

BJÖRN HALVARSSON

Applications of Coupling Analysis on Bioreactor Models



UPPSALA
UNIVERSITET

Master of Science Thesis, Engineering Physics Programme,
School of Engineering, Uppsala University, Sweden.

ABSTRACT

This master thesis treats the problem of selecting proper pairings of input and output signals to obtain satisfactory control of multivariable systems. Particularly, multivariable bioreactor systems are considered.

In multivariable systems, interaction between the different parts of the system has to be considered when designing a controller since inappropriate pairings of input and output signals may give unstable systems.

Two measurements that quantify the degree of interaction and give pairing suggestions are the commonly used Relative Gain Array (RGA) and the recently suggested Hankel Interaction Index Array (HIIA). To make a comparison between their ability to give reasonable pairing suggestions these methods were utilized in the analysis of bioreactor models. The considered models were simplified versions of the IAWQ Activated Sludge Model No. 1 that models an activated sludge process configured for nitrogen removal.

It was found that the RGA method was unable to give reasonable pairing suggestions in some cases. This is due to the nearly triangular structure of the transfer function matrices of the considered models. It was also found that an analysis of condition numbers did not fully manage to indicate that the RGA was misleading.

The HIIA method did perform well in all cases if the system transfer function matrices were scaled in a physically relevant way and thereafter low-pass filtered.

Keywords: Coupling analysis, Relative Gain Array, Hankel Interaction Index Array, Gramian based interaction measure, bioreactor models.

© Björn Halvarsson 2003

ISSN 1401-5757
UPTEC F03 024

Acknowledgements

I would herby like to thank my supervisors, Tech. Lic. Pär Samuelsson and Professor Bengt Carlsson for all their help and encouragement during this project. I also want to express my gratitude to all of the other very friendly and helpful members of Systems and Control, Department of Information Technology.

Contents

1	Introduction	1
1.1	Thesis Outline	2
2	Theory	5
2.1	The Relative Gain Array (RGA)	5
2.1.1	Definition	5
2.1.2	Algebraic Properties	6
2.1.3	Pairing Recommendation	7
2.1.4	A Dynamic Extension of the RGA	9
2.1.5	Generalization for Non-Square Plants	9
2.1.6	The Link Between the RGA and the Condition Number	11
2.1.7	Conclusions	15
2.2	Gramian Based Interaction Measure	16
2.2.1	Gramian Fundamentals	16
2.2.2	The Hankel Interaction Index Array (HIIA)	17
3	Description of the Analysed Bioreactor Models	19
3.1	Wastewater Treatment	20
3.2	The Activated Sludge Process	20
3.3	The IAWQ Activated Sludge Model No. 1	23
3.4	The COST Benchmark WWTP	24
3.5	Control of WWTPs	24
3.5.1	Control Handles for Nitrogen Removal	24
3.5.2	Controlled Output Signals for Nitrogen Removal	25
3.6	Simplified ASM1 Models	25
4	Analysis of the Bioreactor Models	29
4.1	Linearising the Models	29
4.2	The Influence of External Carbon Dosage	33
4.3	RGA Analysis of Model (3.1)	34
4.3.1	Steady-state Analysis	34
4.3.2	Dynamic Analysis	35
4.3.3	Conclusions	38

4.4	RGA Analysis of Model (3.3)	38
4.4.1	Steady-state Analysis	38
4.4.2	Dynamic Analysis	40
4.4.3	Conclusions	43
4.5	RGA Analysis of a Modified Version of Model (3.1)	43
4.5.1	Steady-state Analysis	43
4.5.2	Dynamic Analysis	45
4.5.3	Conclusions	49
4.6	HIIA Analysis of Model (3.1)	49
4.7	HIIA Analysis of Model (3.3)	52
4.8	HIIA Analysis of the Modified version of Model (3.1)	52
4.9	General Conclusions	53
4.9.1	The Condition Number as a Warning Signal When the RGA May Not Work	53
4.9.2	Why the RGA Fails	53
4.9.3	The RGA versus the HIIA	55
4.9.4	Practical Implications	56
A	Parameter values	57
B	HIIA Matrices Obtained for the Analysed Bioreactor Mod- els	59
B.1	HIIA Matrices for Model (3.3)	59
B.2	HIIA Matrices for the Modified Version of Model (3.1)	60

Chapter 1

Introduction

Many control systems of today are multivariable. This means that they have both multiple inputs as well as multiple outputs. Such systems are called multiple-input multiple-output (MIMO) systems. Compared to single-input single-output (SISO) systems, the control design for MIMO systems is a bit more elaborate. One reason for this is that different parts of a multivariable system may interact and cause couplings in the system. For example, consider a shower with separate flow control for hot and cold water. This is a MIMO system since the two inputs, the flow of hot water and the flow of cold water, are utilized to control the two outputs (i.e. the flow from the tap and the temperature of the effluent water). Evidently, when changing one of the inputs, both of the outputs will be affected. Hence, there are significant couplings in the system. In other words, interaction occurs if a change in one input affects several outputs.

Often, an easy way to control a fairly decoupled MIMO system is to use a multi-loop strategy, i.e. to separate the control problem into several single-loop SISO systems and then use conventional SISO control on each of the loops, see Kinnaert (1995) and Wittenmark *et al.* (1995). This gives rise to the *pairing problem*:

Which input signal should be selected to control which output signal to get the most efficient control with a low degree of interaction?

In real-life applications the considered MIMO system could be rather complex: In the chemical process industry a complexity of several hundred control loops is not unusual, see Wittenmark *et al.* (1995). The proper pairing selection is thus often not at all obvious. Also, the choice of pairing is crucial since a bad choice may give unstable systems even though each loop separately is stable. This problem could arise due to interaction between the different loops. Generally, the stronger the couplings, the harder it is to obtain satisfactory control performance. Evidently, there is a need for a measurement that can both give some advise when solving the pairing

problem and that also quantifies the level of interaction occurring in the system.

One such measurement is the Relative Gain Array (RGA) developed by Bristol (1966). Here, the RGA considers steady-state properties of the plant and gives a suggestion on how to solve the pairing problem in the case of a decoupled (diagonal) control structure. It also indicates which pairings that should be avoided due to possible stability and performance problems.

Later, a dynamic extension of the RGA was proposed in the literature, see e.g. Kinnaert (1995) for a survey. With this extension, the RGA could be used to analyse the considered plant at any frequency but still only at one single frequency at a time.

Furthermore, the RGA can be generalized for non-square plants and be employed as a screening tool to get a suggestion on what inputs or outputs that should be removed in the case of excess signals, see Skogestad and Postlethwaite (1996).

The above mentioned drawback, as well as other drawbacks such as the inability of the RGA to deal with other control structures than decoupled (diagonal) ones, has nourished the development of other resembling tools. One such example is the Partial Relative Gain (PRG) suggested by Häggblom (1997) that should be able to deal with the pairing problem for larger systems in a more reliable way than the conventional RGA. Other examples are given by Kinnaert (1995) where a survey of interaction measures for MIMO systems can be found.

A somewhat different approach was employed by Conley and Salgado (2000) when considering observability and controllability Gramians to find out how to solve the pairing problem. In this way, the full dynamics of the considered system is incorporated in a single measure.

Recently, Wittenmark and Salgado (2002), refined this work and proposed a new interaction measure, the Hankel Interaction Index Array (HIIA). This measure seems to be able to overcome most of the disadvantages that the RGA possesses.

In this thesis, the RGA and the HIIA will be employed in the selection of input–output signal pairings for some rather demanding MIMO linear systems, namely bioreactor systems. The RGA suggestions will be compared with those of the HIIA and with results obtained from physical insights of the considered systems.

1.1 Thesis Outline

Chapter 2 deals theoretically with different interaction measures for MIMO systems (the RGA and the HIIA). Other important measures, such as the condition number, and the issue of selecting a proper scaling of the considered system are also treated.

In Chapter 3 the objective is to describe the models that will be analysed in Chapter 4. The reader is also introduced to the topic of wastewater treatment from a Swedish point of view.

Finally, in Chapter 4 the RGA method and the HIA method are used to analyse the bioreactor models.

Chapter 2

Theory

2.1 The Relative Gain Array (RGA)

2.1.1 Definition

The most widely used interaction measure for MIMO linear systems so far, is the Relative Gain Array (RGA) introduced by Bristol (1966). The RGA for a quadratic plant is given by

$$\text{RGA}(G) = G(0) \cdot * (G(0)^{-1})^T \quad (2.1)$$

where $G(0)$ is the steady-state transfer function matrix and “ $\cdot *$ ” denotes the Hadamard or Schur product (i.e. elementwise multiplication). To see that the RGA provides information relevant in the selection of input-output pairing, the RGA will be derived following Bristol (1966) (see also Kinnaert (1995) and Skogestad and Postlethwaite (1996)).

Consider a square plant with transfer function matrix $G(s)$ with elements $[G(s)]_{ij} = g_{ij}(s)$ and denote input j by u_j and output i by y_i . In the sequel, the Laplace-variable s will be omitted for $G(s)$ and $G_{ij}(s)$.

The *open loop gain* between input u_j and output y_i , $g_{ij}(0)$, can be obtained by letting all other inputs except u_j , be constant:

$$g_{ij}(0) = \left(\frac{\partial y_i}{\partial u_j} \right)_{u_k = \text{constant}, \forall k \neq j} \quad (2.2a)$$

Similarly, the *closed loop gain*, \hat{g}_{ij} is obtained, by holding all outputs except y_i constant, using feedback control:

$$\hat{g}_{ij}(0) = \left(\frac{\partial y_i}{\partial u_j} \right)_{y_k = \text{constant}, \forall k \neq i} \quad (2.2b)$$

Now, ideally, if no interaction between the loops are present, the gain between input u_j and output y_i would remain the same when the other loops

are closed, so the *relative gain* $g_{ij}/\hat{g}_{ij} = 1$. On the other hand, if there is loop interaction in the system, g_{ij} and \hat{g}_{ij} will differ. Hence, the quotient

$$\lambda_{ij}(0) = g_{ij}(0)/\hat{g}_{ij}(0)$$

can be used as an interaction measure, and a relative gain array (RGA) with elements given by $\lambda_{ij}(0)$ can be formed. $y = Gu$ gives $g_{ij} = [G]_{ij}$, and similarly, $u = G^{-1}y$ gives

$$\frac{1}{\hat{g}_{ij}(0)} = \left(\frac{\partial u_j}{\partial y_i} \right)_{y_k = \text{constant}, \forall k \neq i} = [G^{-1}(0)]_{ji}$$

Each element in the RGA can therefore be calculated using

$$\lambda_{ij}(0) = \frac{g_{ij}(0)}{\hat{g}_{ij}(0)} = [G(0)]_{ij}[G^{-1}(0)]_{ji} \quad (2.3)$$

Hence, the whole RGA matrix can directly be calculated using (2.1). Clearly, to minimize undesired interactions, pairings corresponding to a RGA-element, $\lambda_{ij}(0)$, as close to one as possible should be selected.

2.1.2 Algebraic Properties

The RGA possesses several useful algebraic properties. Some of the most important are listed below.

Property 1 If rows and columns are permuted in the transfer function matrix, G , then the rows and columns in the RGA are permuted in the same way.

Property 2 The division in (2.3) with \hat{g}_{ij} ensures the RGA to be *scaling independent*, i.e.

$$\text{RGA}(G) = \text{RGA}(S_1GS_2) \quad (2.4)$$

where S_1 and S_2 are diagonal scaling matrices of the same dimension as G .

Property 3 The division in (2.3) *normalizes* the RGA, in such a way that the numerical sum of each column, as well as the numerical sum of each row, in the RGA is equal to one, i.e. for a $n \times n$ matrix RGA

$$\sum_{i=1}^n \lambda_{ij} = \sum_{j=1}^n \lambda_{ij} = 1 \quad (2.5)$$

Property 4 If the transfer function matrix, G , is diagonal or triangular, and if the rows in the transfer function matrix are permuted to get nonzero elements along the diagonal in the case of a triangular G ,

then the RGA equals the identity matrix.¹ The first case gives a *decoupled* system and the second case gives a *one way interactive* system, (Grosdidier *et al.*, 1985). Thus the RGA does not differ between diagonal and certain triangular plants. Clearly, this is a drawback.

Property 5 For the case of a 2×2 plant, G , with nonzero elements only, the following holds: (a) If the number of positive elements in $G(0)$ is *odd* then $\lambda_{ij} \in (0, 1)$; (b) If the number of positive elements in $G(0)$ is *even* then $\lambda_{ij} \in (-\infty, 0) \cup (1, \infty)$, (Grosdidier *et al.*, 1985).

Property 1, 2 and 4 can very easily be shown by using the definition of the RGA in equation (2.1). Additional properties as well as proofs to some of them can be found in e.g. Skogestad and Postlethwaite (1996).

2.1.3 Pairing Recommendation

Due to the normalization property (Property 3) for a 2×2 system a symmetric matrix is obtained:

$$\text{RGA}(G(0)) = \begin{bmatrix} \lambda & 1 - \lambda \\ 1 - \lambda & \lambda \end{bmatrix} \quad (2.6)$$

Depending on the value of λ , five different cases occur, (Kinnaert, 1995):

$\lambda = 1$: This is the ideal case when no interaction between the loops is present. The pairing should be along the diagonal, i.e. $u_1 - y_1$ and $u_2 - y_2$;

$\lambda = 0$: This is the same situation as above, except that now the suggested pairing is along the anti-diagonal, i.e. $u_1 - y_2$, $u_2 - y_1$

$0 < \lambda < 1$: This case is not desirable since the gain increases (i.e. \hat{g}_{ij} increases) when the loops are closed, hence, there is interaction;

$\lambda > 1$: Now, the gain decreases when the loops are closed. This situation is therefore also undesirable.

$\lambda < 0$: This situation corresponds to the worst case scenario since now, even the sign changes when the loops are closed and this is highly undesirable.

The conclusion is that u_1 should only be paired with y_1 when $\lambda \geq 0.5$, otherwise it should be paired with y_2 . For the higher-dimensional case, the rule should be to choose pairings that have a RGA-element close to one. Negative pairings should definitely be avoided.

¹For anti-diagonal and triangular systems with nonzero elements along the anti-diagonal, the RGA equals the anti-identity matrix with zeros in all positions except along the anti-diagonal.

Stability Considerations

As concluded above, a pairing with negative RGA-element is not desirable since this will cause the gain to change sign when the other loops are closed. Further investigations made by Grosdidier *et al.* (1985) have shown that pairings with negative RGA-elements may cause instability. Assuming a system with a feedback configuration consisting of a compensator, $K(s)$, with the structure

$$K(s) = \frac{k}{s}C(s) \quad (2.7)$$

and also assuming $C(s)$ to be diagonal, $G(s)C(s)$ to be proper and $k > 0$, Grosdidier *et al.* (1985) showed that if any diagonal RGA-element, λ_{jj} , is negative then at least one of the following properties applies to the closed loop system:

1. The closed loop system is unstable;
2. Loop j is unstable when all other loops are opened;
3. The closed loop system is unstable when loop j is removed (i.e. open).

In the 2×2 case, the rows of the transfer function matrix can always be permuted to get a system without negative diagonal elements in the RGA matrix. This is clearly seen from equation (2.6).

Hence, with the exception for the 2×2 case, it is not possible to obtain a good decentralized controller in the presence of a negative RGA-element on the principal diagonal, (Maciejowski, 1989). Instead, one may consider a cross-coupled controller structure.

Observe also, that a negative diagonal element in the RGA only gives a *sufficient* (but not *necessary*) condition for instability.

The next result to consider to avoid pairings that may lead to instability is Niederlinski's theorem. If the same compensator structure as above is used and it is assumed that $C(s)$ is diagonal and $G(s)$ (with elements g_{ij} , $i, j = 1, \dots, n$) is stable, and each individual control loop remains stable when any of the other loops are opened², then a sufficient condition for instability is

$$\det G(0) / \prod_{i=1}^n g_{ii}(0) < 0 \quad (2.8)$$

As pointed out by Kinnaert (1995), it is advisory to combine the use of the RGA with a check of the Niederlinski theorem to see if the chosen pairings *may* be alright regarding stability. Note that since (2.8) is only a *sufficient* condition for instability, there might be some G which do not fulfill the test but still causes instability.

²For the case of a 2×2 plant this assumption can be relaxed as shown by Chiu and Arkun (1991).

The RGA-number

To characterize the chosen pairings with a scalar measurement, the *RGA-number* can be used:

$$\text{RGA-number} = \|\text{RGA}(G) - I\|_{\text{sum}} \quad (2.9)$$

where the sum matrix norm is defined as

$$\|A\|_{\text{sum}} = \sum_{i,j} |a_{ij}| \quad (2.10)$$

if the $n \times n$ matrix A has elements a_{ij} , $i, j = 1, \dots, n$. Evidently, the RGA-number measures how dominant the diagonal in the RGA is.³ A proper pairing choice should thus have a RGA-number close to zero.

2.1.4 A Dynamic Extension of the RGA

Bristol (1966) only used the plant steady-state gain, $G(0)$, when calculating the RGA. The reason for this, was probably that in the process industry this steady-state measure is often far more easy to obtain than the corresponding dynamic measure, $G(i\omega)$, (Maciejowski, 1989). However, later, a dynamic extension of the RGA was proposed (see e.g. Kinnaert (1995) for a survey):

$$\Lambda(G(i\omega)) = G(i\omega) * (G(i\omega)^{-1})^T \quad (2.11)$$

This definition is obviously the same as for the original RGA, except that now the plant gain, G , is allowed to be measured at any frequency ω . Not surprisingly, this dynamic version of the RGA possesses the same properties as the steady-state RGA. Therefore, in the sequel, both of these RGA-versions will be denoted by Λ .

