Supporting Information

Integrated Optimisation of Upstream and Downstream Processing in Biopharmaceutical Manufacturing under Uncertainty: A Chance Constrained Programming Approach

Songsong Liu[a], Suzanne S. Farid[b], Lazaros G. Papageorgiou[a,*]

[a] *Centre for Process Systems Engineering, Department of Chemical Engineering, UCL (University College London), London WC1E 7JE, United Kingdom*

[b] *The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, UCL (University College London), London WC1E 7JE, United Kingdom*

* *Corresponding Author*

1. Literature model

The literature MILP model[19] for the optimal chromatography strategies of antibody purification processes is presented as follows.

The integer variables, CN_{si}, CYN_s and BN, are expressed by binary variables:

$$CN_{si} = \sum_{j=1}^{j_s} j \cdot W_{sij}\,,\ \forall s \in CS, i \tag{S.1}$$

$$\sum_{j=1}^{j_s} W_{sij} = X_{si},\ \forall s \in CS, i \tag{S.2}$$

$$CYN_s = \sum_{k=1}^{k_s} k \cdot Y_{sk}\,,\ \forall s \in CS \tag{S.3}$$

$$\sum_{k=1}^{k_s} Y_{sk} = 1,\ \forall s \in CS \tag{S.4}$$

$$BN = \sum_{n=1}^{q} 2^{n-1} \cdot Z_n \tag{S.5}$$

For each packed-bed chromatography step, only one resin can be used:

$$\sum_{r\in R_s} U_{sr} = 1,\ \forall s \in CS \tag{S.6}$$

At most one resin of each resin type can be used:

$$\sum_{s\in CS} \sum_{r\in R_s \cap R_t} U_{sr} \le 1,\ \forall t \tag{S.7}$$

The initial product protein mass, M_0, is the protein mass from the upstream bioreactors:

$$M_0 = titre \cdot \alpha \cdot brv \tag{S.8}$$

The product protein mass remaining after step s, M_s is related to the yield of the step:

$$M_s = ncy_s \cdot M_{s-1},\ \forall s \notin CS \tag{S.9}$$

$$M_s = \sum_{r\in R_s} (cy_{sr} \cdot \overline{UM}_{s-1,r})\,,\ \forall s \in CS \tag{S.10}$$

$$\overline{UM}_{s-1,r} \le titre \cdot \alpha \cdot brv \cdot U_{sr},\ \ \forall s \in CS, r \in R_s \tag{S.11}$$

$$\sum_{r\in R_s} \overline{UM}_{s-1,r} = M_{s-1},\quad \forall s \in CS \tag{S.12}$$

The annual product output, AP, is amount of product produced per year by the facility:

$$AP = \sum_{n=1}^{q} \sigma \cdot 2^{n-1} \cdot \overline{ZM}_{sn}\,,\ \forall s = bf \tag{S.13}$$

$$\overline{ZM}_{sn} \le titre \cdot \alpha \cdot brv \cdot Z_n,\ \forall s = bf, n = 1, \dots, q \tag{S.14}$$

$$\overline{ZM}_{sn} \le M_s,\ \forall s = bf, n = 1, \dots, q \tag{S.15}$$

$$\overline{ZM}_{sn} \ge M_s - titre \cdot \alpha \cdot brv \cdot (1 - Z_n),\ \forall s = bf, n = 1, \dots, q \tag{S.16}$$

The number of completed batches, BN, is limited by an upper bound:

$$BN \leq maxbn \tag{S.17}$$

The total column volume of chromatography step s, TCV_s, is the number of columns, CN_{si}, multiplied by the single column volume, cv_{si}:

$$TCV_s = \sum_i cv_{si} \cdot CN_{si}\,, \ \forall s \in CS \tag{S.18}$$

Only one column size is allowed at each chromatography step.

