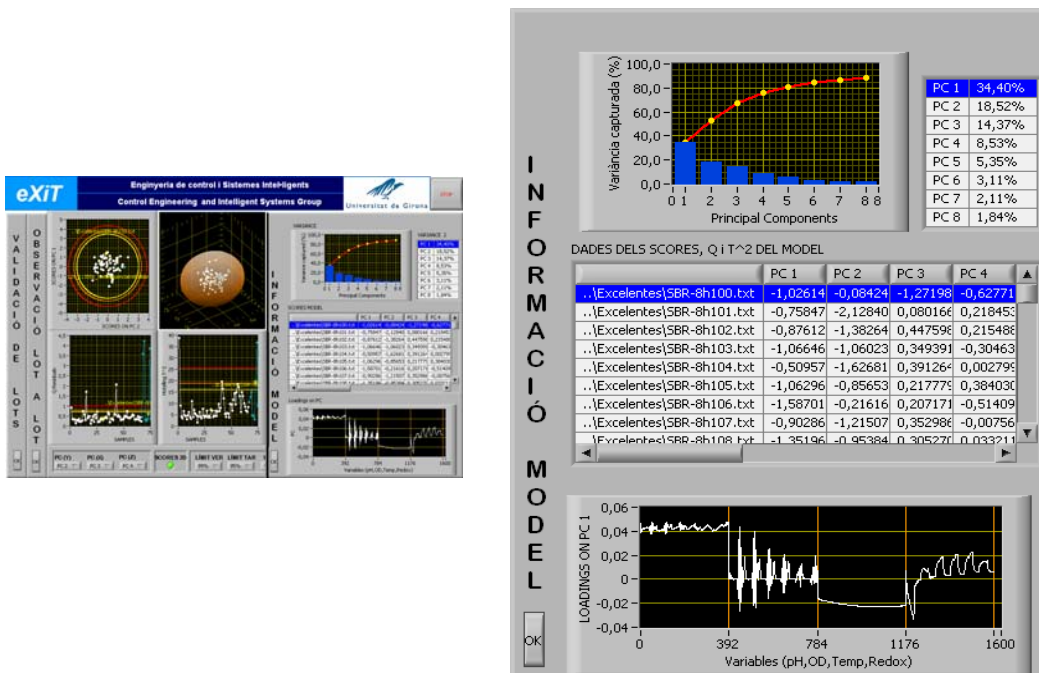


Figura 71. Finestra de la interfície on es mostra informació addicional del model de referència

Gràcies a aquesta finestra l'usuari pot accedir a una informació ràpida del model sense la necessitat d'adreçar-se al mòdul destinat a la creació de models.

La barra de comandaments d'aquesta interfície es mostra a la Figura 72 . Aquesta barra és similar a la presentada a la resta d'interfícies.



Figura 72 . Barra de comandaments de la interfície per a la validació de lots

6.VALORACIÓ ECONÒMICA

El cost total del projecte és de:

12.099,00 € (DOTZE MIL NORANTA NOU EUROS)

Girona, gener 2006

Miquel Àngel Castro Ortega

7. CONCLUSIONS

L'objectiu d'aquest projecte era la realització d'una aplicació que permetés portar a terme el control estadístic multivariable en un SBR per a la depuració d'aigües residuals. Aquest objectiu inicial s'ha complert de forma satisfactòria.

Aquesta aplicació permet realitzar un anàlisi estadístic multivariable complet dels lots processats a una planta SBR. Aquest anàlisi es pot portar a terme tant a nivell d'observació com a nivell del lot. D'aquesta manera es pot realitzar un seguiment complet del processat del lot des del seu inici i un cop finalitzat. L'aplicació realitza els anàlisis corresponents al lot en procés i als darrers lots finalitzats a la planta de manera automàtica.

Es presenten de manera clara i ràpida un conjunt de gràfics de control i gràfics de contribució on es recull la informació obtinguda dels anàlisis estadístics corresponents. L'aplicació també presenta uns llistats dels lots que es troben fora de control i permet accedir a informació detallada de cadascun dels lots analitzats.

Aquesta aplicació ha estat provada en unes condicions similars a les condicions de funcionament normal a la planta SBR. La correcció dels resultats obtinguts ha estat contrastada a partir d'altres aplicacions professionals de control estadístic. L'aplicació no ha pogut ser instal·lada ni a la planta experimental ni a la planta pilot degut a la paral·lela actualització del mòdul de monitorització en l'entorn LabVIEW.