When analyzing a system it is advisory to use this dynamic RGA and hence study the behaviour of $\Lambda(G)$ in the interesting frequency range. As pointed out by Skogestad and Postlethwaite (1996), to avoid instability it is often enough to require $\Lambda(G)$ to be near the identity matrix at the crossover-frequency. But, of course, a pairing that results in negative RGA-elements should not be tolerated for any frequency of interest.

2.1.5 Generalization for Non-Square Plants

The RGA can be generalized for use with $m \times l$ non-square plants, H , by replacing the normal matrix inverse by the pseudo-inverse (Moore-Penrose inverse)⁴, denoted by " \dagger ". Then, the RGA will be given by

$$\Lambda(H) = H * (H^\dagger)^T \quad (2.12)$$

³Before calculating the RGA-number it is assumed that the rows of the RGA are permuted to get the pairing suggestions along the diagonal.

⁴The pseudo inverse of a matrix A is given by $(A^H A)^{-1} A^H$ or $A^H (A A^H)^{-1}$ if the inverse in those expressions exists (Glad and Ljung, 1997). A^H is the transposed and complex conjugated A .

For a non-square plant, the RGA will only partly have the same properties as the RGA for a square plant. Especially, if the plant either has full column rank or full row rank then, for the two cases respectively, the following will apply (Skogestad and Postlethwaite, 1996):

$$\sum_{j=1}^m \lambda_{ij} = \|e_i^T U_r\|_2^2 \quad (2.13a)$$

$$\sum_{i=1}^l \lambda_{ij} = \|e_j^T V_r\|_2^2 \quad (2.13b)$$

where e_i and e_j are unit column vectors of length l and m , respectively, with value one at position i and j , respectively, and zeros in every other position; U_r is a matrix containing the first r *output* singular vectors for H , and V_r is a matrix containing the first r *input* singular vectors for H .⁵ Thus, $e_i^T U_r$ can be seen as the projection of an unit output on the effective output space of H spanned by the columns in U_r , and in the same way, $e_j^T V_r$ can be seen as the projection of an unit input on the effective input space of H spanned by the columns of V_r .

Often, the selection of inputs and outputs from a large group of candidates, can be tedious. However, if the above results, (2.13a) and (2.13b), are used it is obvious that the RGA is a very efficient tool in the selection of inputs and outputs. If the desire is to remove some input candidates, then simply calculate the RGA for the "full" transfer matrix containing every input and output candidate, and then calculate the column sums of the RGA. According to (2.13b), those columns with sums much lower than one will correspond to inputs with a low degree of influence on the system, and thus, they can be dismissed. Similarly, too many output candidates can be dealt with by considering the row sums of the RGA and then, according to (2.13a), remove outputs corresponding to low row sums.

Note however that, of course, a physical understanding of the plant should be the main guidance in the selection of what signals that should

⁵Input and output singular vectors can be obtained by performing a singular value decomposition (SVD) of H as follows

$$H = U \Sigma V^H = U_r \Sigma_r V_r^H \quad (2.14)$$

where Σ_r is a diagonal square matrix consisting of the first $r = \text{rank}(H)$ non-zero singular values, U_r consists of the first r columns of U and finally, V_r consists of the first r columns of V . As concluded by Skogestad and Postlethwaite (1996), the columns in V_r represent the input directions that can affect the outputs, and similarly, the columns in U_r represent the output directions that can be affected by the inputs. If the SVD of H in (2.14) is rewritten as

$$H V = U \Sigma \quad (2.15)$$

the reason for the columns in V and U to be called *input* singular vectors and *output* singular vectors, respectively, is found.

be used. This also applies to the pairing problem, see e.g. Maciejowski (1989). Also, note that the scaling independency is lost for a non-square plant: For the case with more inputs than outputs, the RGA depend on the *input* scaling, and if there are more outputs than inputs, the RGA will depend on the *output* scaling (Skogestad and Postlethwaite, 1996).⁶

2.1.6 The Link Between the RGA and the Condition Number

As previously seen in Section 2.1.3, the RGA may reveal information regarding possible control difficulties and also information about the plant's robustness. This kind of information is linked to the condition number of the system. In this section, the condition number is defined and then, the main results presented in the literature concerning the RGA and its link to the condition number and robustness are briefly discussed.

The Condition Number

The condition number of a matrix measures, loosely speaking, the difficulty of inverting the matrix or equivalently, the difficulty to numerically compute the solution to the linear equation

$$Ax = b \quad (2.16)$$

where A is a matrix and, x and b column vectors of proper length; x is the unknown vector. Now assume that A is nonsingular and is subject to perturbations δA . Then obviously, the unknown vector x is also subject to perturbations, δx , since $x = A^{-1}b$ according to (2.16). The following relation between the relative uncertainty in A and x can then be shown (Grosdidier *et al.*, 1985):

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\delta A\|}{\|A\|} \quad (2.17)$$

The condition number is then defined as

$$\gamma(A) = \|A\| \cdot \|A^{-1}\| \quad (2.18)$$

and from (2.17) it can be seen that it represents the largest number by which a relative uncertainty in A will be amplified and transmitted to the solution x . If the Euclidean matrix norm (the induced 2-norm, i.e. "the maximum singular value norm") is used, the condition number will be given by

$$\gamma(A) = \|A\|_e \cdot \|A^{-1}\|_e = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \quad (2.19)$$

⁶Input scaling represents *post*multiplying with a diagonal scaling matrix, and output scaling represents *pre*multiplying with a diagonal scaling matrix.

where $\|A^{-1}\|_e = 1/\sigma_{\min}(A)$ is used. If the condition number for A is "large" (typically > 10 according to Kinnaert (1995)) then A is said to be *ill-conditioned*.

For MIMO linear systems, the gain at a given frequency, ω , will be bounded by the smallest and largest singular values for the process transfer function matrix G , see e.g. Glad and Ljung (1997) or Skogestad and Postlethwaite (1996). Thus, if the condition number is large, this span will be large and the process model will show high directionality, i.e. the actual gain will highly depend on the direction of the input vector (Waller *et al.*, 1994).

As shown by Waller and Waller (1995) a high condition number may lead to control difficulties when using an inverting controller. This can easily be understood since a high condition number simply means that the inverse is "hard" to find. Contrary to the RGA, the condition number is scaling dependent and thus, to draw any useful conclusions, first a proper scaling of the model has to be selected.

Scaling

To obtain a model that is numerically easier to overview, the model can be scaled. This can be done by introducing scaled variables given by

$$u = D_u^{-1}u^o \quad (2.20a)$$

$$y = D_y^{-1}y^o \quad (2.20b)$$

where the original model is given by

$$y^o(t) = G^o(p)u^o(t) \quad (2.21)$$

and the superscript "*o*" denotes the original (or physical) variables; $G^o(p)$ denotes the original transfer function matrix between output $y^o(t)$ and input $u^o(t)$; D_u and D_y are diagonal scaling matrices.⁷ Thus the scaled model is given by

$$G(s) = S_1 G^o(s) S_2 \quad (2.22)$$

with the scaling matrices $S_1 = D_y^{-1}$ and $S_2 = D_u$.

The choice of a "proper scaling" is not at all easy since there are no general scaling procedures (Waller and Waller, 1995). The scaling should be done carefully since a bad scaling may make the model ill-conditioned even though the plant itself is well-conditioned. Besides, there are different opinions in the literature on what a "proper scaling" is. Some authors (as Grosdidier *et al.* (1985)) seems to be of the opinion that the scaling that gives the minimized condition number should be used while others favorites

⁷Normally this model is expanded to include disturbances and errors as well, see e.g. Glad and Ljung (1997) or Skogestad and Postlethwaite (1996).

a scaling that takes physical properties in the plant into account. Waller and Waller (1995) and Waller *et al.* (1994) discuss this ambiguity thoroughly and to bring some clarity, a "refinement" of the definition of directionality is given: The main idea is to use different scalings when considering stability aspects and when considering performance aspects. For stability considerations Waller and Waller (1995) shows that a proper scaling would be the one that results in the minimized condition number, γ_{min} , i.e. scaling independent measures such as γ_{min} and RGA will provide useful information. On the other hand, a scaled model, with scaling matrices chosen according to the relative importance of the different variables, will be adequate when investigating performance aspects. Thus, when a proper scaling is used, the condition number may give insights that the RGA (or γ_{min}) does not.

The Minimized Condition Number

To obtain the minimized condition number the scaling matrices, S_1 and S_2 , are chosen according to

$$\gamma_{min}(G) = \min_{S_1, S_2} \gamma(S_1 G S_2) \quad (2.23)$$

As shown by Grosdidier *et al.* (1985) γ_{min} are closely related to the RGA. For the case of a 2×2 plant, G , Grosdidier *et al.* (1985) showed that the minimized condition number is given by

$$\gamma_{min} = \|\Lambda(G)\|_1 + \sqrt{\|\Lambda(G)\|_1^2 - 1} \quad (2.24)$$

where the 1-norm is defined as

$$\|\Lambda\|_1 = \max_j \sum_{i=1}^m |\lambda_{ij}| \quad (2.25)$$

i.e. "the maximum column sum". It can also be shown that γ_{min} is bounded by $\|\Lambda(G)\|_1$ according to

$$\gamma_{min} \leq 2\|\Lambda(G)\|_1 \quad (2.26)$$

with equality when $\|\Lambda(G)\|_1 \rightarrow \infty$.

For larger quadratic systems the following conjecture is valid (Grosdidier *et al.*, 1985):

$$\gamma_{min} \leq 2 \max(\|\Lambda(G)\|_1, \|\Lambda(G)\|_\infty) \quad (2.27)$$

where the ∞ -norm is defined as

$$\|\Lambda\|_\infty = \max_i \sum_{j=1}^m |\lambda_{ij}| \quad (2.28)$$

i.e. "the maximum row sum". The work of finding γ_{min} by means of optimization theory is often rather tedious and therefore it is handy to first calculate the RGA and then use (2.24), (2.26) and (2.27).

For the case of a 2×2 plant, using equation (2.24) and the definition of the RGA, equation (2.1), it is very easy to prove that diagonal and triangular systems and also systems with an odd number of positive elements (use Property 5) always have $\gamma_{min} = 1$. This leads Grosdidier *et al.* (1985) to conclude that these kind of plants are well-behaved with no sensitivity problems. However, as shall be seen in the next section, $\gamma_{min} = 1$ does not always imply a system that behaves well.

Robustness and the RGA

Another way to see that large RGA elements may cause controllability problems is to consider robustness aspects. Robustness is the ability for a system to remain stable when model errors are introduced. As shown by Grosdidier *et al.* (1985), the following relations between relative changes in RGA-elements and in elements of the transfer function matrix, G , and in elements in the inverse of the transfer function matrix, exist:

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = (1 - \lambda_{ij}) \frac{dg_{ij}}{g_{ij}} \quad (2.29a)$$

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = \frac{\lambda_{ij} - 1}{\lambda_{ij}} \frac{d\tilde{g}_{ji}}{\tilde{g}_{ji}} \quad (2.29b)$$

where $\tilde{g}_{ij} = [G^{-1}]_{ij}$. From (2.29a) it is concluded that a large $|\lambda_{ij}|$ results in a large sensitivity to errors in g_{ij} . Similarly, from (2.29b) it is seen that for large $|\lambda_{ij}|$ the errors in λ_{ij} and g_{ij} are approximately equal. If (2.29a) and (2.29b) are combined the result is

$$\frac{d\tilde{g}_{ji}}{\tilde{g}_{ji}} = \lambda_{ij} \frac{dg_{ij}}{g_{ij}} \quad (2.30)$$

and thus, a large λ_{ij} will cause a small change in g_{ij} to result in a large change in \tilde{g}_{ji} . Hence, large RGA-elements are undesirable.

However, small RGA-elements do not always imply good robustness. An example of this is the case of a triangular plant. According to Property 4 these plants always have a RGA that equals the identity matrix⁸ and γ_{min} will be 1. But as mentioned by both Chien *et al.* (1992) and Morari and Zafriou (1989) the plant may still be sensitive to uncertainties.

If a diagonal plant with a transfer function matrix given by

$$G = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \quad (2.31)$$

⁸The rows in the transfer function matrix are assumed to be permuted to get nonzero elements along the diagonal.

where α is a complex number, is considered then it can be shown that if α is large (compared to the diagonal elements) the "worst case deviation" given by Chien *et al.* (1992) will be large. Obviously, neither the RGA nor the minimized condition number is able to indicate this. A triangular plant of the form in (2.31) with a large off-diagonal element will have a large condition number, and thus, when using the RGA, it is advisory to also calculate the condition number to make sure the information in the RGA is valid (Kinnaert, 1995). Thus, in this case, the best information will be given by the condition number when the system is properly scaled.

For a more detailed discussion on the scaling choice, robustness and ill-conditionedness, see Waller and Waller (1995).

2.1.7 Conclusions

As previously seen, the RGA provides a very simple way of characterizing interactions present in a MIMO linear system. The RGA gives a suggestion on how to pair the input and output signals if a decentralized controller is intended to be used. It may also give warnings by means of large RGA-elements when there may be stability and robustness problems. In the selection of suitable input and output signals, the generalized version of the RGA for non-square plants is also very useful.

However, the RGA suffers from at least two main disadvantages as pointed out by Birk (2002), namely that the use of RGA requires:

1. A decoupled (i.e. diagonal) controller structure;
2. Zero steady-state control error.⁹

The requirements above certainly reduce the freedom of choice when designing a controller using the RGA.

In some situations, the RGA also fails to give reliable information: In the case of a triangular plant, the RGA does not indicate the presence of couplings through off-diagonal elements. In Kinnaert (1995) it is mentioned that some authors do not regard this as being a drawback since the RGA still gives the *best possible* decoupled controller. This is certainly true, but if the objective is to find the best possible controller among all controller structures – MIMO controllers included – then this feature of the RGA is a clear drawback. Another drawback is that the RGA only considers one separate frequency; it would of course be better to have an interaction measure that considers information given by all of the interesting frequencies.

⁹In Grosdidier *et al.* (1985) it is shown for a 2×2 plant, that one inherent requirement for the RGA to work properly, is the use of an integrating controller when closing the loops.

2.2 Gramian Based Interaction Measure

In the previous section it was seen that the RGA suffers from some important disadvantages, such as its inability to deal properly with triangular plants, and its limitation to consider each frequency separately. To deal with the first disadvantage, modified versions of the RGA, like the Partial Relative Gain (Hägglblom, 1997) and the Block Relative Gain have been proposed. However, neither of these can be considered as dynamic, since still, they only consider what happens in one single frequency. Then, Conley and Salgado (2000) proposed a new interaction measure based on Gramians, able to handle both of the above mentioned pitfalls. Recently, a modified version of the interaction measure was suggested by Wittenmark and Salgado (2002) where the Hankel norm is used.

2.2.1 Gramian Fundamentals

In order to understand the Gramian based interaction measures, this section provides some fundamental definitions on Gramians.

Consider a linear system, with inputs given by the $n \times 1$ vector $u(t)$ and outputs given by the $p \times 1$ vector $y(t)$. Given the state vector $x(t)$, the system can be described as a state space realization

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t)\end{aligned}\tag{2.32}$$

where A , B , C and D are real matrices of dimension $n \times n$, $n \times m$, $p \times n$ and $p \times m$, respectively. Given an initial state $x(t_0)$ and an input $u(t)$, then for $t \geq t_0$ the solution to (2.32) is given by (Skogestad and Postlethwaite, 1996)

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau\tag{2.33}$$

For a system to be state controllable it is required that there exists an input $u(t)$ such that $x(t_1) = x_1$ for some final state x_1 . It can be verified using (2.33) that one input that satisfies this criterion is given by (Skogestad and Postlethwaite, 1996)

$$u(t) = -B^T e^{A^T(t_1-t)}W_c(t_1)^{-1}(e^{At_1}x_0 - x_1)\tag{2.34}$$

where $W_c(t)$ is a *Gramian* defined as

$$W_c(t) = \int_0^t e^{A\tau}BB^T e^{A^T\tau}d\tau\tag{2.35}$$

Then, for the solution in (2.34) to exist, the inverse of $W_c(t)$ needs to exist, i.e. $W_c(t)$ must have full rank for every $t > 0$. For a stable system it is

enough to require $W_c(\infty)$ to have full rank. Thus, define the *controllability Gramian*, Γ_c , as

$$\Gamma_c = \int_0^\infty e^{A\tau} B B^T e^{A^T \tau} d\tau \quad (2.36)$$

and require this to have full rank for the system to be state controllable.

Similarly, a stable system will be *state observable* if the observability Gramian, Γ_o , defined as

$$\Gamma_o = \int_0^\infty e^{A^T \tau} C^T C e^{A\tau} d\tau \quad (2.37)$$

has full rank. These Gramians can be obtained by solving the following continuous Lyapunov equations (Conley and Salgado, 2000):

$$A\Gamma_c + \Gamma_c A^T + B B^T = 0 \quad (2.38a)$$

$$A^T \Gamma_o + \Gamma_o A + C^T C = 0 \quad (2.38b)$$

The rank of Γ_c is the dimension of the controllable subspace corresponding to the given system (2.32), and correspondingly, the rank of Γ_o is the dimension of the observable subspace of the same system.

A more thorough discussion on Gramians as well as precise definitions of state controllability and state observability can be found in many text books such as Skogestad and Postlethwaite (1996) and Rugh (1996).

2.2.2 The Hankel Interaction Index Array (HIIA)

Obviously, the Gramians presented above are measures of how hard it is to observe and control the states of the given system. As shown by Conley and Salgado (2000) and Wittenmark and Salgado (2002) it is possible to split the system given by (A, B, C, D) into fundamental subsystems (A, B_j, C_i, D_{ij}) where B_j is the j :th column in B , C_i is the i :th row in C and D_{ij} is the (i, j) :th element of D . Then for each of these, the Lyapunov equations in (2.38) can be solved to obtain the controllability and observability Gramians for each subsystem. Hence, solve

$$A\Gamma_c^{(j)} + \Gamma_c^{(j)} A^T + B_j B_j^T = 0 \quad (2.39a)$$

$$A^T \Gamma_o^{(i)} + \Gamma_o^{(i)} A + C_i^T C_i = 0 \quad (2.39b)$$

The controllability and observability Gramians for the full system will then be the sum of the Gramians for all the subsystems.

