Adv Biochem Eng Biotechnol (2021) 177: 1–28
https://doi.org/10.1007/10_2020_154
© Springer Nature Switzerland AG 2020
Published online: 31 December 2020

# Potential of Integrating Model-Based Design of Experiments Approaches and Process Analytical Technologies for Bioprocess Scale-Down


**Peter Neubauer, Emmanuel Anane, Stefan Junne, and Mariano Nicolas Cruz Bournazou**

## Contents



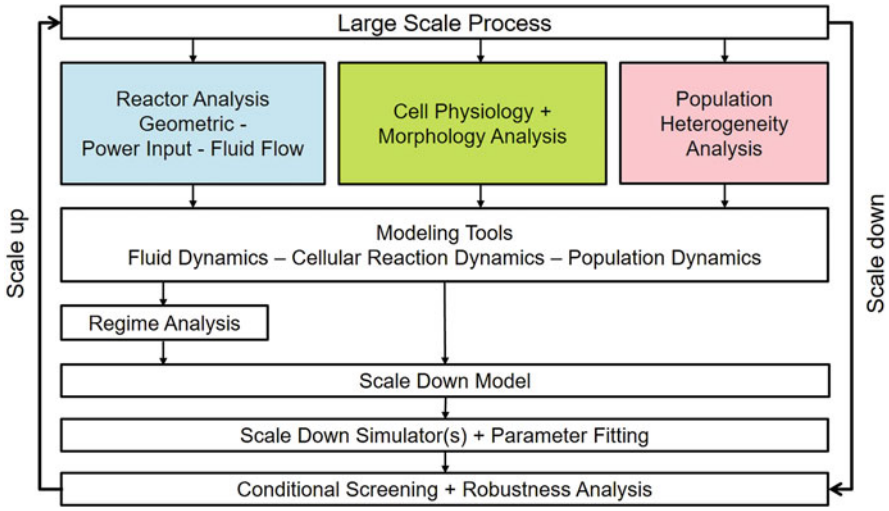
**Abstract** Typically, bioprocesses on an industrial scale are dynamic systems with a certain degree of variability, system inhomogeneities, and even population heterogeneities. Therefore, the scaling of such processes from laboratory to industrial scale and vice versa is not a trivial task. Traditional scale-down methodologies consider several technical parameters, so that systems on the laboratory scale tend to qualitatively reflect large-scale effects, but not the dynamic situation in an industrial bioreactor over the entire process, from the perspective of a cell. Supported by the enormous increase in computing power, the latest scientific focus is on the

P. Neubauer (✉), E. Anane, and S. Junne
Bioprocess Engineering, Faculty III Process Sciences, Institute of Biotechnology, Technische Universität Berlin (TU Berlin), Berlin, Germany
e-mail: peter.neubauer@tu-berlin.de

M. N. Cruz Bournazou
DataHow AG, Zurich, Switzerland

application of dynamic models, in combination with computational fluid dynamics to quantitatively describe cell behavior. These models allow the description of possible cellular lifelines which in turn can be used to derive a regime analysis for scale-down experiments. However, the approaches described so far, which were for a very few process examples, are very labor- and time-intensive and cannot be validated easily. In parallel, alternatives have been developed based on the description of the industrial process with hybrid process models, which describe a process mechanistically as far as possible in order to determine the essential process parameters with their respective variances. On-line analytical methods allow the characterization of population heterogeneity directly in the process. This detailed information from the industrial process can be used in laboratory screening systems to select relevant conditions in which the cell and process related parameters reflect the situation in the industrial scale. In our opinion, these technologies, which are available in research for modeling biological systems, in combination with process analytical techniques are so far developed that they can be implemented in industrial routines for faster development of new processes and optimization of existing ones.

**Graphical Abstract**




**Keywords** Bioprocess scale-up, Process analytical techniques, Process modeling, Scale-down

## 1 Status of Bioprocess Scale-Down: The Need for a Model-Based Design

Studies on the scale-down of bioprocesses have received much attention in recent years. On the one side, this is due to an increasing implementation of the concept of a circular bio-economy. Within this framework there is a boost in new processes and

strategies. In the field of biotechnological production of basic chemicals and biocatalysts, the reactor volumes are getting steadily larger to ensure the necessary yields. Newly conceived processes must be robustly feasible on a large scale [1]. On the other side, in the field of biopharmaceutical production, product yields in the reactor originally had a lower priority for originator products compared to the costs incurred in downstream processing. The expiry of many patents and the development of generic or biosimilar products has led to a price pressure that companies can only counteract with very efficient and less variable bioprocesses. Additionally, the implementation of single-use strategies, especially in connection with the lower power input in single-use bioreactors, introduces scaling phenomena, i.e. imperfect mixing issues already at much smaller scales, i.e. in reactors with 1–5 $m^3$ [2]. Still, the main reason for the increased interest in scale-down investigations are failures, lower yields, higher batch-to-batch variability, or even changes in the product quality of processes performed in industrial scale. The difficulties to predict the outcomes at this scale, the large number of unexpected responses of cells to different environments, and the impossibility of computer based tools to foresee the changes on the phenotype throughout scales based on laboratory data are the driving forces behind scale-down popularity [1].