$$\sum_i X_{si} = 1, \ \forall s \in CS \tag{S.19}$$

$$CN_{si} \leq maxcn_s \cdot X_{si}, \ \forall s \in CS, i \tag{S.20}$$

The total amount of resin available is no less than the minimum required amount, RV_s:

$$\sum_{k=1}^{k_s} k \cdot \overline{YV}_{sk} \geq RV_s, \ \forall s \in CS \tag{S.21}$$

$$\overline{YV}_{sk} \leq maxcn_s \cdot maxcv_s \cdot Y_{sk}, \ \forall s \in CS, k = 1, \dots, k_s \tag{S.22}$$

$$\sum_{k=1}^{k_s} \overline{YV}_{sk} = TCV_s, \ \forall s \in CS \tag{S.23}$$

$$RV_s = \sum_{r \in R_s} \frac{\overline{UM}_{s-1,r}}{dbc_r \cdot \mu}, \ \forall s \in CS \tag{S.24}$$

The number of cycles, CYN_s, at each chromatography step has an upper bound:

$$CYN_s \leq maxcyn_s, \ \forall s \in CS \tag{S.25}$$

Volumetric flow rate, VFR_s, is determined by the velocity of flow and the diameter of the column:

$$VFR_s = \frac{1}{1000} \cdot \frac{1}{60} \cdot \sum_{r \in R_s} \sum_i vel_r \cdot \pi \cdot (\frac{dm_{si}}{2})^2 \cdot \overline{UX}_{sri}\,, \ \forall s \in CS \tag{S.26}$$

$$\sum_{r \in R_s} \overline{UX}_{sri} = X_{si}, \ \forall s \in CS, i \tag{S.27}$$

$$\sum_i \overline{UX}_{sri} = U_{sr}, \ \forall s \in CS, r \in R_s \tag{S.28}$$

The initial product volume entering downstream processes, PV_0, is the working volume of bioreactor:

$$PV_0 = \alpha \cdot brv \tag{S.29}$$

The product volume remaining after each step s, PV_s, and the required buffer material volume at each step s, BV_s, are given in Eq. (A30)-(S.43):

$$PV_s = (fvr_s + 1) \cdot PV_0, \ \forall s = h \tag{S.30}$$

$$BV_s = fvr_s \cdot PV_0, \ \forall s = h \tag{S.31}$$

$$PV_s = \sum_{r \in R_s \cap BER} \sum_{k=1}^{k_s} ecv_r \cdot k \cdot \overline{UYV}_{srk} + \sum_{r \in R_s \cap FTR} \overline{UV}_{s-1,r}, \ \forall s \in CS \tag{S.32}$$

$$BV_s = \sum_{r \in R_s} \sum_{k=1}^{k_s} bcv_r \cdot k \cdot \overline{UYV}_{srk}, \ \forall s \in CS \tag{S.33}$$

$$\overline{UYV}_{srk} \leq maxcn_s \cdot maxcv_s \cdot U_{sr}, \ \forall s \in CS, r \in R_s, k = 1, \ldots, k_s \tag{S.34}$$

$$\sum_{r \in R_s} \overline{UYV}_{srk} = \overline{YV}_{sk}, \ \forall s \in CS, k = 1, \ldots, k_s \tag{S.35}$$

$$\overline{UV}_{s-1,r} \leq maxpv_{s-1} \cdot U_{sr}, \ \forall s \in CS, r \in R_s \tag{S.36}$$

$$\sum_{r \in R_s} \overline{UV}_{s-1,r} = PV_{s-1}, \ \forall s \in CS \tag{S.37}$$

$$PV_s = (nvr_s + 1) \cdot PV_{s-1}, \ \forall s = vi \tag{S.38}$$

$$BV_s = nvr_s \cdot BV_{s-1}, \ \forall s = vi \tag{S.39}$$

$$PV_s = (fvr_s + 1) \cdot PV_{s-1}, \ \forall s = vf \tag{S.40}$$

$$BV_s = fvr_s \cdot BV_{s-1}, \ \forall s = vf \tag{S.41}$$

$$PV_s = \frac{M_s}{fconc}, \ \forall s = ufdf \tag{S.42}$$

$$BV_s = dvr_s \cdot \frac{M_s}{fconc}, \ \forall s = ufdf \tag{S.43}$$

The total buffer usage per batch, BBV, is the summation of buffer usage, BV_s, in all downstream steps:

$$BBV = \sum_s BV_s \tag{S.44}$$

The annual total buffer volume, ABV, is related to the number of completed batches:

$$ABV = \sum_{n=1}^{q} 2^{n-1} \cdot \overline{ZV}_n \tag{S.45}$$

$$\overline{ZV}_n \leq maxbbv \cdot Z_n, \ \forall n = 1, \ldots, q \tag{S.46}$$

$$\overline{ZV}_n \leq BBV, \ \forall n = 1, \ldots, q \tag{S.47}$$

$$\overline{ZV}_n \geq BBV - maxbbv \cdot (1 - Z_n), \ \forall n = 1, \ldots, q \tag{S.48}$$

The total processing time at each chromatography step, T_s, is comprised of processing time for both adding buffer (PLT_s) and loading product (BAT_s):

$$T_s = PLT_s + BAT_s, \ \forall s \in CS \tag{S.49}$$

The processing time for loading product, PLT_s is related to the incoming product volume:

$$\frac{1}{1000}\cdot\frac{1}{60}\cdot\sum_{r\in R_s}\sum_i\sum_{j=1}^{j_s} vel_r\cdot\pi\cdot(\frac{dm_{si}}{2})^2\cdot j\cdot\overline{UWT}_{srij} = PV_{s-1},\ \forall s\in CS \tag{S.50}$$

$$\overline{UWT}_{srij}\leq brt\cdot W_{sij},\ \forall s\in CS, r\in R_s, i, j=1,\ldots,j_s \tag{S.51}$$

$$\overline{UWT}_{srij}\leq brt\cdot U_{sr},\ \forall s\in CS, r\in R_s, i, j=1,\ldots,j_s \tag{S.52}$$

$$\sum_i\sum_{j=1}^{j_s}\sum_{r\in R_s}\overline{UWT}_{srij} = PLT_s,\ \forall s\in CS \tag{S.53}$$

The processing time for adding buffer, BAT_s is related to the required buffer volume:

$$BAT_s = \sum_{r\in R_s}\sum_i\sum_{k=1}^{k_s}\frac{bcv_r\cdot cv_{si}\cdot k\cdot\overline{UXY}_{srik}}{\frac{1}{1000}\cdot\frac{1}{60}\cdot vel_s\cdot\pi\cdot(\frac{dm_{si}}{2})^2},\ \forall s\in CS \tag{S.54}$$

$$\sum_{r\in R_s}\sum_i\overline{UXY}_{srik} = Y_{sk},\ \forall s\in CS, k=1,\ldots,k_s \tag{S.55}$$

$$\sum_{k=1}^{k_s}\overline{UXY}_{srik} = \overline{UX}_{sri},\ \forall s\in CS, r\in R_s, i \tag{S.56}$$

The processing time per batch, BT, is the summation of processing time of all downstream steps:

$$BT = \frac{\sum_s T_s}{60\cdot sfd\cdot sfn} \tag{S.57}$$

The annual DSP time, AT, is related to the number of completed batches:

$$AT = \sum_{n=1}^{q} 2^{n-1}\cdot\overline{ZT}_n \tag{S.58}$$

$$\overline{ZT}_n\leq(aot-st-brt)\cdot Z_n,\ \forall n=1,\ldots,q \tag{S.59}$$

$$\overline{ZT}_n\leq BT,\ \forall n=1,\ldots,q \tag{S.60}$$

$$\overline{ZT}_n\geq BT-(aot-st-brt)\cdot(1-Z_n),\ \forall n=1,\ldots,q \tag{S.61}$$