Unfortunately, both the controllability and the observability Gramian will depend on the chosen state space realization. However, the eigenvalues of the product of these will not. The Hankel norm for a system with transfer function $G(s)$ is defined as

$$\|G(s)\|_H = \sqrt{\lambda_{max}(\Gamma_c \Gamma_o)} = \sigma_1^H \quad (2.40)$$

where σ_1^H is the maximum Hankel singular value. Hence, this measure is invariant with respect to the state space realization and it is therefore well suited as a combined measure for controllability and observability. In Wittenmark and Salgado (2002) it is shown that the Hankel norm of $G(s)$ given in (2.40) also can be interpreted as a gain between past inputs and future outputs. Then, if the Hankel norm is calculated for each fundamental subsystem and arranged in a matrix Σ_H given by

$$[\Sigma_H]_{ij} = \|G_{ij}(s)\|_H \quad (2.41)$$

this matrix can be used as an interaction measure. In Wittenmark and Salgado (2002) a normalized version, the Hankel Interaction Index Array (HIIA), is proposed:

$$[\Sigma_H]_{ij} = \frac{\|G_{ij}(s)\|_H}{\sum_{kl} \|G_{kl}(s)\|_H} \quad (2.42)$$

The pairing rule will be the same as for the RGA, i.e. for each row (i.e. output) select the largest element. Even though not directly stated by Wittenmark and Salgado (2002), expected performance for different control structures can certainly be compared by summing the elements in Σ_H : Clearly, due to the normalization, the aim is to find the simplest controller structure that gives a sum as near one as possible. In the slightly different interaction measure proposed by Conley and Salgado (2000) this is used.

When $G_{ij} = 0$ the Gramian product, $\Gamma_c^{(j)} \Gamma_o^{(i)}$, will be zero and so will the corresponding element in the matrix Σ_H . This implies that the structure of Σ_H will be the same as the structure of G and thus, non-diagonal elements will not be hidden as in the case of the RGA. Hence, the HIIA can also be used to evaluate other controller structures than just the diagonal ones.

As shown in examples given by Wittenmark and Salgado (2002) the HIIA outperforms the RGA when dealing with systems that have interactions with non-monotonic frequency behaviour. The reason for this is that the full dynamics of the system will be taken into account when using Gramians. If the objective is to study the interactions in a specific frequency range only, then the original system can be filtered before the HIIA is calculated, see Wittenmark and Salgado (2002).

However, the HIIA may not give reliable pairing proposals when the state space realization has a nonzero direct term D . The reason is that D is not used when computing the Gramians using (2.39).

It should also be noted that both the HIIA and the RGA requires a linear system model. Since most physical processes are nonlinear even better performance can be expected if a nonlinear control strategy is applied. Hence, there is a need for an interaction measure that can deal with nonlinear models. Progress towards this has been made by Gray and Scherpen (1998) by the development of non-linear Gramian extensions. But beforehand, the linear HIIA must be thoroughly tested and perhaps also further developed.

Chapter 3

Description of the Analysed Bioreactor Models

In the complex process of wastewater treatment, many different cause-effect relationships exist, and therefore, there are many possible choices of input and output signals, see Olsson and Jeppsson (1994). This could hence motivate the study of wastewater treatment plant (WWTP) models with respect to the selection of input and output signals.

When treating wastewater, the aim is to reduce as much as possible of the undesired constituents such as organic matter, nitrogen and phosphorous. This is commonly done using wastewater treatment plants. In a WWTP several multivariable processes are going on simultaneously. These processes often need to be properly controlled in order to maintain the concentrations of undesired constituents in the outlet water within the legislated limits. As the public awareness of environmental issues increases, the environmental legislation becomes stricter, and thus, the requirements on WWTPs become even harder to fulfill. The used control strategies need then to be as efficient as possible, see e.g. Olsson and Newell (1999). Therefore, models of the WWTP processes are interesting to study with respect to the choice of e.g. control structure. An example of such models are the *bioreactor models*.

From a theoretical point of view, the bioreactor models are non-linear multivariable systems that may contain a significant degree of couplings. Hence, this also gives an interesting opportunity to test the performance of the methods for input-output pairing selection discussed in the previous chapter.

This chapter gives a description of the bioreactor models that will be analysed in the next chapter, as well as a short introduction to wastewater treatment.

3.1 Wastewater Treatment

Until some time during the 19:th century, the activity of man had not affected the environment to any appreciable extent. When the industrial revolution came, a rapidly increased standard of living as well as a substantially population growth followed. The society became more and more urbanized and the problem of taking care of the human waste products and waste disposal became a serious (hygienic) problem. The introduction of the water closet solved the problem locally, but only locally, since the problem was instead moved to the surrounding environment with an increased load on the recipients (e.g. lakes and rivers). This could not be handled by the recipients without heavily disturbed local ecosystems. The degradation of organic material present in the wastewater, consumes loads of oxygen and the recipient will thus suffer from lack of oxygen after some while. Even if most of the organic matter is removed before the wastewater reaches the recipient, chemical compounds such as phosphorous and nitrogen are still present, and may cause eutrophication (i.e. over-fertilization). Eventually, this will also result in a lack of oxygen. Thus the aim of wastewater treatment should be to remove both the content of organic matter and suspended solids as well as the content of nitrogen and phosphorous to a reasonable extent.

In the beginning of the 20:th century, the first wastewater treatment plants were introduced in Sweden. They were simple plants using only a *mechanical treatment step*. This step could consist of a grid and a sand filter to remove larger objects and particles. Then, in the late 1950's, step two, *the biological treatment step*, was introduced. Hereby, microorganisms (e.g. bacteria) are used to remove organic matter present in the incoming wastewater. Later, in the 1970's, a third step, *the chemical treatment step*, was employed to reduce the content of phosphorous. Nowadays, the biological step are also utilized to reduce the content of nitrogen and phosphorous. A general WWTP, consisting of the above mentioned steps, is given in Figure 3.1.

The sludge also needs to be treated. The main procedures are depicted in Figure 3.1. In the thickening procedure, the sludge is concentrated. Then, the sludge is stabilized in order to reduce odor and pathogenic content. Finally, the moisture content of sludge can be reduced by the use of dewatering. For a description of how to practically realize these steps, see e.g. Metcalf and Eddy Inc. (1991).

3.2 The Activated Sludge Process

The biological treatment step can be realized in several different ways. One of the most common is the *activated sludge process* where activated sludge, i.e. microorganisms (mainly bacteria), is employed to degrade (i.e. oxidize)

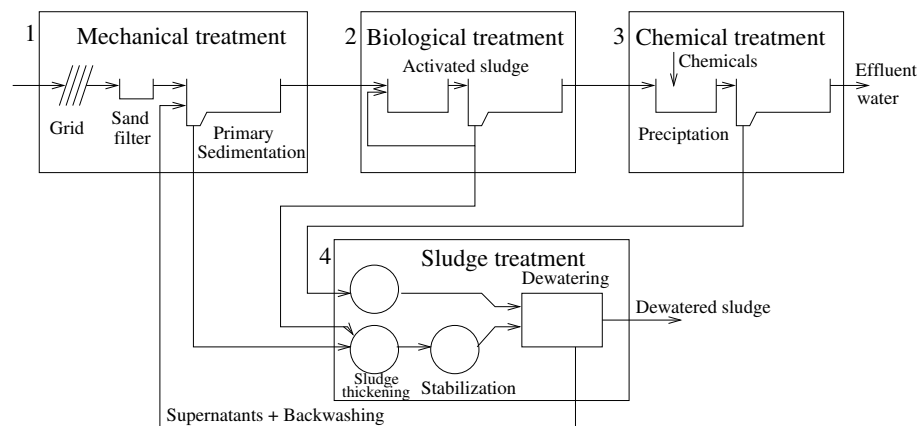


Figure 3.1: A general WWTP (Kommunförbundet, 1988).

organic material. The basic set-up consists of an aerated basin where oxygen is added by blowing air into the water, and a settler tank, see Figure 3.2. In the aerated basin, the bacteria degrade the incoming organic material while consuming oxygen. In this way the microorganisms fulfill their need of energy and as a result bacterial growth will occur. Together with dead microorganisms and other particulate material, the living microorganisms form sludge. To separate the sludge from the purified water, a *settler*, where the sludge settles, can be used directly after the aerated tank. Since the amount of microorganisms needs to be kept at a high level, some sludge is often recirculated as shown in Figure 3.2, while the rest is removed as excess sludge. With the excess sludge, some nitrogen (and phosphorus) is removed, but still far too much remains.

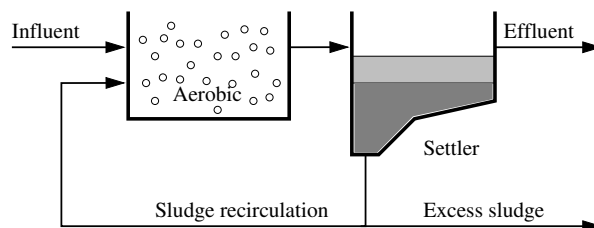
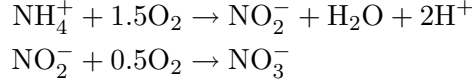


Figure 3.2: A basic activated sludge process with an aerated basin and a settler.

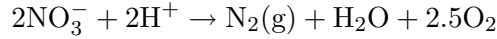
However, if the activated sludge process is extended to consist of both aerated and non-aerated (anoxic) basins, then bacteria may be employed for efficient nitrogen removal. In the aerated basins, bacteria oxidize ammonium

to nitrate in a two-step process called *nitrification* (Lindberg, 1997):



For these processes to occur, the concentration of dissolved oxygen (DO) must be sufficiently high and a long sludge age (the average time each particle stays in the system) is required due to slow bacteria growth.

In the anoxic tanks, another type of bacteria is employed in the *denitrification* process, described by



i.e., the bacteria change nitrate into nitrogen gas using the oxygen in the nitrate ions. However, no dissolved oxygen should be present for this process to take place, instead, a sufficient amount of readily biodegradable substrate is needed. Hence, together, nitrification and denitrification change ammonium into nitrogen gas which is totally harmless to the environment.

Nitrogen removal can be performed in several different types of WWTPs. One of the most popular is the *pre-denitrification* system, see Ingildsen (2002). In this design, the anoxic tanks are placed before the aerated basins, and thus, denitrification is performed before the nitrification process, see Figure 3.3.

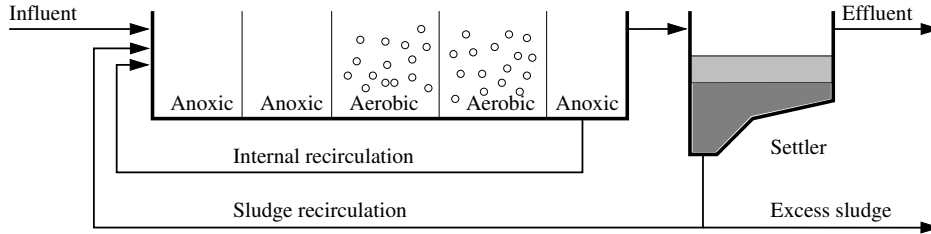


Figure 3.3: An activated sludge process configured for nitrogen removal (*pre-denitrification*).

To supply the denitrification process with nitrate, there is a feedback flow from the last tank as shown in Figure 3.3. In some cases, when the influent water has a low content of carbon, the bacteria in the anoxic tank need to be fed with carbon. For this purpose, methanol or ethanol is often used. The last tank in Figure 3.3 aims to reduce the DO concentration before the water is recirculated to the first anoxic tank.

For a deeper discussion about the ASP, see e.g. Olsson and Newell (1999) and Metcalf and Eddy Inc. (1991).

3.3 The IAWQ Activated Sludge Model No. 1

For the previously discussed ASP, there are several different models. The most common ASP model today, is the IAWQ¹ Activated Sludge Model No. 1 (ASM1), developed by the International Association on Water Pollution Research and Control (IAWPRC), see Henze *et al.* (1986). Here eight different processes are modelled:

- P_1 Aerobic growth of heterotrophs;
- P_2 Anoxic growth of heterotrophs;
- P_3 Aerobic growth of autotrophs;
- P_4 Decay of heterotrophs;
- P_5 Decay of autotrophs;
- P_6 Ammonification;
- P_7 Hydrolysis of entrapped organic materials;
- P_8 Hydrolysis of entrapped organic nitrogen.

Further, thirteen state variables are used in the model:

- S_I Soluble inert organic matter;
- S_S Soluble readily biodegradable substrate;
- X_I Particulate inert organic matter and products;
- X_S Slowly biodegradable substrate;
- $X_{B,A}$ Active autotrophic biomass;
- $X_{B,H}$ Active heterotrophic biomass;
- X_P Particulate products arising from biomass decay;
- S_O Dissolved oxygen;
- S_{NO} Soluble nitrate nitrogen;
- S_{NH} Soluble ammonium nitrogen;
- S_{ND} Soluble biodegradable organic nitrogen;
- X_{ND} Particulate biodegradable organic nitrogen;
- S_{ALK} Alkalinity.

¹International Association for Water Quality. This association was formerly called International Association on Water Pollution Research and Control (IAWPRC).

Obviously, the ASM1 contains a lot of states and is hence rather complex. For this reason, several simplified models based on the ASM1 have been developed. For the investigations carried out in this thesis, two reduced models will be used, as done by Ingildsen (2002).

A more detailed description of the ASM1 model can be found in Henze *et al.* (1986).

3.4 The COST Benchmark WWTP

The comparison between different control strategies for a WWTP is often difficult due to the variable influent conditions and the high complexity of a WWTP. Therefore, to enable objective comparisons between different control strategies, a simulation Benchmark WWTP has been developed by the COST 682 Working Group No.2, see Copp (2002) and the website <http://www.ensic.inpl-nancy.fr/COSTWWTP/>.

In the Benchmark WWTP a typical WWTP with pre-denitrification is implemented. It consists of five biological reactor tanks configured in-series. The first two tanks have a volume of 1000 m³ each, and are anoxic and assumed to be fully mixed. The remaining three tanks are aerated and have a volume of 1333 m³ each. All biological reactors are modelled according to the ASM1 model. Finally, there is a secondary settler modelled using the double-exponential settling velocity function of Takács *et al.* (1991).

To get an objective view of the performance of the applied control strategy, it is important to run the Benchmark simulation with different influent disturbances. Therefore, influent input files for three different weather conditions – dry weather, stormy weather and rainy weather – are available together with the Benchmark implementation.

3.5 Control of WWTPs

As previously seen, WWTPs are complex multivariable systems. Therefore, to obtain satisfactory control performance, it is often necessary to use more advanced control strategies. However, since wastewater treatment is "non-productive" compared to the industry, the extra investments needed to employ such advanced control strategies have been hard to justify economically. Nowadays, as the effluent demands get tighter, the interest for more advanced control strategies is again awakening, see Olsson and Newell (1999).

3.5.1 Control Handles for Nitrogen Removal

When removing nitrogen, there are several variables that can be used as actuators, or *control handles*, to control the outputs. In a pre-denitrification

system, there are five main control handles, as stated by Ingildsen (2002):

1. The airflow rate (in the aerated compartments);
2. The internal recirculation flow rate, Q_{int} ;
3. The sludge outtake flow rate (excess sludge);
4. The external carbon dosage, $S_{S_{dosage}}$;
5. The sludge recirculation flow rate.

The first control handle, the airflow rate, is employed to affect the DO concentration in the aerated compartments. Hereby, the performance of the autotrophic nitrification bacteria will be affected. Most common today is to control the airflow rate to maintain a specific DO level. Another way is to make use of online-measurements of the ammonium concentration in the last aerated compartment, and let these control the time-varying DO setpoint, see e.g. Lindberg (1997).

The internal recirculation flow rate, Q_i , affects the supply of nitrate for the denitrification process but also the DO concentration in the anoxic compartments since some DO may be transported from the last aerated compartment. The DO transportation between the processes, can however, be reduced by introducing an anoxic tank after the last aerated basin, see Figure 3.3.

External carbon dosage, $S_{S_{dosage}}$, can be applied when the influent water does not have enough readily biodegradable substrate to feed the denitrification bacteria.

3.5.2 Controlled Output Signals for Nitrogen Removal

The primary outputs from a WWTP are the effluent ammonium concentration, the organic matter, the nitrate concentration and the suspended solids, see Ingildsen (2002). For a more thorough discussion on cause-effect relationships in activated sludge plants, see Olsson and Jeppsson (1994).

3.6 Simplified ASM1 Models

In this thesis two simplified bioreactor models based on the ASM1 model are analyzed. Similar models are analyzed by Ingildsen (2002).