S'ha desenvolupat una eina vàlida per a tot tipus d'usuari. Per una banda, permet a l'operari de planta realitzar un control de càlcul complex de manera senzilla, obtenint un conjunt de gràfics que li permeten la detecció de possibles anomalies a la planta. Per altra banda, aquesta aplicació també és útil per a investigadors en el camp de les tècniques de control basades en PCA, ja que permet l'anàlisi detallat de nous models per tal de perfeccionar-los i obtenir nova informació provinent de les dades de la planta. La base de dades utilitzada en el mòdul de monitorització permet l'exportació de les dades mitjançant arxius d'extensió *sql* que permeten la utilització de l'aplicació fora de línia.

Una de les dificultats més importants en la realització d'aquesta aplicació es basen en la baixa capacitat del programa LabVIEW alhora de realitzar càlculs matricials i càlculs iteratius de grans dimensions. En versions més recents s'observen millores en disminució de temps i recurs de màquina en el procés d'aquest tipus de càlcul. Altres dificultats en la realització

d'aquesta aplicació es basen en la utilització del bloc *Matlab Script* per a establir comunicació entre LabVIEW i Matlab. Aquesta eina no disposa d'una eina de depuració d'errors des de l'entorn LabVIEW, fet que dificulta la seva implementació. El fet que l'aplicació estigui dotada d'una elevada autonomia a nivell de càlcul, permet el funcionament normal de gran part de l'aplicació enfront una fallada de connexions entre LabVIEW i Matlab.

L'aplicació està dotada de flexibilitat alhora d'adaptar-se a configuracions de la planta SBR diferents a l' actual com podria ser un augment de les variables de procés degut a la instal·lació d'una nova sonda. En aquest sentit, l'aplicació permet la seva utilització per al control estadístic multivariable en altres tipologies de processos per lots. L'aplicació també podria modificar-se en certs aspectes per la seva utilització per al control de processos continus.

L'aplicació realitzada podria ser objecte de certes ampliacions que la dotarien d'una major consistència alhora de realitzar qualsevol tipus d'anàlisi. En aquest sentit, es podrien incloure eines que permetessin l'adequació de lots de durada variable per tal de poder-los projectar sobre els models establerts. També es podria dotar l'aplicació d'eines que permetessin l'anàlisi estadístic per a cadascuna de les etapes que conformen el procés de lot i d'aquesta manera obtenir més informació abans de la finalització del mateix.

També es podria dotar l'eina de la implementació necessària per a la creació de nous models PCA des del propi entorn LabVIEW. En aquest cas, l'eina esdevindria completament autònoma a nivell de càlcul. D'aquesta manera, desapareixeria la necessitat d' utilitzar aplicacions externes i l'establiment de connexions amb altres programes.

Girona, gener de 2006

Miquel À. Castro Ortega

8.RELACIÓ DE DOCUMENTS

DOCUMENT 1: MEMÒRIA I ANNEX

MEMÒRIA

ANNEX

ANNEX A: ESTUDI ECONÒMIC

9.BIBLIOGRAFIA

AGUADO D. Aplicación de métodos estadísticos multivariantes para la modelación y la monitorización de un reactor discontinuo secuencial para el tratamiento de aguas residuales. Tesis Doctoral. Departamento de ingeniería hidráulica medio ambiente. Universidad de Valencia. 2005

BALAGUER, M.D., COLPRIM, J., MARTÍN, M., POCH, M., RODRÍGUEZ-RODA, I. Tractament biològic d'aigües residuals urbanes. Programa de doctorat "Tecnologies del Medi Ambient". EQATA. 1998

FERRER, A. . Control estadístico megavariante para los procesos del siglo XXI. *27 Congreso Nacional de Estadística e Investigación Operativa*, Lleida (Espanya). 2003

KOURTI, T.; MACGREGOR, J.F. Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, 28 (4), 409-428.1996

LabVIEW Database Connectivity Toolset User Manual. National Instruments. 2001

LabVIEW Style Guide. National Instruments. 2003

LabVIEW 7.0 User Manual. National Instruments. 2003

LENNOX, J.; ROSEN, C. Adaptive multiscale principal component analysis for on-line monitoring of wastewater treatment. *Wat. Sci. Tech.*, 45 (4-5), 227-235. 2002