The annual DSP time, AT, cannot exceed the annual available time:

$$AT\leq aot-st-brt \tag{S.62}$$

The labour cost, LC, involves the direct labour cost, DLC, supervisors cost, SC, quality control and quality assurance (QCQA) cost, QC, and management cost, MC:

$$LC = DLC+SC+QC+MC \tag{S.63}$$

$$DLC = 24\cdot uon\cdot w\cdot brt\cdot BN+\ don\cdot w\cdot sfd\cdot sfn\cdot AT \tag{S.64}$$

$$SC = s\lambda\cdot DLC \tag{S.65}$$

$$QC = q\lambda\cdot DLC \tag{S.66}$$

$$MC = m\lambda\cdot DLC \tag{S.67}$$

The chemical reagents cost, CRC, is assumed to include the cost for buffer, BC, and bioreactor media, MEC:

$$CRC = BC + MEC \tag{S.68}$$

$$BC = bpc \cdot ABV \tag{S.69}$$

$$MEC = \theta \cdot mepc \cdot \alpha \cdot brv \cdot BN \tag{S.70}$$

The key consumables cost, CC, in this study is the resin cost:

$$CC = \sum_{s \in CS} \sum_{r \in R_s} \sum_{n=1}^{q} \sum_{k=1}^{k_s} \frac{of \cdot rpc_r \cdot 2^{n-1} \cdot k \cdot \overline{ZUYV}_{srkn}}{l_r} \tag{S.71}$$

$$\overline{ZUYV}_{srkn} \leq maxtcv_s \cdot Z_n, \ \forall s \in CS, r \in R_s, k = 1, \dots, k_s, n = 1, \dots, q \tag{S.72}$$

$$\overline{ZUYV}_{srkn} \leq \overline{UYV}_{srk}, \ \forall s \in CS, r \in R_s, k = 1, \dots, k_s, n = 1, \dots, q \tag{S.73}$$

$$\overline{ZUYV}_{srkn} \geq \overline{UYV}_{sk} - maxtcv_s \cdot (1 - Z_n), \ \forall s \in CS, r \in R_s, k = 1, \dots, k_s, n = 1, \dots, q \tag{S.74}$$

The miscellaneous material cost, MIC, is proportional to the total chemical reagents cost, CRC, and consumables cost, CC.

$$MIC = mi\lambda \cdot (CRC + CC) \tag{S.75}$$

The utilities cost, UC, can be expressed as the summation of three terms:

$$UC = a \cdot brn \cdot brv + b \cdot brv \cdot BN + c \cdot ABV \tag{S.76}$$

The annualised capital cost, CAC, is calculated by the fixed capital investment, FCI, and the capital recovery factor:

$$CAC = FCI \cdot \frac{r \cdot (1+r)^{el}}{(1+r)^{el} - 1} \tag{S.77}$$

$$FCI = lang \cdot (1 + gef) \cdot (brn \cdot brc + \sum_{s \in CS} \sum_i cc_{si} \cdot CN_{si} + oe\lambda \cdot brc \cdot brn) \tag{S.78}$$

Other indirect costs include the annual maintenance cost, MAC, insurance cost, IC, local tax costs, TC, and general utilities cost, GUC:

$$MAC = ma\lambda \cdot FCI \tag{S.79}$$

$$IC = i\lambda \cdot FCI \tag{S.80}$$

$$TC = t\lambda \cdot FCI \tag{S.81}$$

$$GUC = gu \cdot brn \cdot brv \tag{S.82}$$

$$OIC = MAC + IC + TC + GUC \tag{S.83}$$

The annual total cost of goods is the summation of the above costs:

$$COG = LC + CRC + CC + MIC + UC + CAC + OIC \quad (S.84)$$

The objective is to minimise COG/g:

$$OBJ = \frac{COG}{AP} \quad (S.85)$$

Nomenclature

Indices

bf	bulk fill step
h	harvest step
i	column size
j	column number
k	cycle number
n	digit of the binary representation
r	resin
s	downstream step
t	resin type
ufdf	UF/DF step
vf	virus filtration step
vi	virus inactivation step

Sets

BER	set of resins in bind-elute mode
CS	set of chromatography steps, = capture, intermediate purification, polishing
FTR	set of resins in flow-through mode
R_s	set of resins suitable to chromatography step *s*
R_t	set of resins of the resin type *t*