The first, and the most simple of these two bioreactor models, only considers processes that are relevant in the medium time scale, i.e. in the time scale of hours to days. Hence, slowly changing variables are assumed constant and quickly varying variables are neglected. The growth of autotrophic and heterotrophic microorganisms can be regarded as slow, and thus, the processes P_4 and P_5 are excluded (see Section 3.3 for nomenclature). The

ammonification and the hydrolysis (P_6 , P_7 and P_8) are also neglected since, when the ASP operates under normal conditions, all of these have an almost constant process rate, see Ingildsen (2002). Further, the ASP is here modelled as a two-tank system, with one anoxic tank and one aerated tank. Anoxic processes (e.g. denitrification) only take place in the anoxic tank, and similarly, the aerobic processes (e.g. nitrification) are only allowed in the aerated tank. Hence, the DO concentration is assumed to be zero in the anoxic tank. The volume of the two tanks were chosen to 2000 m³ and 3999 m³, respectively, and hence these volumes equals the total volumes of the anoxic and the aerobic compartments in the Benchmark implementation. The differential equations describing the model will hence be given by:

$$\begin{aligned}
\frac{dS_{\text{NH}}(1)}{dt} &= \frac{Q}{V_1} S_{\text{NH},\text{in}} - \frac{Q + Q_i}{V_1} S_{\text{NH}}(1) + \frac{Q_i}{V_1} S_{\text{NH}}(2) - i_{XB} P_2(1) \\
\frac{dS_{\text{NH}}(2)}{dt} &= \frac{Q + Q_i}{V_2} S_{\text{NH}}(1) - \frac{Q + Q_i}{V_2} S_{\text{NH}}(2) - i_{XB} P_1(2) - (i_{XB} + \frac{1}{Y_A}) P_3(2) \\
\frac{dS_{\text{NO}}(1)}{dt} &= -\frac{Q + Q_i}{V_1} S_{\text{NO}}(1) + \frac{Q_i}{V_1} S_{\text{NO}}(2) - \frac{1 - Y_H}{2.86 Y_H} P_2(1) \\
\frac{dS_{\text{NO}}(2)}{dt} &= \frac{Q + Q_i}{V_2} S_{\text{NO}}(1) - \frac{Q + Q_i}{V_2} S_{\text{NO}}(2) + \frac{1}{Y_A} P_3(2) \\
\frac{dS_S(1)}{dt} &= \frac{Q}{V_1} S_{S,\text{in}} - \frac{Q + Q_i}{V_1} S_S(1) + \frac{Q_i}{V_1} S_S(2) - \frac{1}{Y_H} P_2(1) \\
\frac{dS_S(2)}{dt} &= \frac{Q + Q_i}{V_2} S_S(1) - \frac{Q + Q_i}{V_2} S_S(2) - \frac{1}{Y_H} P_1(2)
\end{aligned} \tag{3.1}$$

where the nomenclature is explained in Table 3.1. The arguments 1 and 2 denote tank 1 and tank 2, respectively.

The used process equations are given by:

$$\begin{aligned}
P_1(1) &= \mu_H \frac{S_S(1)}{K_S + S_S(1)} \frac{S_O(1)}{K_{O,H} + S_O(1)} X_{B,H} \\
P_1(2) &= \mu_H \frac{S_S(2)}{K_S + S_S(2)} \frac{S_O(2)}{K_{O,H} + S_O(2)} X_{B,H} \\
P_2(1) &= \mu_H \frac{S_S(1)}{K_S + S_S(1)} \frac{K_{O,H}}{K_{O,H} + S_O(1)} \frac{S_{\text{NO}}(1)}{K_{\text{NO}} + S_{\text{NO}}(1)} \eta_g X_{B,H}
\end{aligned}$$

$$\begin{aligned}
P_2(2) &= \mu_H \frac{S_S(2)}{K_S + S_S(2)} \frac{K_{O,H}}{K_{O,H} + S_O(2)} \frac{S_{NO}(2)}{K_{NO} + S_{NO}(2)} \eta_g X_{B,H} \\
P_3(1) &= \mu_A \frac{S_{NH}(1)}{K_{NH} + S_{NH}(1)} \frac{S_O(1)}{K_{O,A} + S_O(1)} X_{B,A} \\
P_3(2) &= \mu_A \frac{S_{NH}(2)}{K_{NH} + S_{NH}(2)} \frac{S_O(2)}{K_{O,A} + S_O(2)} X_{B,A}
\end{aligned} \tag{3.2}$$

The model given in (3.1) can be extended if the assumptions of zero DO concentration in the anoxic compartment is relaxed. Hence, both processes, the nitrification and the denitrification, are allowed to take place in both reactor tanks. Mathematically, the DO concentration in the anoxic tank, $S_O(1)$, is added as an extra state. The model is then given by the following differential equations (Ingildsen, 2002):

$$\begin{aligned}
\frac{dS_{NH}(1)}{dt} &= \frac{Q}{V_1} S_{NH,in} - \frac{Q + Q_i}{V_1} S_{NH}(1) + \frac{Q_i}{V_1} S_{NH}(2) - i_{XB} P_2(1) \\
&\quad - i_{X,B} P_1(1) - (i_{XB} + \frac{1}{Y_A}) P_3(2) \\
\frac{dS_{NH}(2)}{dt} &= \frac{Q + Q_i}{V_2} S_{NH}(1) - \frac{Q + Q_i}{V_2} S_{NH}(2) - i_{XB} P_2(2) - i_{XB} P_1(2) \\
&\quad - (i_{XB} + \frac{1}{Y_A}) P_3(2) \\
\frac{dS_{NO}(1)}{dt} &= \frac{Q + Q_i}{V_1} S_{NO}(1) + \frac{Q_i}{V_1} S_{NO}(2) - \frac{1 - Y_H}{2.86 Y_H} P_2(1) + \frac{1}{Y_A} P_3(1) \\
\frac{dS_{NO}(2)}{dt} &= \frac{Q + Q_i}{V_2} S_{NO}(1) - \frac{Q + Q_i}{V_2} S_{NO}(2) - \frac{1 - Y_H}{2.86 Y_H} P_2(2) + \frac{1}{Y_A} P_3(2) \\
\frac{dS_S(1)}{dt} &= \frac{Q}{V_1} S_{S,in} - \frac{Q + Q_i}{V_1} S_S(1) + \frac{Q_i}{V_1} S_S(2) - \frac{1}{Y_H} P_2(1) - \frac{1}{Y_H} P_1(1) \\
\frac{dS_S(2)}{dt} &= \frac{Q + Q_i}{V_2} S_S(1) - \frac{Q + Q_i}{V_2} S_S(2) - \frac{1}{Y_H} P_2(2) - \frac{1}{Y_H} P_1(2) \\
\frac{dS_O(1)}{dt} &= -\frac{Q + Q_i}{V_1} S_O(1) + \frac{Q_i}{V_1} S_O(2) - \frac{1 - Y_H}{Y_H} P_1(1) - (\frac{4.57}{Y_A} + 1) P_3(1)
\end{aligned} \tag{3.3}$$

The used parameter values are listed in Appendix A.

Table 3.1: *Nomenclature for the bioreactor models.*

Symbol	Explanation
S_{NH}	Ammonium concentration
S_{NO}	Nitrate and nitrite concentration
S_{O}	Dissolved oxygen (DO) concentration
S_{S}	Readily biodegradable substrate
Q	Influent flow rate
Q_i	Internal recirculation flow rate
V_1	Volume of tank 1, the anoxic tank
V_2	Volume of tank 2, the aerobic tank
η_g	Correction factor for anoxic growth of heterotrophs
i_{XB}	Quotient between the mass of nitrogen and the mass of the chemical oxygen demand
K_{NH}	Ammonium half saturation constant for autotrophs
K_{NO}	Nitrate half saturation constant for heterotrophs
$K_{\text{O,A}}$	Oxygen half saturation constant for autotrophs
$K_{\text{O,H}}$	Oxygen half saturation constant for heterotrophs
K_{S}	Half saturation constant for heterotrophs
μ_{A}	Autotrophic max. specific growth rate
μ_{H}	Heterotrophic max. specific growth rate
$X_{\text{B,A}}$	Active autotrophic biomass
$X_{\text{B,H}}$	Active heterotrophic biomass
Y_{A}	Autotrophic yield
Y_{H}	Heterotrophic yield

Chapter 4

Analysis of the Bioreactor Models

In this chapter the previously described bioreactor models will be analysed using the theoretical tools from Chapter 2.

The first objective is to solve the pairing problem and to find a suitable control structure. If the couplings between the different control handles in the system are sufficiently low, then a controller selection involving several decoupled (diagonal) SISO controllers may be suitable. If this is not the case, a MIMO control structure will provide a better solution. The MIMO solution will, however, generally be much more complex.

The second objective is to make a performance comparison between the RGA and the HIIA for the given bioreactor models: Which method will give the most appropriate pairing selection?

4.1 Linearising the Models

To calculate measures such as the RGA and the HIIA, the used model must be linear. Therefore, the simplified ASM1 models given by (3.1) and (3.3) are linearised around some operating point.

Before this, suitable input and output signals in the control problem must be selected. Since the simplified ASM1-models given by (3.1) and (3.3) are valid in the medium time scale (hours and days), a suitable set of input signals is the DO concentration ($S_O(2)$), the internal recirculation flow rate (Q_i) and the external carbon dosage ($S_{S_{dosage}}$). The denitrification is mainly influenced by Q_i and $S_{S_{dosage}}$, while the nitrification is mainly influenced by $S_O(2)$, according to Ingildsen (2002). Hence, if the couplings between Q_i and $S_O(2)$, and between $S_{S_{dosage}}$ and $S_O(2)$ are low, then the denitrification and the nitrification process may be considered separately when choosing control structure and SISO controllers may hence be selected.

The effluent ammonium concentration, $S_{NH}(2)$, and the total effluent

concentration of inorganic nitrogen, $S_{N_{tot}}(2) = S_{NH}(2) + S_{NO}(2)$, were selected as output signals. The effluent ammonium concentration can be regarded as a measure of how well the nitrification is performed, and similarly, the nitrate concentration can be regarded as a measure of how well the denitrification works. In practice, $S_{N_{tot}}(2)$ is often measured instead of $S_{NO}(2)$ and hence this signal is selected.

To obtain quadratic 2×2 -systems, only two of the three input signals were considered at a time. Here, only $S_O(2)$ and Q_i were considered.

The linearisation was performed using the MATLAB function `linmod`. The stationary operating points were found from simulations. The obtained linear models can be represented in standard state space form as:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t)\end{aligned}$$

where $x(t)$ is the state vector given by

$$x(t) = [S_{NH}(1) \ S_{NH}(2) \ S_{NO}(1) \ S_{NO}(2) \ S_S(1) \ S_S(2)]^T \quad (4.1a)$$

for the model (3.1), and by

$$x(t) = [S_{NH}(1) \ S_{NH}(2) \ S_{NO}(1) \ S_{NO}(2) \ S_S(1) \ S_S(2) \ S_O(1)]^T \quad (4.1b)$$

in the case with the extended model given by (3.3). The input signal vector $u(t)$ is given by:

$$u(t) = \begin{bmatrix} Q_i \\ S_O(2) \end{bmatrix} \quad (4.1c)$$

where $S_O(2)$ is the same as the DO setpoint. The output signal vector is given by:

$$y(t) = \begin{bmatrix} S_{NH}(2) \\ S_{N_{tot}}(2) \end{bmatrix} \quad (4.1d)$$

and

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (4.1e)$$

for the model given in (3.1), and for the extended model given in (3.3):

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (4.1f)$$

Note that since there is no direct term, D , in the state space representation, the HIIA method is applicable.

Three different operating points were selected. These corresponds to the input signals:

- $u_1 = [10000 \text{ m}^3/\text{day} \quad 2 \text{ mg/l}]^T$,
- $u_2 = [36892 \text{ m}^3/\text{day} \quad 2 \text{ mg/l}]^T$,
- $u_3 = [50000 \text{ m}^3/\text{day} \quad 2 \text{ mg/l}]^T$.

In the sequel, these input signals will be referred to as operating points. The influent flow rate, Q , is $18446 \text{ m}^3/\text{day}$, and hence the internal recirculation flow rate, Q_i , is twice as large for the second operating point u_2 .

The steady-state operational maps for the two models, (3.1) and (3.3), and for the Benchmark WWTP are shown in Figure 4.1–4.3. The output signals, $S_{\text{NH}}(2)$ and $S_{\text{N}_{\text{tot}}}(2)$, and the effluent nitrate concentration, $S_{\text{NO}}(2)$ are plotted against the two input signals $S_{\text{O}}(2)$ and Q_i . In Figure 4.1 and 4.2 the nitrate concentration in the first reactor, the anoxic, is also plotted against $S_{\text{O}}(2)$ and Q_i . Note also that the concentration variables for the Benchmark model refers to the effluent concentrations even though the argument '2' is omitted.

As can be clearly seen in the operational maps for the Benchmark WWTP in Figure 4.3, the ASP with pre-denitrification behaves in a non-linear way. Both the simplified models, (3.1) and (3.3), are also non-linear even though they do not fully manage to model the dynamics of the ASM1-model used in the Benchmark WWTP. This divergence is clearly seen in the operational map for $S_{\text{N}_{\text{tot}}}(2)$ for model (3.1) in Figure 4.1. The extended model (3.3) performs better, even though the numerical values and the inclination of the contours differ somewhat from the results obtained from the Benchmark WWTP in Figure 4.3.

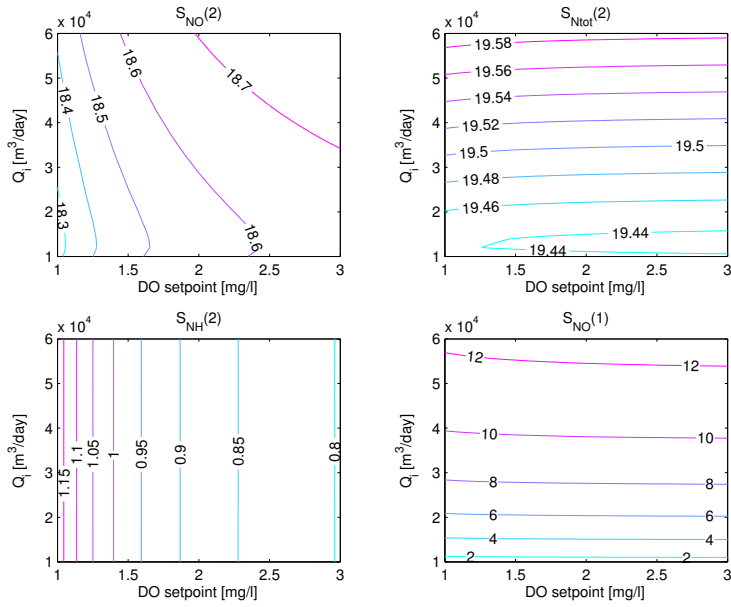


Figure 4.1: Operational maps for the simplified ASM1-model given by equation (3.1).

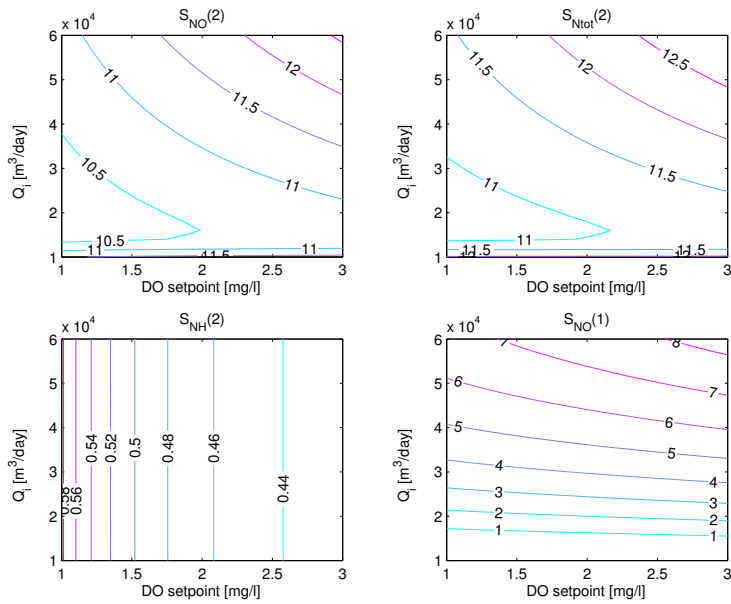


Figure 4.2: Operational maps for the extended simplified ASM1-model given by equation (3.3).

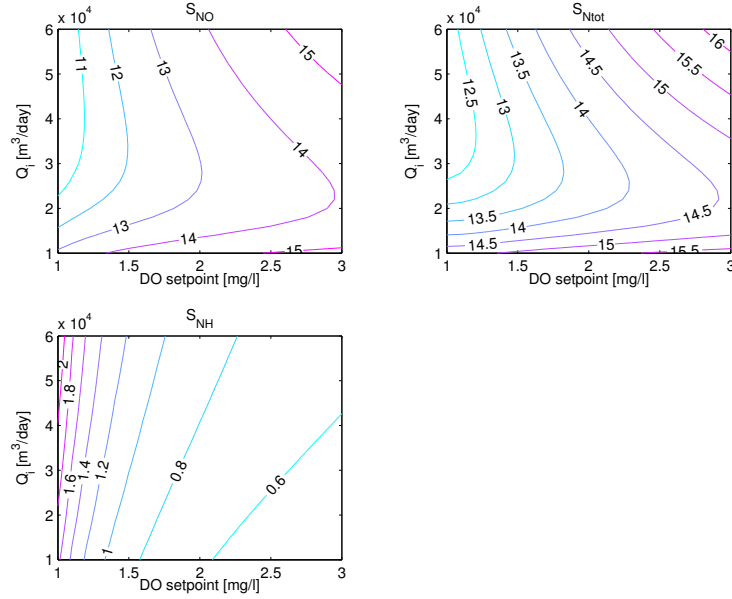


Figure 4.3: Operational maps obtained from the Benchmark WWTP.

4.2 The Influence of External Carbon Dosage

The coupling between the external carbon dosage concentration ($S_{S_{dosage}}$) and the DO concentration ($S_O(2)$) is not analysed using the RGA method or the HIIA method. However, to illustrate the influence of $S_{S_{dosage}}$, operational maps for model (3.3) with a $S_{S_{dosage}}$ of 90 g COD/m^3 , are given in Figure 4.4. Note that the total influent concentration of carbon, S_{S_i} , now is 159.5 g COD/m^3 .