MACGREGOR, J.F.; JAECKLE, C. ; KIPARISSIDES, C.; KOUTOUDI, M. Process monitoring and diagnosis by multiblock PLS methods. *AIChE J.*, 40 (5), 826-838.1994

MATLAB 6.5 User Manual. Mathworks. 2003

METCALF Y EDDY . *Ingeniería de aguas residuales. Tratamiento, vertido y reutilización*. 3ª edición. Editorial McGraw-Hill. 2000

MySQL User Manual. MySQL AB. 2004

- NOMIKOS, P.; MACGREGOR, J.F. Monitoring of batch processes using multi-way principal component analysis. *AIChE J.*, 40, 1361-1375.1994
- NOMIKOS, P.; MACGREGOR, J.F. Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics*, 37 (1), 41-59.1995
- OLSSON, G.; NEWELL R. *Wastewater Treatment Systems. Modelling, Diagnosis and Control*. IWA Publishing, London, UK. 1999
- RIERA, R. On line monitoring of a Sequencing Batch Reactor. Master Thesis. Faculteit Bio-Ingenieurswetenschappen. Universiteit Gent. 2005
- ROSEN, C.; LENNOX, J.A. Multivariate and multiscale monitoring of wastewater treatment operation. *Wat. Res.*, 35 (14), 3402-3410. 2001
- RUBIO, M. "Qualitative Trends for Situation Assessment in a SBR Wastewater Treatment Process", Research Report, Department of Electronics, Computer Science and Automatic Control, University of Girona. 2004
- RUIZ, M.L. Multivariate Statistical Process Control for situation assessment of a Sequencing Batch Reactor (SBR). Research Report. Department of Electronics, Computer Science and Automatic Control, University of Girona.2004
- RUIZ, G.; CASTELLANO, M.; GONZÁLEZ, W.; ROCA, E.; LEMA, J.M. Algorithm for steady state detection of multivariate process: Application to wastewater anaerobic digestion process. Proceedings of *2nd International IWA Conference on Automation in Water Quality Monitoring*, AutMoNet, 19-20 April, Vienna, Austria. 2004
- RUIZ, M.L.; COLOMER, J.; RUBIO, M.; MELÉNDEZ, J.; COLPRIM, J. Situation Assessment of a sequencing batch reactor using multiblock MPCA and Fuzzy classification. *BESAI 4th ECAI Workshop on Binding Enviromental Science and Artificial Intelligence*, Valencia, Spain. 2004
- RUSELL, E.L.,CHIANG L.H,BRAATZ R.D. Data-driven techniques for fault detection and diagnosis in chemical processes.Springer-Verlag. 2000

SIMCA-P User Manual.Umetrics. 2005

SIMCA-P+ User Manual. Umetrics. 2005

VIVES,M.T. Sbr technology for wastewater treatment: suitable operational conditions for a nutrient removal.Tesis Doctoral. Lequia.Universitat de Girona. 2005

WESTERHUIS, J.A.; KOURTI, T.; MACGREGOR, J.F. (1999). Comparing alternative approaches for multivariate statistical analysis of batch process data. *J. Chemometrics*, 13, 397-413.

WISE,B., GALLAGHER,N.B., BRO,R., SHAVER J. PLS_Toolbox 3.0 User Manual. Eigenvector Research Incorporated. 2003

WOLD, S.; KETTANEH, N.; FRIDEN, H.; HOLMBERG, A. (1998). Modelling and diagnosis of batch processes and analogous kinetic experiments. *Chemom. Intell. Lab. Syst.*, 44, 331-340.

YOO, C.K.; LEE, D.S.; VANROLLEGHEM P.A. (2004). Application of multiway ICA for on line process monitoring of a sequencing batch reactor. *Wat. Res.*, 38, 1715-1732.

10. GLOSSARI

DBO	Demanda bioquímica d' oxigen
EDAR	Estació depuradora d' aigües residuals
MSPC	Control estadístic multivariable de processos
PCA	Anàlisi de components principals
PCR	Regressió en components principals
PLS	Regressió en mínims quadrats parcials
SBR	Reactor discontinu seqüencial
SPC	Control estadístic de processos
SPE	Error quadràtic de predicció