Parameters

a, b, c	utilities cost coefficients
aot	annual operating time, day
bcv_r	buffer usage of resin *r*, CV
bpc	buffer price, £/L
brc	bioreactor cost at given discrete volume for piecewise approximation, £
brf	scale-up factor of bioreactor cost
brn	number of bioreactors
brt	bioreaction time, day
brv	given discrete bioreactor volume for piecewise approximation, L
cc_{si}	column cost of size *i* at chromatography step *s*, £
cf	scale-up factor of column cost
cv_{si}	volume of column size *i* at chromatography step *s*, L
cy_{sr}	product yield of resin *r* at chromatography step *s*
dbc_r	dynamic binding capacity of resin *r*, g/L
dm_{si}	diameter of column size *i* at chromatography step *s*, L

don	number of operators for downstream processing
dvr_s	diafiltration volume ratio of step *s*
ecv_r	elute volume of resin *r*, CV
el	equipment lifetime, year
$fconc$	final concentration of product, g/L
fvr_s	flush volume ratio of step *s*
gef	general equipment factor
gu	general utility unit cost, £/L
h_{si}	height of column size *i* at step *s*, cm
$i\lambda$	insurance cost ratio to the fixed capital investment
j_s	maximum number of columns at chromatography step *s*, $maxcn_s$
k_s	maximum number of cycles at chromatography step *s*, $maxcyn_s$
l_r	life time of resin *r*, cycle
$lang$	Lang factor
$maxbbv$	maximum buffer volume per batch
$maxbn$	maximum number of batches
$maxbrv$	maximum bioreactor volume
$maxcn_s$	maximum number of columns at chromatography step *s*
$maxcv_s$	maximum column volume at chromatography step *s*
$maxcyn_s$	maximum number of cycles at chromatography step *s*
$maxpv_s$	maximum product volume at step *s*
$ma\lambda$	maintenance cost ratio to the fixed capital investment
$mepc$	media price, £/L
$mi\lambda$	miscellaneous material cost ratio to chemical reagent and consumable costs
$m\lambda$	management cost ratio to direct labour cost
ncy_s	product yield of non-chromatography step *s*
nvr_s	neutralisation volume ratio of step *s*
$oe\lambda$	other equipment cost ratio to the bioreactor cost
of	overpacking factor of resin
q	maximum digit number in the binary representation of number of batches, $\lceil \log_2 maxbn \rceil$
$q\lambda$	QCQA cost ratio to direct labour cost
r	interest rate
rpc_r	resin price of resin *r*, £/L
$refbrc$	reference cost of a bioreactor, £
$refbrv$	reference volume of a bioreactor, L
$refcc$	reference cost of a chromatography column, £
$refdm$	reference diameter of a chromatography column, cm
sfd	duration per shift, hour
sfn	number of shifts per day
st	seed train bioreaction time, day
$s\lambda$	supervisors cost ratio to direct labour cost
$titre$	upstream product titre, g/L
$t\lambda$	tax cost ratio to the fixed capital investment
uon	number of operators per bioreactor in upstream processing
vel_r	linear velocity of flow for resin *r*, cm/h
w	wage of an operator, £/h
α	bioreactor working volume ratio
θ	media overfill allowance

μ chromatography resin utilisation factor
σ batch success rate

Continuous Variables

ABV annual buffer volume, L
AP annual product output, g
AT annual downstream operating time, day
BAT_s time for adding buffer per batch at chromatography step *s*, min
BBV buffer volume added per batch, L
BC buffer cost, £
BRC bioreactor cost, £
BT downstream processing time per batch, day
BV_s buffer volume per batch in chromatography step *s*, L
CAC capital cost, £
CC consumables cost, £
COG annual cost of goods, £
CRC chemical reagents cost, £
DLC direct labour cost, £
FCI fixed capital investment, £
GUC general utility cost, £
IC insurance cost, £
LC labour cost, £
M_0 initial product mass entering downstream processes per batch, g
M_s initial product mass per batch after step *s*, g
MAC maintenance cost, £
MC management cost, £
MEC media cost, £
MIC miscellaneous material cost, £
OBJ objective
OIC other indirect costs, £
PLT_s time for loading product per batch at chromatography step *s*, min
PV_0 initial product volume entering downstream processes per batch, L
PV_s product volume per batch after step *s*, L
QC QCQA cost, £
RV_s resin volume required at chromatography step *s*, L
SC supervisors cost, £
T_s processing time per batch of step *s*, min
TC tax cost, £
TCV_s total column volume at chromatography step *s*, L
UC utilities cost, £
VFR_s volumetric flow rate at chromatography step *s*, L/min