Evidently, the nitrate concentration in the anoxic compartments, $S_{NO}(1)$, is now reduced to a value below 0.10 g COD/m^3 (compare Figure 4.4 and 4.2). The reason for this is an improved denitrification (since the denitrification bacteria is fed with more readily biodegradable substrate (carbon)). This also results in a lower nitrate concentration, $S_{NO}(2)$, in the aerobic compartments, see Figure 4.4. It is also evident that the influence of the DO concentration ($S_O(2)$) on $S_{NO}(2)$ and $S_{N_{tot}}(2)$ is much less with $S_{S_{dosage}} = 90 \text{ g COD/m}^3$ (compare Figure 4.4 and 4.2). In this case, the operational maps in Figure 4.4 clearly indicates that the couplings between Q_i and $S_O(2)$ are low since Q_i mainly affects $S_{N_{tot}}(2)$, and $S_O(2)$ mainly affects $S_{NH}(2)$. Instead, there is a significant coupling between $S_{NH}(2)$ and $S_O(2)$, and between $S_{N_{tot}}(2)$ and Q_i as seen in Figure 4.4.

In the forthcoming analysis, no dosage of external carbon is assumed to take place. The total influent concentration of carbon will hence be 69.5 g COD/m^3 .

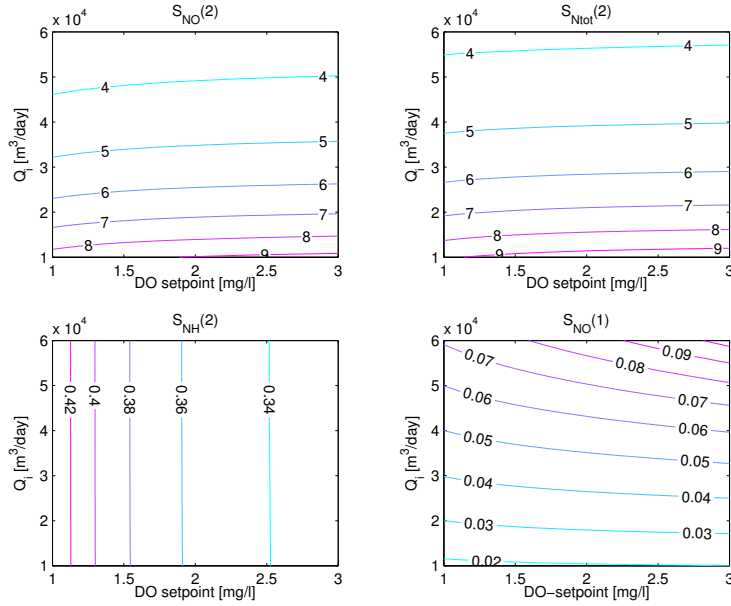


Figure 4.4: Operational maps for model (3.3) but with $S_{S_{dosage}} = 90$ mg COD/l.

4.3 RGA Analysis of Model (3.1)

In this section, the simplified ASM1-model given by equation (3.1) will be analysed using the RGA method.

4.3.1 Steady-state Analysis

Inspired by Ingildsen (2002), the analysis starts in the operating point u_2 . The steady-state transfer function matrix, $G(0)$, obtained for this operating point, is given by:

$$G_{u_2}(0) = \begin{bmatrix} 0.0000000 & -0.1301435 \\ 0.0000033 & -0.0024930 \end{bmatrix} \quad (4.2)$$

and the steady-state RGA, $\Lambda(0)$, for the same point is given by:

$$\Lambda_{u_2}(0) = \begin{bmatrix} -0.0000018 & 1.0000018 \\ 1.0000018 & -0.0000018 \end{bmatrix} \quad (4.3)$$

From (4.3), algebraic Property 3 and 5 of the RGA matrix, Λ , given in Section 2.1.2, can very easily be verified.

Clearly, the RGA suggests an anti-diagonal input-output pairing, i.e. $S_{NH(2)}-S_O(2)$ and $S_{N_{tot}(2)}-Q_i$. A diagonal pairing should definitely be avoided since the diagonal elements are negative.

If the rows in $G(0)$ given by (4.2) are switched in order to get the selected pairing along the diagonal, the RGA-number and the Niederlinsky index can be calculated.¹ The RGA-number is found to be very close to zero (7.2655×10^{-6}) which indicates that the selected control structure is satisfactory with a very low degree of cross-couplings. The Niederlinsky index equals 1 in this particular case, so this does not give any indications of stability problems. The condition number for $G(0)$ is however very large: 39155. But, as discussed in Section 2.1.6, the condition number is scaling dependent. Hence, the analysed system needs to be properly scaled before doing this kind of investigations. When considering stability questions, it is sufficient to consider a scaling that minimizes the condition number, see Section 2.1.6. In the case of a 2×2 matrix, the minimized condition number, γ_{min} , can be calculated using equation (2.24). Here, γ_{min} equals 1.0027 and hence, no stability problems can be expected.

Another common choice is to scale the variables in the system so that their maximum deviation from their average point lies in the interval $[-1, 1]$. Here, this scaling choice can be realized by choosing the scaling matrices as:

$$D_u = \begin{bmatrix} 36892 & 0 \\ 0 & 2 \end{bmatrix} \quad (4.4a)$$

$$D_y = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \quad (4.4b)$$

The last scaling matrix, D_y , simply states that a maximum deviation in the outputs y of 3 units is accepted. If this scaling procedure is applied to $G(0)$, the condition number becomes much lower (2.1229). Hence, the condition number does not indicate any problem whatsoever.

4.3.2 Dynamic Analysis

So far, only steady-state results have been considered, but as discussed in Section 2.1.4, the RGA can also be used for a dynamic study. Figure 4.5 shows the behaviour of the real part of λ_{11} for the three operating points over different frequency ranges. For low frequencies, which are the interesting ones here, the real part of λ_{11} is very close to 0 for operating point u_2 . The same results are obtained for the two other, but are omitted here. For higher frequencies the two upper operating points, u_2 and u_3 , have a real part of λ_{11} with a deep valley. The curve corresponding to the operating point u_1 does not have this property. However, since the studied model is only valid in the medium time scale (which means frequencies in the range

¹In the sequel, when calculating the RGA-number and the Niederlinsky index, it is assumed that the rows in G are switched (if necessary) to get the suggested input-output pairing along the diagonal.

of approximately 10^{-5} rad/s to 10^{-3} rad/s), there is no need to pay much attention to these discrepancies.

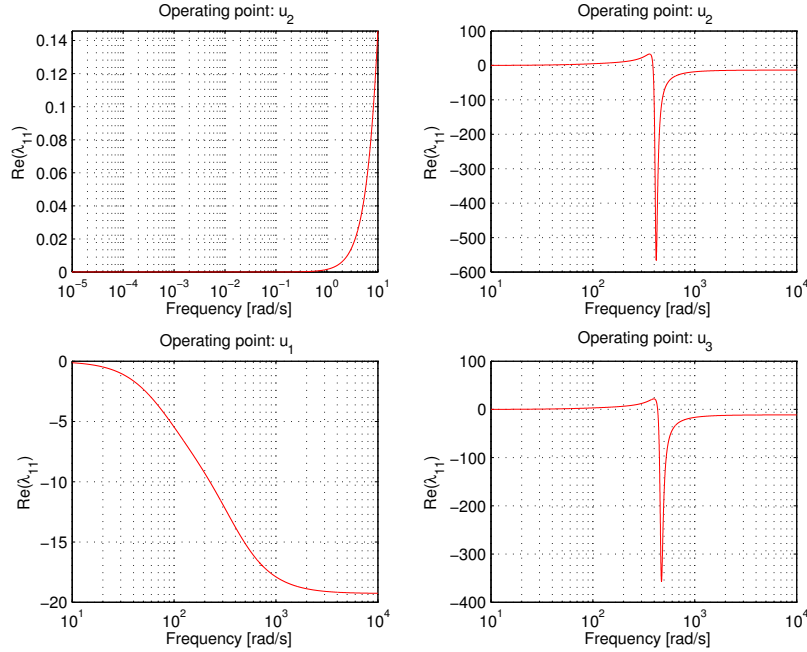


Figure 4.5: Plots of the the real part of λ_{11} for the three different operating points for model (3.1).

Figure 4.6 and Figure 4.7 shows the dynamic behaviour of the condition numbers for the system transfer function matrices. In the upper diagram of Figure 4.6 the condition number for the unscaled system at operating point u_2 , is shown for low frequencies (between 10^{-5} rad/s and 10^1 rad/s). This condition number is clearly very large, but if the system is scaled with the suggested scaling matrices in (4.4) then the condition number goes low as seen in the lower diagram in the same figure. However, for higher frequencies even the scaled system gets ill-conditioned. The condition number for the scaled system, passes value 10 already slightly above 7 rad/s. The minimized condition number, γ_{min} , is also plotted in the lower diagram, and has a low value until about 51 rad/s where it passes 10. In Figure 4.7 the same types of condition numbers are plotted but for higher frequencies (between 10^1 rad/s and 10^4 rad/s). Note especially that the peak is located at (almost) the same frequency as the dip in the plot of the real part of λ_{11} in Figure 4.5. This is a clear indication of the close relationship between the RGA and the condition number discussed in Section 2.1.6. Again, for frequencies corresponding to the interesting time scale here, the medium time scale, the condition number analysis does not indicate any problems.

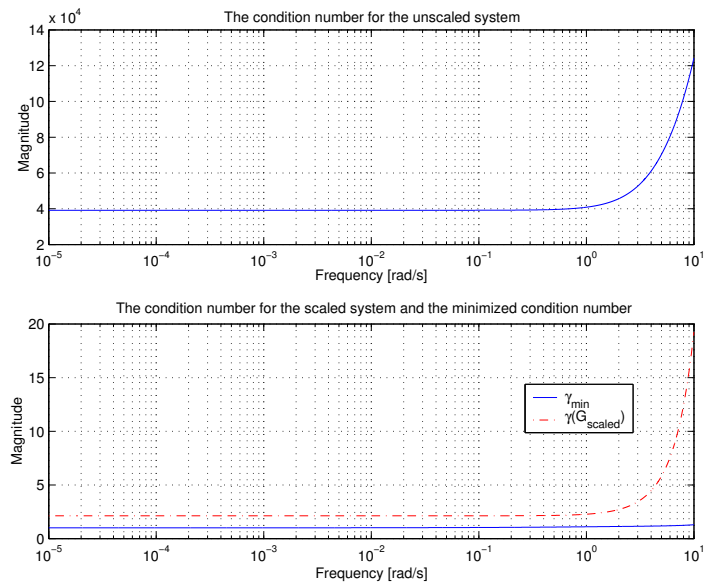


Figure 4.6: Plots of the condition numbers for the system obtained from model (3.1) at operating point u_2 . The scaled system is scaled according to (4.4).

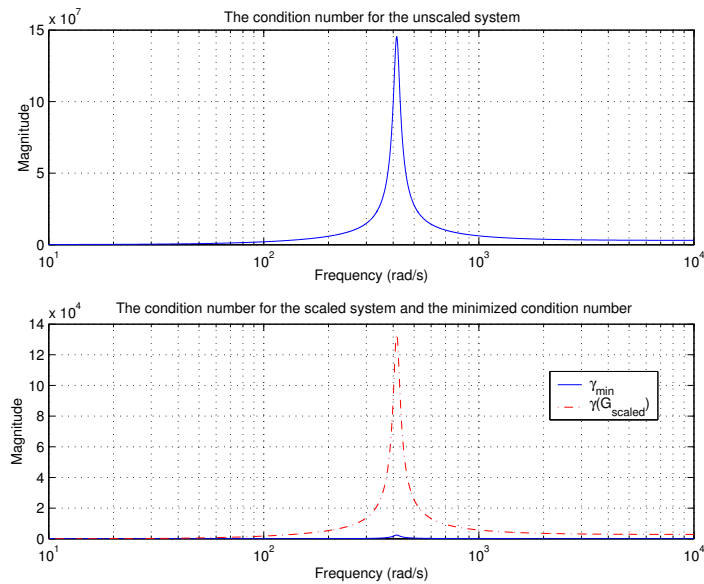


Figure 4.7: Plots of the condition numbers for the system obtained from model (3.1) at operating point u_2 . The scaled system is scaled according to (4.4).

4.3.3 Conclusions

The RGA suggests the same anti-diagonal input-output pairing for all of the three operating points in the considered time-scale. This conclusion seems reasonable when considering the operational maps in Figure 4.1: Clearly, $S_{NH}(2)$ should be paired with $S_O(2)$, and $S_{N_{tot}}$ is mostly affected by Q_i and should hence be controlled by Q_i .

4.4 RGA Analysis of Model (3.3)

The steady-state operational maps in Figure 4.2 for the extended model given by (3.3) differ from the operational maps obtained from the previously analysed model (3.1). From the operational map for $S_{N_{tot}}$ in Figure 4.2, it is obvious that $S_{N_{tot}}$ is affected by both $S_O(2)$ and Q_i . It can hence be expected that this property should be reflected in the RGA matrix. It is also obvious that the system will behave somewhat differently depending on which operating point that is chosen. Hence, all of them will be studied more carefully compared to the study of model (3.1).

4.4.1 Steady-state Analysis

The (unscaled) steady-state gain matrices obtained for the three operating points, u_3 , u_2 and u_1 , respectively, are given by:

$$G_{u_3}(0) = \begin{bmatrix} 0.0000000 & -0.0552867 \\ 0.0000295 & 0.6516180 \end{bmatrix} \quad (4.5a)$$

$$G_{u_2}(0) = \begin{bmatrix} 0.0000000 & -0.0552865 \\ 0.0000294 & 0.4780803 \end{bmatrix} \quad (4.5b)$$

$$G_{u_1}(0) = \begin{bmatrix} -0.0000000 & -0.0553009 \\ -0.0003721 & 0.0134283 \end{bmatrix} \quad (4.5c)$$

The gain matrices for u_2 and u_3 are rather similar, while the gain matrix for u_1 differs more. The difference in magnitude is largest for the lower right element, i.e. the gain between $S_{N_{tot}}$ and $S_O(2)$. This difference is expected, since according to the operational map for $S_{N_{tot}}$ in Figure 4.2, $S_{N_{tot}}$ is much less influenced by $S_O(2)$ in u_1 than it is in u_2 and u_3 .

However, to be able to draw any useful conclusions directly from the gain matrices, they should first be properly scaled. If the same scaling procedure as defined in equation (4.4) (with the upper left element in D_u changed to 10000 and 50000 for u_1 and u_3 , respectively) is applied, the following scaled gain matrices are obtained:

$$G_{u_3}^{scaled}(0) = \begin{bmatrix} 0.0000018 & -0.0368578 \\ 0.4915713 & 0.4344120 \end{bmatrix} \quad (4.6a)$$

$$G_{u_2}^{scaled}(0) = \begin{bmatrix} 0.0000013 & -0.0368577 \\ 0.3618929 & 0.3187202 \end{bmatrix} \quad (4.6b)$$

$$G_{u_1}^{scaled}(0) = \begin{bmatrix} -0.0000514 & -0.0368673 \\ -1.2404721 & 0.0089522 \end{bmatrix} \quad (4.6c)$$

Once again, it is obvious from $G_{u_1}^{scaled}$ that $S_{N_{tot}}$ is far mostly influenced by Q_i in u_1 . It is now also obvious that for the upper two operation points, u_3 and u_2 , $S_{N_{tot}}$ is almost equally influenced by Q_i and $S_O(2)$. According to the operational maps in Figure 4.2 this is reasonable. However, it should be noted that nothing else should be expected since the operational maps indeed are plots of the gains (multiplied with the input signals) calculated for a grid of different operating points. Hence, the analysis of the gains in (4.6) does not provide any extra information than the operational maps in Figure 4.2. Even though a steady-state gain analysis provides some information, it does not consider cross-coupling properties like the RGA does, see the derivation of the RGA in Section 2.1.1 (Remember that the RGA considers what happens with the gain when other loops are closed and opened.). Hence, a RGA analysis is suitable even in this case.

The steady-state RGA matrices for the three operation points are found to be:

$$\Lambda_{u_3}(0) = \begin{bmatrix} 0.0000428 & 0.9999572 \\ 0.9999572 & 0.0000428 \end{bmatrix} \quad (4.7a)$$

$$\Lambda_{u_2}(0) = \begin{bmatrix} 0.0000318 & 0.9999682 \\ 0.9999682 & 0.0000318 \end{bmatrix} \quad (4.7b)$$

$$\Lambda_{u_1}(0) = \begin{bmatrix} 0.0000101 & 0.9999899 \\ 0.9999899 & 0.0000101 \end{bmatrix} \quad (4.7c)$$

Obviously, the RGA suggests an anti-diagonal pairing, $S_{NH}(2)$ – $S_O(2)$ and $S_{N_{tot}}(2)$ – Q_i , for all of the three operation points. Hence, the RGA does not make any difference between u_1 and the upper two points u_2 and u_3 . The RGA-number is in the order of 10^{-4} for u_2 and u_3 , and 10^{-5} for u_1 .

The Niederlinski index is found to be 1 for all of the operation points. Hence this does not indicate any stability problems.

The condition numbers for the unscaled systems are high (157 , 1.42×10^5 and 2.62×10^5 for u_1 , u_2 and u_3 , respectively) but much lower for the scaled systems given in (4.6c)–(4.6a) (33.6 , 17.5 , and 23.8 for u_1 , u_2 and u_3 , respectively) but can still be regarded as large. Note that the condition number is large (33.6) in u_1 despite that the RGA here gives a reasonable pairing suggestion. The minimized condition number, γ_{min} , is, however, the lowest possible for all operating points, with a value of 1. Hence, this does not suggest any stability problems. Note, that since the number of positive elements in $G_{u_1}(0)$, $G_{u_2}(0)$ and $G_{u_3}(0)$, is odd, γ_{min} will always be 1, see Section 2.1.6. The same thing applies to diagonal and triangular systems.

4.4.2 Dynamic Analysis

Seen over a larger frequency interval, there are some differences in the RGA matrices as can be seen in Figure 4.8 where the real part of λ_{11} is plotted in various frequency ranges. For the interesting frequency range, i.e. between about 10^{-5} rad/s and 10^{-3} rad/s, the real part of λ_{11} is very close to zero. Hence, the RGA does not suggest a different pairing dynamically than statically.

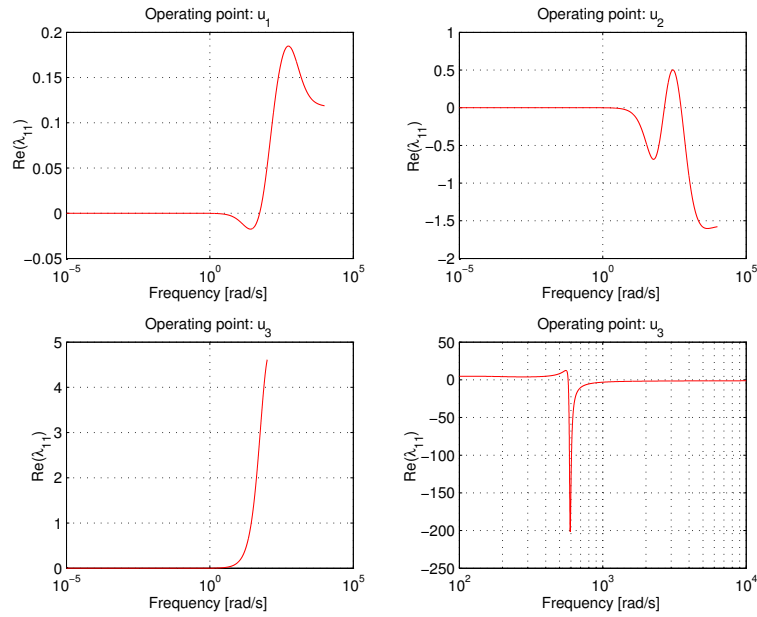


Figure 4.8: Plots of the the real part of λ_{11} for the systems obtained from model (3.3) in the three different operating points.

Figures 4.9 – 4.12 show the condition numbers for the three operating points for different frequency ranges. The condition numbers of the unscaled systems are always very high. For the scaled systems, the condition numbers becomes much smaller in magnitude but can still be considered as high with magnitudes a bit over 10 for most frequencies. The minimized condition number is however always low in the interesting frequency range and does hence not indicate any stability problems. Also note the similarities between the real part of λ_{11} and the condition numbers (compare for example Figure 4.8 and Figure 4.10). For a theoretical discussion of the link between the RGA and the condition number, see Section 2.1.6.

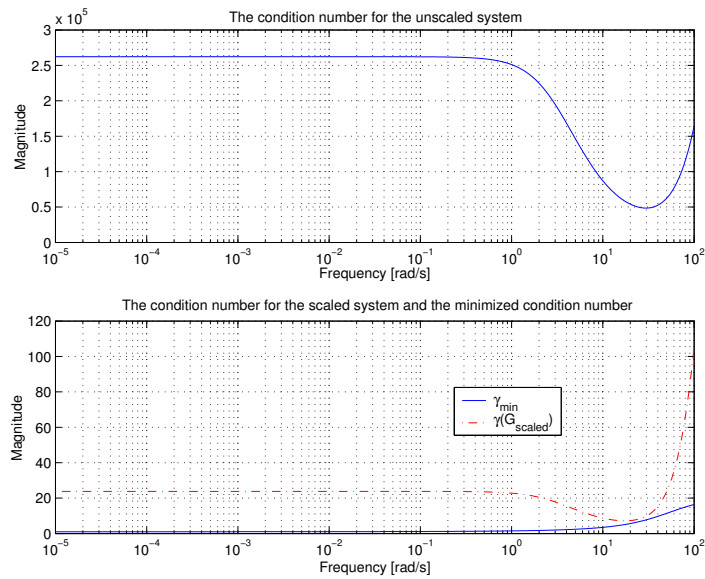


Figure 4.9: Plots of the condition numbers for the system obtained from model (3.3) at operating point u_3 . The scaled system is scaled according to (4.4) but with the upper left element in D_u replaced by 50000.

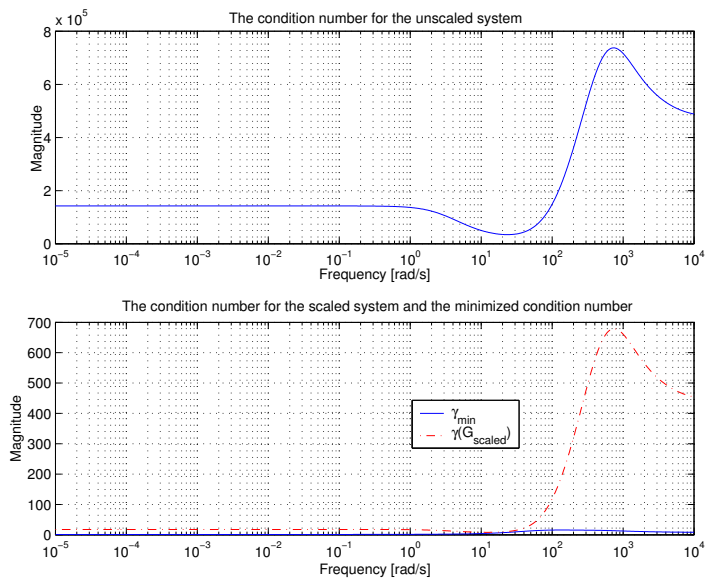


Figure 4.10: Plots of the condition numbers for the system obtained from model (3.3) at operating point u_2 . The scaled system is scaled according to (4.4).

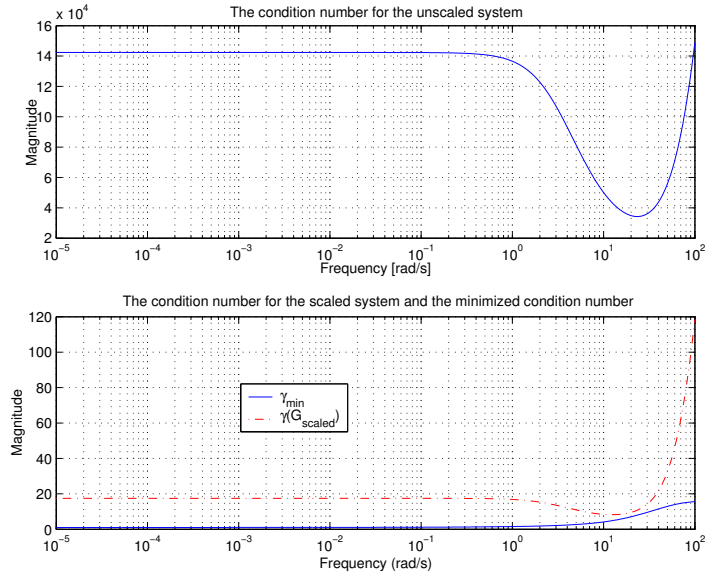


Figure 4.11: Plots of the condition numbers for the system obtained from model (3.3) at operating point u_2 . The scaled system is scaled according to (4.4).

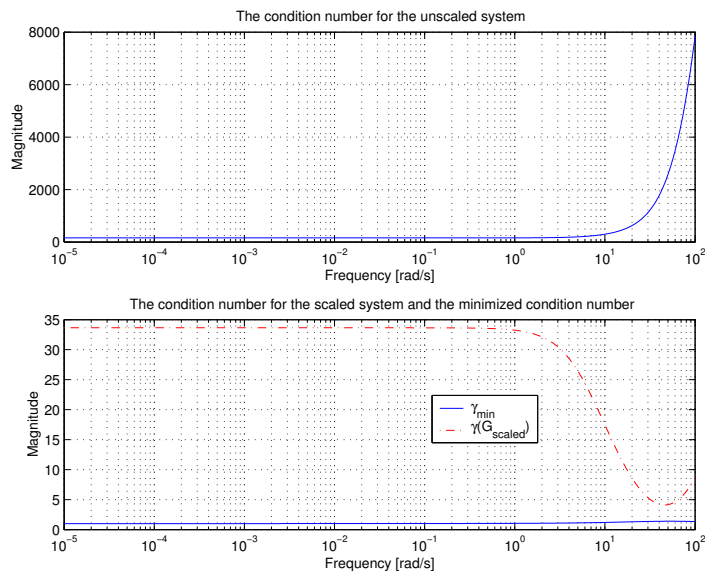


Figure 4.12: Plots of the condition numbers for the system obtained from model (3.3) at operating point u_1 . The scaled system is scaled according to (4.4) but with the upper left element in D_u replaced by 10000.

4.4.3 Conclusions

The RGA does not show any significant differences between the three operating points, and, as for the previous model given by (3.1), the RGA suggests the anti-diagonal pairing $S_{\text{NH}}(2)$ – $S_{\text{O}}(2)$ and $S_{N_{\text{tot}}}(2)$ – Q_i regardless of operating point or frequency (as long as the frequency lies in the frequency range that corresponds to the medium time scale). Clearly, the steady state operational maps in Figure 4.2 indicate that there should be a difference between at least the two upper operation points, u_2 and u_3 , and the lower one, u_1 . This is more thoroughly discussed in Section 4.9.2.

4.5 RGA Analysis of a Modified Version of Model (3.1)

To further test the ability of the RGA method to provide reasonable pairing suggestions for non-linear systems, the simplified ASM1 model (3.1) is modified: The volumes of both the bioreactors, V_1 and V_2 , are equally set to 1000 m^3 . Also, the output signal $S_{N_{\text{tot}}}(2)$ is replaced by $S_{\text{NO}}(2)$. For this modified version of (3.1) the operational maps for the two output signals, $S_{\text{NO}}(2)$ and $S_{\text{NH}}(2)$ in Figure 4.13, clearly indicate that different pairings should be used in the different operating points (at least, in the lower operating point, u_1 , a different pairing should be used than in the other two upper operating points).

4.5.1 Steady-state Analysis

The unscaled steady-state transfer function matrices in the different operating points are given by:

$$G_{u_3}(0) = \begin{bmatrix} 0.0000000 & -1.0572495 \\ 0.0000045 & 1.0011595 \end{bmatrix} \quad (4.8a)$$

$$G_{u_2}(0) = \begin{bmatrix} -0.0000000 & -1.0574619 \\ -0.0000008 & 0.9633135 \end{bmatrix} \quad (4.8b)$$

$$G_{u_1}(0) = \begin{bmatrix} -0.0000019 & -1.0603624 \\ -0.0002951 & 0.6788788 \end{bmatrix} \quad (4.8c)$$

If the systems are scaled in the same way as for the other two models (i.e. using the scaling matrices in (4.4) and replacing the upper left element in D_u with 10000 and 50000 for u_1 and u_3 , respectively) the following transfer function matrices are obtained:

$$G_{u_3}^{\text{scaled}}(0) = \begin{bmatrix} 0.0004401 & -0.7048330 \\ 0.0748164 & 0.6674396 \end{bmatrix} \quad (4.9a)$$

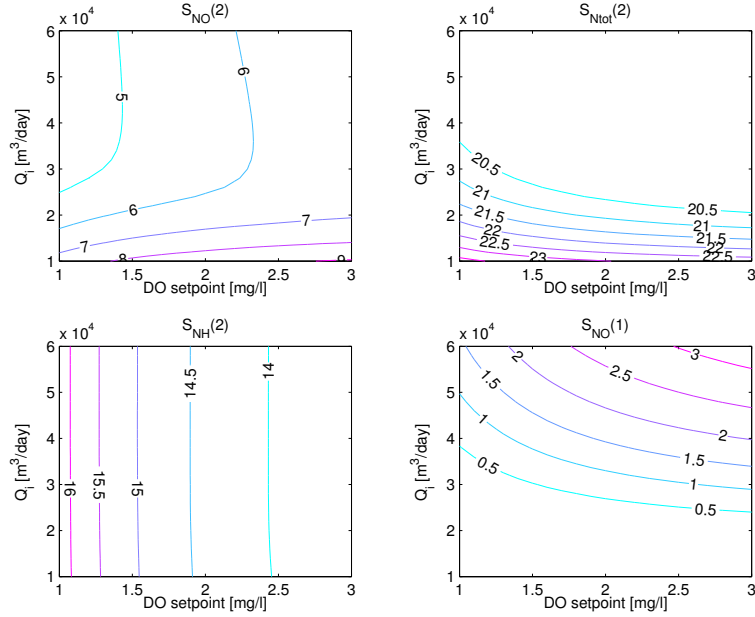


Figure 4.13: Operational maps for the modified version of (3.1) with $V_1 = V_2 = 1000 \text{ m}^3$.

$$G_{u_2}^{scaled}(0) = \begin{bmatrix} -0.0000557 & -0.7049746 \\ -0.0099721 & 0.6422090 \end{bmatrix} \quad (4.9b)$$

$$G_{u_1}^{scaled}(0) = \begin{bmatrix} -0.0062569 & -0.7069083 \\ -0.9835153 & 0.4525859 \end{bmatrix} \quad (4.9c)$$

As seen from the scaled transfer function matrices in (4.9a)–(4.9c), the main difference between the two upper operation points, u_3 and u_2 , and the lower operation point, u_1 , is the lower left element, i.e. the gain between Q_i and $S_{NO}(2)$. The gain between $S_O(2)$ and $S_{NO}(2)$ also changes when going upwards from u_1 to u_3 .

The steady-state RGA matrices, $\Lambda(0)$, are given by:

$$\Lambda_{u_3}(0) = \begin{bmatrix} 0.0055390 & 0.9944610 \\ 0.9944610 & 0.0055390 \end{bmatrix} \quad (4.10a)$$

$$\Lambda_{u_2}(0) = \begin{bmatrix} 0.0050592 & 0.9949408 \\ 0.9949408 & 0.0050591 \end{bmatrix} \quad (4.10b)$$

$$\Lambda_{u_1}(0) = \begin{bmatrix} 0.0040565 & 0.9959435 \\ 0.9959435 & 0.0040565 \end{bmatrix} \quad (4.10c)$$

The RGA-number are found to be 0.0222, 0.0202 and 0.0162 for u_3 , u_2 and u_1 , respectively. Even though the RGA-numbers are a bit larger for this model than for the previously analysed models, still the RGA suggests the

anti-diagonal pairing $S_{\text{NH}}(2)\text{--}S_{\text{O}}(2)$ and $S_{\text{NO}}(2)\text{--}Q_i$. This clearly contradicts the results from the operational maps in Figure 4.13.

The Niederlinski index does not give any warnings: Here it has a value of slightly above 1 for all of the three operation points.

The condition numbers for the unscaled steady-state transfer function matrix are all very high: 4.4×10^5 , 2.4×10^6 and 5.0×10^3 for u_3 , u_2 and u_1 , respectively. The condition numbers for the scaled systems are very low for u_1 (1.9), much higher for u_2 (129) and high for u_3 (18). Evidently, in this case, the condition number is large for the operating points where the RGA does not give reasonable pairing suggestions, and low when the RGA works. The minimized condition number is, however, as low as possible: the value is 1 for u_1 , u_2 and u_3 . Hence, regarding stability there should not be any problems, but from a performance point of view, there might be some problems, especially in the middle operating point, u_2 . Note also, that here, the number of positive elements in $G_{u_1}(0)$, $G_{u_2}(0)$ and $G_{u_3}(0)$ are odd, and that is the reason why $\gamma_{\min} = 1$.

4.5.2 Dynamic Analysis

Figure 4.14 shows the dynamic behaviour of the real part of λ_{11} . Nothing of interest happens for the relevant low frequency part and it can be concluded that the RGA suggests the same pairings even dynamically when considering frequencies corresponding to the medium time-scale.

The condition numbers for the three operating points for different frequency ranges are shown in Figure 4.19–4.16. As for the previously analysed models, the condition number for the unscaled systems is large. For the scaled systems the condition number becomes much smaller: For the system around u_1 , the condition number is now low (around 2 in the interesting frequency range between 10^{-5} rad/s and 10^{-3} rad/s), but it is still large around u_2 (around 130 in the interesting frequency range between 10^{-5} rad/s and 10^{-3} rad/s) and around u_3 (around 18 in the interesting frequency range between 10^{-5} rad/s and 10^{-3} rad/s). The minimized condition number is, however, very low in all of the three operating points. Hence, no stability problems should be expected.