Binary Variables

U_{sr} 1 if resin *r* is selected at chromatography step *s*; 0 otherwise
W_{sij} 1 if there are *j* columns of size *i* at chromatography step *s*; 0 otherwise
X_{si} 1 if column size *i* is selected at chromatography step *s*; 0 otherwise
Y_{sk} 1 if there are *k* cycles at chromatography step *s*; 0 otherwise
Z_n 1 if the *n*th digit of the binary representation of variable BN is equal to 1; 0 otherwise

Integer Variables

BN	number of completed batches
CN_{si}	number of columns of size *i* at chromatography step *s*
CYN_s	number of cycles at chromatography step *s*

Auxiliary Variables

$\overline{UM}_{s-1,r}$	$\equiv U_{sr} \cdot M_{s-1}$
$\overline{UV}_{s-1,r}$	$\equiv U_{sr} \cdot PV_{s-1}$
$\overline{UWT}_{srij}$	$\equiv U_{sr} \cdot W_{sij} \cdot PLT_s$
$\overline{UX}_{sri}$	$\equiv U_{sr} \cdot X_{si}$
$\overline{UXY}_{srik}$	$\equiv U_{sr} \cdot X_{si} \cdot Y_{sk}$
$\overline{UYV}_{srk}$	$\equiv U_{sr} \cdot Y_{sk} \cdot TCV_s$
$\overline{YV}_{sk}$	$\equiv Y_{sk} \cdot TCV_s$
$\overline{ZM}_{sn}$	$\equiv Z_n \cdot M_s$
$\overline{ZT}_n$	$\equiv Z_n \cdot BT$
$\overline{ZV}_n$	$\equiv Z_n \cdot BBV$
$\overline{ZUYV}_{srkn}$	$\equiv Z_n \cdot U_{sr} \cdot Y_{sk} \cdot TCV_s$

2. Case study data

More data of the case study are presented in Tables S1 and S2.

Table S1. Data for non-chromatographic unit operations

Unit operation parameter	Value	Unit operation parameter	Value
Cell culture		processing time (h)	1.5
bioreaction time (days)	15	*Virus filtration*	
seed train bioreaction time (days)	29	yield (%)	95
bioreactor working volume ratio (%)	75	flush volume ratio	0.3
media overfill factor	1.2	processing time (h)	4
media price (£/L)	32	*Ultra/Diafiltration*	
Harvest		yield (%)	90
yield (%)	95	processing time (h)	4
flush volume ratio	0.1	final concentration (g/L)	75
processing time (h)	4	diafiltration volume	7
Virus inactivation		*Bulk fill*	
yield (%)	90	yield (%)	98
neutralisation volume ratio	1.75	filling time (h)	6

Table S2. More data for cost and time

Parameters	**Value**	**Parameters**	**Value**
aot (day)	340	a (£/L)	14.145
don (day)	15	b (£/L)	4.234
uon (day)	3	c (£/L)	0.071
sfd (hours/shift)	8	$s\lambda$	0.2
sfn (shift/day)	1	$q\lambda$	1
w (£/h)	20	$m\lambda$	1
bpc (£/L)	1	$mi\lambda$	0.1
σ	90%	$oe\lambda$	0.8
gef	0.7	$ma\lambda$	0.05
gu (£/L)	90	$i\lambda$	0.005
$lang$	6	$t\lambda$	0.01