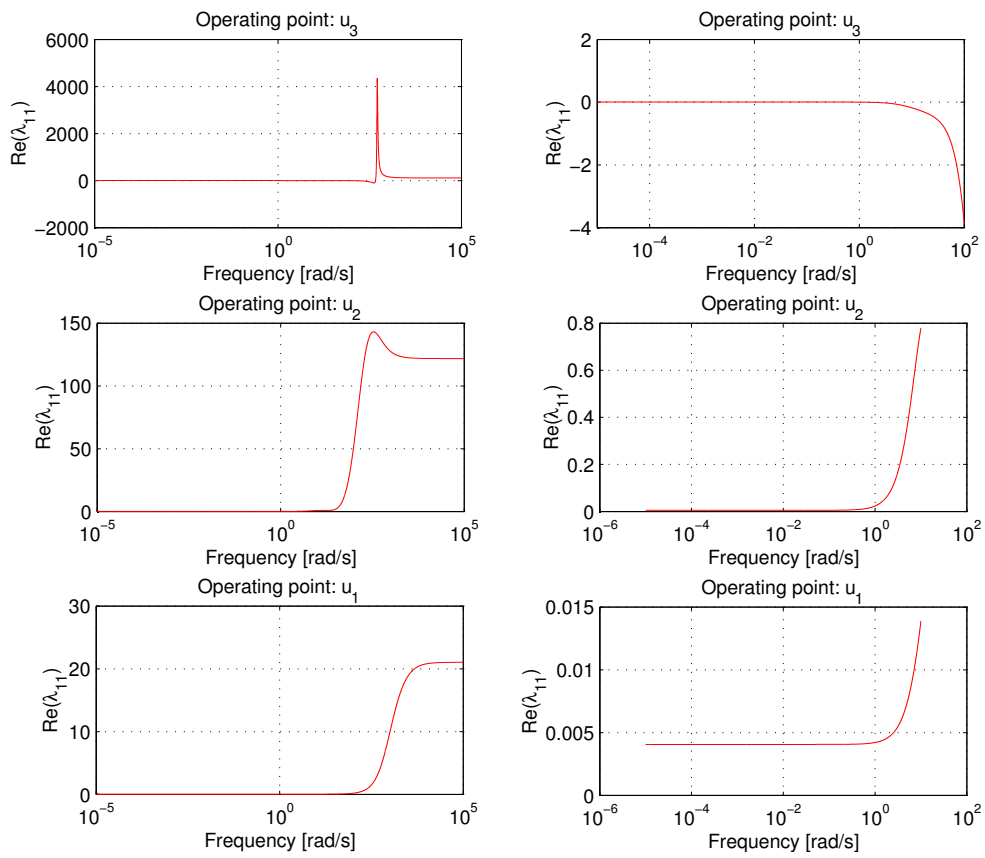


Figure 4.14: Plots of the the real part of λ_{11} for the three different operating points for the modified version of model (3.1).

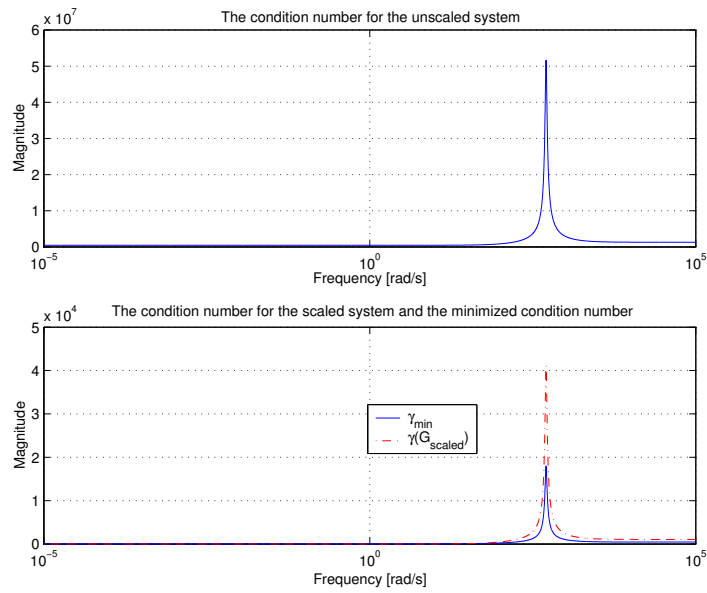


Figure 4.15: Plots of the condition numbers for the system obtained from the modified version of model (3.1) at operating point u_3 . The scaled system is scaled according to (4.4) but with the upper left element in D_u replaced by 50000.

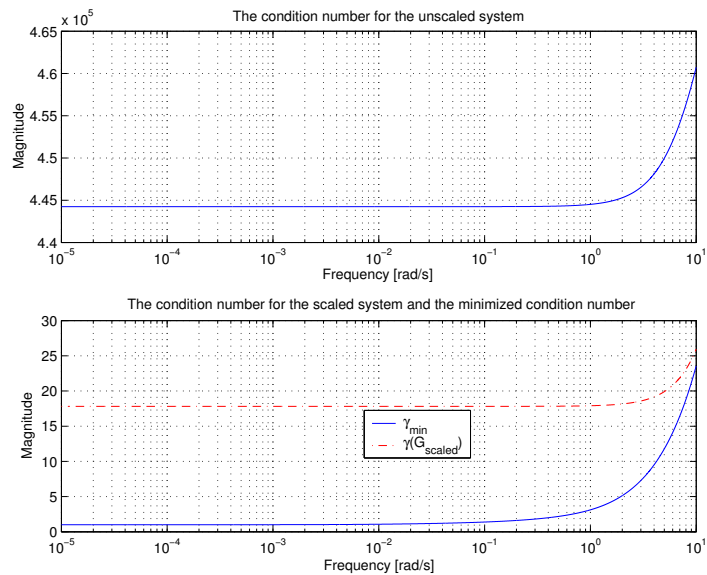


Figure 4.16: Plots of the condition numbers for the system obtained from the modified version of model (3.1) at operating point u_3 . The scaled system is scaled according to (4.4) but with the upper left element in D_u replaced by 50000.

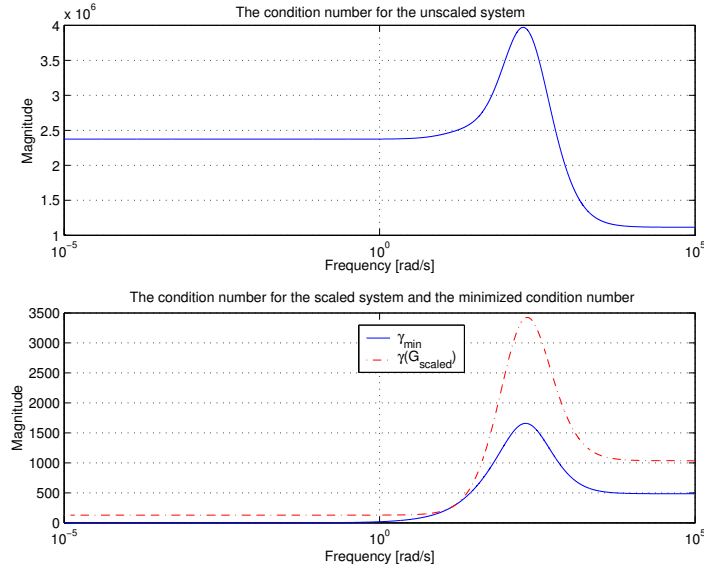


Figure 4.17: Plots of the condition numbers for the system obtained from the modified version of model (3.1) at operating point u_2 . The scaled system is scaled according to (4.4).

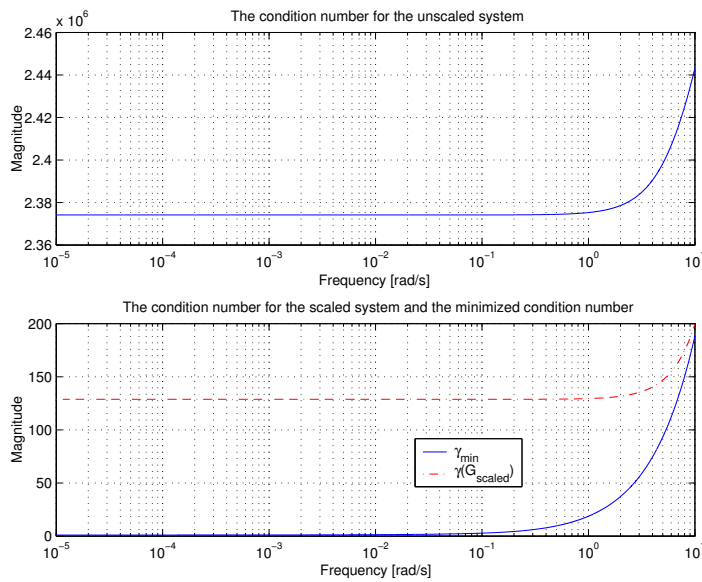


Figure 4.18: Plots of the condition numbers for the system obtained from the modified version of model (3.1) at operating point u_2 . The scaled system is scaled according to (4.4).

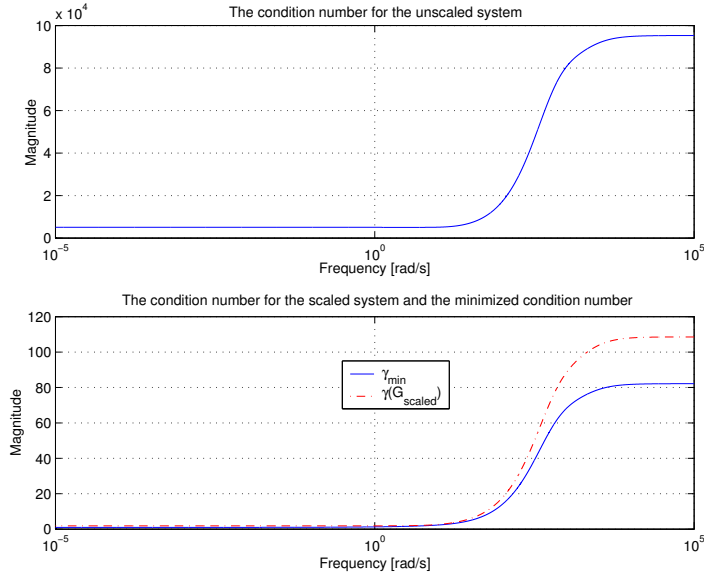


Figure 4.19: Plots of the condition numbers for the system obtained from the modified version of model (3.1) at operating point u_1 . The scaled system is scaled according to (4.4) but with the upper left element in D_u replaced by 10000.

4.5.3 Conclusions

The steady-state analysis as well as the dynamic analysis suggests the same input-output pairings for all of the three operating points, even though the operational maps given in Figure 4.13 clearly indicate that different pairings should be used around at least the lowest operating point u_1 compared to what should be used around the highest one, u_3 . This is further discussed in Section 4.9.2.

4.6 HIIA Analysis of Model (3.1)

The normalized Hankel Interaction Index Array (HIIA) given by (2.42), was calculated for the system given by (3.1). For the considered operating points, the following HIIA, Σ_H , were obtained:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0001162 & 0.9813764 \\ 0.0000274 & 0.0184800 \end{bmatrix} \quad (4.11a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0001543 & 0.9813048 \\ 0.0000271 & 0.0185138 \end{bmatrix} \quad (4.11b)$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0003492 & 0.9795406 \\ 0.0001013 & 0.0200089 \end{bmatrix} \quad (4.11c)$$

Since the right-hand side elements dominate in (4.11a)–(4.11c), the HIIA method suggests that the DO setpoint, $S_O(2)$, should be used to control both $S_{NH}(2)$ and $S_{N_{tot}}$ in all of the considered operation points. According to the operational maps in Figure 4.1 this is not reasonable.

The HIIA is, however, a scaling dependent tool. If the system is scaled in the same way as before, i.e. according to (4.4) for u_2 and with the same scaling matrices except that the upper left element in D_u is replaced by 10000 and 50000 for u_1 and u_3 , respectively, then the HIIA matrices become:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.6330149 & 0.2138250 \\ 0.1491336 & 0.0040265 \end{bmatrix} \quad (4.12a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.6547984 & 0.2257735 \\ 0.1151685 & 0.0042595 \end{bmatrix} \quad (4.12b)$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.5368849 & 0.3012409 \\ 0.1557207 & 0.0061534 \end{bmatrix} \quad (4.12c)$$

For the scaled systems, the dominant elements in the HIIA matrices are now located on the left-hand side, suggesting that Q_i should control both $S_{NH}(2)$ and $S_{N_{tot}}(2)$. Neither this is reasonable.

However, the HIIA is a dynamic measure that considers all possible frequencies. The considered models are, however, only valid in a limited frequency band, approximately 10^{-5} rad/s up to 10^{-3} rad/s. Therefore it is reasonable to perform a band-pass filtering before calculating the HIIA. Here, it is enough to do a low-pass filtering since only the high-frequency response deviates much from the response in the considered frequency range. This was done using a simple first-order low-pass filter, $F(s)$, given by:

$$F(s) = \frac{0.001}{s + 0.001} \quad (4.13)$$

where s is the Laplace-variable. This filter has a 3 dB cut-off frequency of approximately 10^{-3} rad/s, see the Bode diagram in Figure 4.20. The filtering can be expressed as:

$$G^{filtered} = GF \quad (4.14)$$

If the scaled systems are filtered using the low-pass filter F given in (4.13) before the HIIA is calculated, then the following HIIA matrices, Σ_H , are obtained for the three operating points:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000889 & 0.6041432 \\ 0.3843713 & 0.0113966 \end{bmatrix} \quad (4.15a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0001108 & 0.6708085 \\ 0.3162310 & 0.0128498 \end{bmatrix} \quad (4.15b)$$

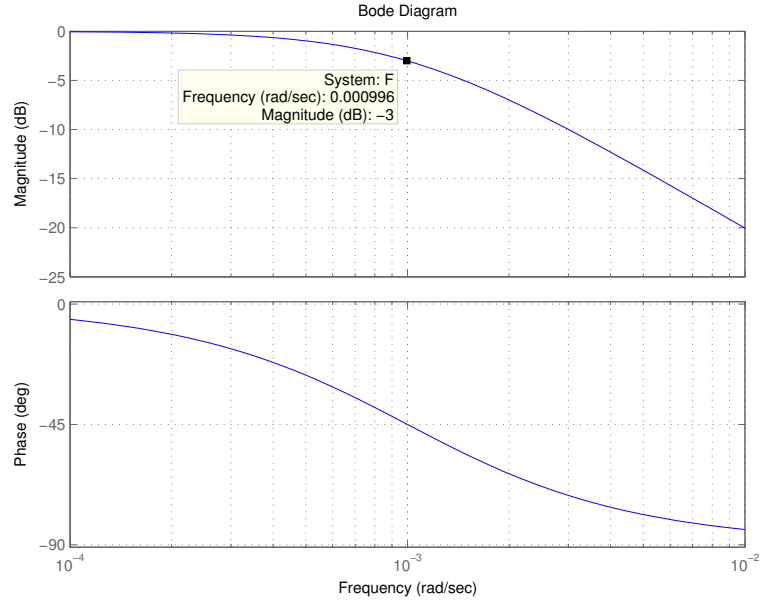


Figure 4.20: Bode diagram for the low-pass filter F .

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0001286 & 0.6872784 \\ 0.2934832 & 0.0191098 \end{bmatrix} \quad (4.15c)$$

Obviously, the suggested input-output pairings, $S_{NH}(2)-S_O(2)$ and $S_{N_{tot}}-Q_i$, do not contradict the steady-state results in the operational maps in Figure 4.1. According to the recently gained knowledge about the behaviour of the considered systems for higher frequencies (see for example the plots of the real part of λ_{11} in Figure 4.5 and the plots of the condition numbers in Figure 4.6 and 4.7), this is not surprising since now, the higher frequencies are excluded from the analysis.

Note that the filtered but unscaled systems do not give reasonable HIIA matrices:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000000 & 0.9814607 \\ 0.0000250 & 0.0185143 \end{bmatrix} \quad (4.16a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000000 & 0.9811798 \\ 0.0000251 & 0.0187951 \end{bmatrix} \quad (4.16b)$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0000000 & 0.9728664 \\ 0.0000831 & 0.0270506 \end{bmatrix} \quad (4.16c)$$

Hence, the scaling is needed.

4.7 HIIA Analysis of Model (3.3)

The results of the HIIA analysis of model (3.3) are similar to the results obtained in the previous section: To get HIIA matrices that are reasonable according to the operational maps given in Figure 4.2, the systems need to be scaled and filtered in the same way as before. The HIIA matrices for the scaled and filtered systems are:²

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000039 & 0.0382667 \\ 0.5105658 & 0.4511637 \end{bmatrix} \quad (4.17a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000058 & 0.0513523 \\ 0.5044358 & 0.4442061 \end{bmatrix} \quad (4.17b)$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0000428 & 0.0286552 \\ 0.9643419 & 0.0069602 \end{bmatrix} \quad (4.17c)$$

In all of the three operation points, the HIIA indicates that $S_{\text{NH}(2)}$ is mostly affected by $S_{\text{O}(2)}$. In u_1 , the HIIA strongly suggests that Q_i should control $S_{\text{N}_{\text{tot}}(2)}$, while in u_2 and u_3 , the magnitude of the elements in the lower row of Σ_H are almost equal. Therefore, best control authority should be achieved if both input signals, Q_i and $S_{\text{O}(2)}$ are employed to control $S_{\text{N}_{\text{tot}}(2)}$. Clearly, all of these suggestions agree with the results from the operational maps in Figure 4.2.

4.8 HIIA Analysis of the Modified version of Model (3.1)

Also, for the modified version of model (3.1), it is found that the HIIA of the filtered and scaled (with the same scaling as before) systems gives the most reasonable input-output pairing suggestions. In this way, the following HIIA matrices are obtained:³

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0002728 & 0.4869227 \\ 0.0517161 & 0.4610883 \end{bmatrix} \quad (4.18a)$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000785 & 0.5194289 \\ 0.0073134 & 0.4731792 \end{bmatrix} \quad (4.18b)$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0029353 & 0.3289062 \\ 0.4575898 & 0.2105686 \end{bmatrix} \quad (4.18c)$$

²Numerical values for the HIIA matrices obtained with the other combinations of scaling and filtering (i.e. unscaled, scaled, and unscaled and filtered) are given in Appendix B.

³Numerical values for the HIIA matrices obtained with the other combinations of scaling and filtering (i.e. unscaled, scaled, and unscaled and filtered) are given in Appendix B.

Hence, the HIA indicates that $S_O(2)$ affects both $S_{NH}(2)$ and $S_{NO}(2)$ ⁴ in u_2 and u_3 . In u_1 the anti-diagonal pairing $S_{NH}(2)$ – $S_O(2)$ and $S_{NO}(2)$ – Q_i is suggested. However, since $[\Sigma_H^{u_1}]_{22}$ is rather large, improved performance can be expected if a MIMO control structure is chosen where $S_{NO}(2)$ is also partly controlled by $S_O(2)$.

None of these conclusions contradict the conclusions that can be drawn from the operational maps in Figure 4.13.

4.9 General Conclusions

4.9.1 The Condition Number as a Warning Signal When the RGA May Not Work

As discussed in Section 2.1.6, the condition number may indicate when the RGA method does not give proper input-output pairing suggestions. Therefore, it is advisory to also calculate the condition number when performing a RGA analysis. For the modified version of model (3.1) this works fine if the transfer function matrices are scaled in a physically reasonable way (i.e. such that the maximum deviation of the considered signals from their average points lies in the interval $[-1, 1]$, see Section 4.5.1). However, as seen in Section 4.4.1, this is not the case for model (3.3) where the condition number is higher for the operating point u_1 than for the other operating points, u_2 and u_3 , even though the RGA gives a reasonable suggestion in u_1 only. It should also be noted that the condition number may be high even if the transfer function matrix is diagonal. One such example is the case of a diagonal transfer function matrix with elements 1 and 100: the condition number is 100. Hence, the condition number should only be considered as a warning tool that *may* indicate when the RGA does not work. To decide if the plant is triangular, it is more reliable to directly investigate the plant transfer function matrix.

4.9.2 Why the RGA Fails

In the RGA analysis of the bioreactor models in Section 4.4–4.5, it was concluded that the RGA method did not provide reasonable input-output pairings in all of the considered operating points. The reason for this can be found if the steady-state gain matrices for the considered systems are studied. According to Property 4 of the RGA matrix (see Section 2.1.2), triangular systems will always give the same RGA, namely the identity matrix (under the assumption that the rows in the transfer function matrix are permuted to get nonzero elements along the diagonal before calculating the RGA).

⁴Remember that $S_{N_{tot}}$ is replaced by $S_{NO}(2)$ in the modified version of model (3.1).

First, consider the modified version of (3.1). In this case the transfer function matrices are almost right under triangular, see (4.8a)–(4.8c) or the scaled transfer functions in (4.9a)–(4.9c). Therefore, the structure of the RGA will be similar for all of them: Almost the anti-identity matrix. The RGA matrices are given in equations (4.10a) – (4.10c), and obviously they are all very close to the anti-identity matrix.

To further investigate how the RGA behaves in this particular case, consider a transfer function matrix given by:

$$G = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (4.19)$$

The RGA of G is then given by

$$\Lambda = \frac{1}{ad - bc} \begin{bmatrix} ad & -bc \\ -bc & ad \end{bmatrix} \quad (4.20)$$

In the operation point u_1 the RGA gives a reasonable input-output pairing suggestion, $S_{\text{NH}}(2)$ – $S_{\text{O}}(2)$ and $S_{\text{NO}}(2)$ – Q_i . Consider the scaled steady-state transfer function matrix, $G_{u_1}^{\text{scaled}}(0)$ given in (4.9c). The relative differences between the elements in this transfer function matrix are given by:

$$\frac{G_{u_1}^{\text{scaled}}(0)}{[G_{u_1}^{\text{scaled}}(0)]_{11}} = \begin{bmatrix} 1.00 & 113.98 \\ 157.19 & -72.33 \end{bmatrix} \quad (4.21)$$

Evidently, the matrix is almost lower triangular.

If going upwards from u_1 to the uppermost operating point, u_3 , the magnitude of c will decrease (since the affection of Q_i on $S_{\text{NO}}(2)$ will decrease, see Figure 4.13). This change in c will affect the anti-diagonal RGA elements, λ_{12} and λ_{21} according to the expression given in (4.20). In this case, $|ad| \ll |bc|$ for all of the three operation points, see Table 4.1. λ_{12} and λ_{21} will hence be close to 1. Therefore, a change in c (b is almost constant) will only make a very small difference in the RGA matrix. This is exactly what is seen in the RGA matrices in (4.10a)–(4.10c).

Similarly, the change in d , when going from u_1 to u_3 , will not be reflected in the RGA matrix to any larger extent, since d is multiplied by a that is very small. Hence, it can be concluded that the RGA does not work properly in this particular case and should hence be avoided.

The transfer function matrices obtained from the other considered models are also almost triangular, and similar arguments can therefore be applied to show that the RGA method does not give reliable results in these particular cases. However, it can be argued that the RGA works for model (3.3) since it suggests the *best possible diagonal* (decoupled) control structure. For the modified version of model (3.1) the RGA method even fails to do this. Hence, the reason for the failure of the RGA should be sought for in the structure of the transfer function matrices as discussed above.

Table 4.1: *The size of the elements in the scaled steady-state transfer function matrices for the different models in the three different operation points (o.p.). a , b , c and d are defined in (4.19). The scaling is performed with the scaling matrices given by (4.4) for u_2 and the same matrices except that the upper left element in D_u is replaced by 10000 and 50000 for u_1 and u_3 , respectively. (3.1)m is the modified model analysed in Section 4.5.*

Model	O.p.	a	b	c	d	ad	bc
(3.1)	u_3	4.6107×10^{-6}	-0.0868	0.0552	-0.0016	-7.5463×10^{-9}	-0.0048
(3.1)	u_2	3.8772×10^{-6}	-0.0868	0.0409	-0.0017	-6.4438×10^{-9}	-0.0035
(3.1)	u_1	-3.3641×10^{-6}	-0.0868	-0.0370	-0.0024	8.1147×10^{-9}	0.0032
(3.3)	u_3	1.7836×10^{-6}	-0.0369	0.4916	0.4344	7.7481×10^{-7}	-0.0181
(3.3)	u_2	1.3301×10^{-6}	-0.0369	0.3619	0.3187	4.2393×10^{-7}	-0.0133
(3.3)	u_1	-5.1397×10^{-5}	-0.0369	-1.2405	0.0090	-4.6012×10^{-7}	0.0457
(3.1)m	u_3	4.4006×10^{-4}	-0.7048	0.0748	0.6674	2.9372×10^{-4}	-0.0527
(3.1)m	u_2	-5.5663×10^{-5}	-0.7050	-0.0100	0.6422	-3.5747×10^{-5}	0.0070
(3.1)m	u_1	-6.2568×10^{-3}	-0.7069	-0.9835	0.4526	2.8317×10^{-3}	0.6953

4.9.3 The RGA versus the HIIA

The RGA method provides a simple way to decide how a set of input signals should be utilized to control a given set of output signals. Often this method performs well, but in the analysis of the considered bioreactor models, it was clearly seen that the RGA method does not work properly for all of them. The reason for this was found to be the almost triangular structure of the transfer function matrices. This explanation did also explain why the RGA even failed to give the best decoupled (diagonal) control structure in some cases. From this it can be concluded that the RGA should be used with care. It is advisory to include an examination of the structure of the considered transfer function matrices in the RGA analysis.

Furthermore, the newly suggested HIIA method was employed to give input-output pairing suggestions for the considered bioreactor systems. It was shown that for the HIIA method to give reasonable pairing suggestions, the considered systems had to be both scaled in a physically relevant way and low-pass filtered. The filtering was performed to select the frequency band of interest – remember that the HIIA is a measure that takes all frequencies into account. When treating the systems according to this procedure, the HIIA method gave reasonable pairing suggestions for all of the considered systems.

Compared to the RGA, the HIIA possesses several advantages. Evidently, the HIIA is able to deal with special transfer function matrix struc-

tures such as the analysed nearly triangular structures. Also, the HIIA does not require decoupled (diagonal) control structures as the RGA does. Hence, the HIIA can be used to suggest MIMO control structures as seen in Section 4.6–4.8.

It was also observed that the HIIA method is scaling dependent. This means that some effort must be spent on finding proper scaling matrices. However, this is not necessarily a drawback, since this gives an opportunity for the user to weight the considered signals according to his own choice. The RGA method is scaling independent and offer hence not this possibility.

Note, however, that the RGA also acts as a warning tool since it indicates if a certain pairing should be avoided (with a negative RGA element, see Section 2.1.2). The HIIA is unable to do this.

4.9.4 Practical Implications

It was seen in the analysis of the considered bioreactor models that the RGA is misleading in some cases. Hence, based on the RGA results, it should not be concluded, in this particular case (in the medium time scale and with no external carbon dosage), that the couplings are low between the DO setpoint ($S_O(2)$) and the internal recirculation flow rate (Q_i) independent of operation point, as concluded by Ingildsen (2002). Instead, the operational maps indicate that there are some couplings between the nitrification and the denitrification process. Hence, a MIMO control structure can be expected to give better control performance compared to a solution involving SISO controllers. The HIIA analysis supports this view.

Appendix A

Parameter values

Table A.1 shows the parameter values that were used in the simulations of the models (3.1), (3.3) and (3.1)m¹.

Table A.1: *Parameter values used for the models (3.1), (3.3) and (3.1)m.*

Parameter	Value	Unit	Comment
S_{NH_i}	31.56	g N m ⁻³	Influent ammonium conc.
S_{S_i}	$69.5 + S_{S_{dosage}}$	g COD m ⁻³	Influent soluble substrate.
$S_{S_{dosage}}$	0	g COD m ⁻³	In Section 4.2 this value is set to 90 g COD m ⁻³ .
Q	18446	m ³ day ⁻¹	
V_1	2000	m ³	1000 m ³ for model (3.1)m.
V_2	3999	m ³	1000 m ³ for model (3.1)m.
η_g	0.8	-	
i_{XB}	0.08	-	
K_{NH}	1.0	g NH ₃ -N m ⁻³	
K_{NO}	0.5	g NO ₃ -N m ⁻³	
$K_{O,A}$	0.4	g O ₂ m ⁻³	
$K_{O,H}$	0.2	g O ₂ m ⁻³	
K_S	10.0	g COD m ⁻³	
μ_A	0.5	day ⁻¹	
μ_H	4.0	day ⁻¹	
$X_{B,A}$	150	g COD m ⁻³	
$X_{B,H}$	2500	g COD m ⁻³	
Y_A	0.24	-	
Y_H	0.67	-	

¹Model (3.1)m is a modified version of model (3.1) with $V_1 = V_2 = 1000$ m³.

Appendix B

HIIA Matrices Obtained for the Analysed Bioreactor Models

In this appendix, numerical values of the HIIA matrices obtained for the different analysed bioreactor models will be given.

B.1 HIIA Matrices for Model (3.3)

The HIIA matrices obtained for the unscaled systems from model (3.3) are:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000092 & 0.0736742 \\ 0.0000489 & 0.9262677 \end{bmatrix} \quad (\text{B.1a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000158 & 0.0963259 \\ 0.0000696 & 0.9035887 \end{bmatrix} \quad (\text{B.1b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0002624 & 0.7271456 \\ 0.0055892 & 0.2670028 \end{bmatrix} \quad (\text{B.1c})$$

If the systems are scaled according to (4.4) (but with the upper left element in D_u replaced by 10000 and 50000 for operating point u_1 and u_3 , respectively), then the following HIIA matrices are obtained:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0942255 & 0.0300417 \\ 0.4980345 & 0.3776983 \end{bmatrix} \quad (\text{B.2a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.1132591 & 0.0374055 \\ 0.4984516 & 0.3508838 \end{bmatrix} \quad (\text{B.2b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0433659 & 0.0240363 \\ 0.9237719 & 0.0088260 \end{bmatrix} \quad (\text{B.2c})$$

60 HIIA Matrices Obtained for the Analysed Bioreactor Models

For the filtered but unscaled systems, the HIIA matrices are:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000000 & 0.0781829 \\ 0.0000417 & 0.9217754 \end{bmatrix} \quad (\text{B.3a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000000 & 0.1036195 \\ 0.0000552 & 0.8963253 \end{bmatrix} \quad (\text{B.3b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0000002 & 0.8002400 \\ 0.0053861 & 0.1943737 \end{bmatrix} \quad (\text{B.3c})$$

B.2 HIIA Matrices for the Modified Version of Model (3.1)

The HIIA matrices obtained for the unscaled systems from model (3.1)m are:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000135 & 0.5137484 \\ 0.0000156 & 0.4862225 \end{bmatrix} \quad (\text{B.4a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000212 & 0.5235154 \\ 0.0000208 & 0.4764426 \end{bmatrix} \quad (\text{B.4b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0000898 & 0.5911514 \\ 0.0001731 & 0.4085858 \end{bmatrix} \quad (\text{B.4c})$$

If the systems are scaled according to (4.4) (but with the upper left element in D_u replaced by 10000 and 50000 for operating point u_1 and u_3 , respectively), then the following HIIA matrices are obtained:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.1954050 & 0.2973094 \\ 0.2259057 & 0.2813800 \end{bmatrix} \quad (\text{B.5a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.2201956 & 0.2950865 \\ 0.2161646 & 0.2685533 \end{bmatrix} \quad (\text{B.5b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.1939536 & 0.2554917 \\ 0.3739667 & 0.1765880 \end{bmatrix} \quad (\text{B.5c})$$

For the filtered but unscaled systems, the HIIA matrices are:

$$\Sigma_H^{u_3} = \begin{bmatrix} 0.0000000 & 0.5136244 \\ 0.0000022 & 0.4863734 \end{bmatrix} \quad (\text{B.6a})$$

$$\Sigma_H^{u_2} = \begin{bmatrix} 0.0000000 & 0.5232969 \\ 0.0000004 & 0.4767027 \end{bmatrix} \quad (\text{B.6b})$$

$$\Sigma_H^{u_1} = \begin{bmatrix} 0.0000011 & 0.6095744 \\ 0.0001696 & 0.3902549 \end{bmatrix} \quad (\text{B.6c})$$

Bibliography

- Birk, W. (2002). Industry Application of Multivariable Control. PhD thesis. Luleå University of Technology. Luleå, Sweden.
- Bristol, E. H. (1966). On a new measure of interaction for multivariable process control. *IEEE Trans. Automatic Control* **AC-11**, 133–134.
- Chien, J., J.S. Freudenberg and C. N. Nett (1992). On relative gain array and condition number. In: *Proceedings of the 31st Conference on Decision and Control*. Tuscon, Arizona, USA. pp. 219–224.
- Chiu, M.-S. and Y. Arkun (1991). A new result on relative gain array, Niederlinski index and decentralized stability condition: 2x2 plant cases. *Automatica* **27**(2), 419–421.
- Conley, A. and M. E. Salgado (2000). Gramian based interaction measure. In: *Proceedings of the 39th IEEE Conference on Decision and Control*. Sydney, Australia. pp. 5020–5022.
- Copp, J. B., Ed.) (2002). *EUR 19993 – COST Action 624 – The COST simulation benchmark – Description and simulator manual*. European Communities. Luxembourg.
- Glad, T. and L. Ljung (1997). *Reglerteori*. Studentlitteratur. Lund, Sweden. In Swedish.
- Gray, W. S. and J. M. A. Scherpen (1998). Hankel operators and gramians for nonlinear systems. In: *Proc. CDC98*. Tampa, Florida, USA. pp. 1416–1421.
- Grosdidier, P., M. Morari and B. R. Holt (1985). Closed-loop properties from steady-state gain information. *Ind. Eng. Chem. Fundam.* **24**, 221–235.
- Hägglom, K. E. (1997). Control structure analysis by partial relative gains. In: *Proceedings of the 36th Conference on Decision and Control*. San Diego, California, USA. pp. 2623–2624.

- Henze, M., C. P. L. Grady Jr., W. Gujer, G. v. R. Marais and T. Matsuo (1986). Activated sludge model no. 1. Scientific and Technical Report No. 1. IAWPRC, London, UK.
- Ingildsen, P. (2002). Realising Full-Scale Control in Wastewater Treatment Systems Using In Situ Nutrient Sensors. PhD thesis. Lund University, Sweden.
- Kinnaert, M. (1995). Interaction measures and pairing of controlled and manipulated variables for multiple-input multiple-output systems: A survey. *Journal A* **36**(4), 15–23.
- Kommunförbundet (1988). *Introduktion till avloppstekniken*. Sweden. In Swedish.
- Lindberg, C-F. (1997). Control and estimation strategies applied to the activated sludge process. PhD thesis. Uppsala University. Uppsala, Sweden.
- Maciejowski, J. M. (1989). *Multivariable feedback design*. Addison-Wesley.
- Metcalf and Eddy Inc. (1991). *Wastewater Engineering, Treatment, Disposal and Reuse*. third ed.. McGraw-Hill International Editions.
- Morari, M. and E. Zafiriou (1989). *Robust Process Control*. Prentice-Hall. London, UK.
- Olsson, G. and B. Newell (1999). *Wastewater Treatment Systems*. IWA publishing. London, UK.
- Olsson, G. and U. Jeppsson (1994). Establishing cause-effect relationships in activated sludge plants – what can be controlled. In: *Proceedings of workshop modelling, monitoring and control of wastewater treatment plants*. pp. 2057–2070.
- Rugh, W. J. (1996). *Linear system theory*. second ed.. Prentice Hall. Upper Saddle River, New Jersey, USA.
- Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control*. John Wiley & Sons. Chichester, UK.
- Takács, I., G. G. Patry and D. Nolasco (1991). A dynamic model of the clarification-thickening process. *Wat. Res.* **25**(10), 1263–1271.
- Waller, J. B. and K. V. Waller (1995). Defining directionality: Use of directionality measures with respect to scaling. *Ind. Eng. Chem. Res.* **34**, 1244–1252.

-
- Waller, J. B., M. F. Sågfors and K. V. Waller (1994). Ill-conditionedness and process directionality – the use of condition numbers in process control. In: *Proceedings, IFAC Symposium ADCHEM '94*. Kyoto, Japan. pp. 465–470.
- Wittenmark, B. and M. E. Salgado (2002). Hankel-norm based interaction measure for input-output pairing. In: *The 2002 IFAC World Congress*. Barcelona, Spain.
- Wittenmark, B., K. J. Åström and S. B. Jørgensen (1995). *Process Control*. Dep of Automatic Control, Lund University, Lund, Sweden.