The discrepancy between laboratory- and industrial-scale fermentation processes results from heterogeneous environments in large-scale bioreactors due to the limited volumetric power input and geometric issues, which result in longer mixing times associated with the increasing volumes at industrial scale, compared to laboratory scale bioreactors [3]. The effects of such process inhomogeneity on microbial physiology and product syntheses, including the quantity and quality of recombinant products, have attracted much attention in the bioprocess research community, due to the mostly unforeseeable impacts of the heterogeneities on process efficiency, e.g. by the accidental incorporation of non-canonical amino acids into the product [4].

In the past three decades, various forms of single-compartment and multi-compartment scale-down bioreactors have been developed to study scale-up effects in fermentation processes [3]. This development, however, is accompanied by a constant discussion about the extent, to which these systems really reflect the conditions at an industrial scale. Concrete proposals for procedures for scaling down a process to laboratory scale have only been developed in recent years, see, e.g. [5], but are generally very sophisticated and therefore unsuitable for broad application.

In parallel, during the last years, there has been a phenomenal increase in the use of high-throughput (HT) miniaturized bioreactor systems for strain screening and bioprocess development, which has significantly reduced the times required for early bioprocess development. These new powerful laboratory tools require, however, new methods for planning, performing, and evaluating these highly parallel experiments. The systems are no longer treatable by manual methods – therefore, standard methods of design, mathematics and statistics, modeling and process engineering as they have been used in other disciplines for a long time have to be implemented and adapted in the field of bioprocessing. Intuitively, when dealing with large data sets,

highly automated systems and closely interconnected devices, concepts like Internet of Things (IoT) and Digital Twin come to mind. Beyond the hype around Digital Twins, its history and modern definition are closely related to High-Throughput Bioprocess Development. The term Digital Twin emerged in the field of Product Lifecycle Management to increase the efficiency in product and process development [6]. Mathematical models, used for process monitoring (observers) in feedback loops, for approximate optimal control applications (MPC), or even in real-time optimization have existed for quite some time now [7]. Yet, the extension of these methods including IoT, big data, and fully autonomous systems might require a new terminology [8].

The Digital Twin, envisioned as a mirror image (an exact copy) of the physical system that follows the complete lifecycle of the product from idea to manufacturing, is possible only if (1) an exact representation of the system in mathematical equations is at hand and (2) the current state of all relevant elements of the real system can be fully monitored through real-time data. In bioprocess development, building a Digital Twin implies joining High Throughput, Omics, PAT, Machine Learning, Bioprocess Automation, and Bioprocess Systems Engineering tools to enable the development and operation of a biomanufacturing plant with a perfect copy of all units from the molecular/intracellular level up to large-scale dynamics. Such a Digital Twin is clearly far beyond the capabilities of current technologies. Still, it defines a clear roadmap that shows the relevance of the integration of different fields and tools to maximize the efficiency of bioprocess development. Mathematical models, which form the basis of digital twins, support all fields of biotechnology and bioprocess engineering [9]. This includes biochemical systems [10], systems biology [11], metabolic engineering [12], flux balance analysis [13], synthetic biology [14], and bioinformatics [15]. A good overview of applications of mathematical models, as well as a proof of their slow advance in bioprocess engineering is given by Jay Bailey [16]. Nevertheless, the complexity of biological systems poses difficult challenges to the direct use of advanced mathematical techniques in bioprocess development [17].

The complexity of the underlying metabolic and physiological phenomena demands large nonlinear equation systems with a large number of unknown and often time-variant parameters. The existing methods are too complex and computationally expensive for application in biotechnology [18–20]. Compared to general applications in engineering [21–24], biotechnological applications typically lack sufficient data, as well as process understanding [25–27].

Finally, the advances in artificial intelligence, especially in data-driven learning tools, offer incredible possibilities, but need to be adapted to the specific needs of bioprocesses, which have peculiarities (e.g., evolution of the biological system [28]), broad population distributions, very complex chemical composition and complicated (metabolic) reaction networks that are not present in mechanical or chemical processes [12]. Nevertheless, such mathematical tools have greatly contributed to our understanding of the interactions between the organism and the constraints of growth in bioreactors, as well as the elucidation of otherwise obscure intracellular processes [13].

The problem is, however, that these tools and concepts, namely scale-down bioreactors, high-throughput mini bioreactors, and model-based tools, have mainly developed in parallel, with little or no interaction among them. In fact, scale-down bioreactors are still operated as standalone, low-throughput devices [3]; and the benefits of mathematical models are not fully exploited in both scale-down and high-throughput systems [29]. Therefore, the actual challenge is to combine these very special techniques such that they can work together efficiently.

In this work, we discuss the current state of process development focusing on scale-down, the typically underestimated milestone. We discuss existing experimental tools, sensor technologies, and latest advances in computational methods for the design of scale-down investigations. Next, we demonstrate the issues related to the current decoupled efforts to address process development. Finally, we describe the required steps to reach a proper integration of all tools to create a digital twin of the bioprocess development procedure together with its potential and future applications.

## 2 The Digital Twin in Bioprocess Development

The answer to the current challenges in advanced bioprocess development (see Fig. 1) is a digital twin that covers all developmental stages and allows an efficient and effortless transfer of knowledge and information throughout the complete process [30]. Thus, the term "digital twin" covers more than just the mathematical model of a single component of the process. With regard to industrial bioprocesses, "digital twins" can describe the biological system itself or parts of the system, e.g. the three-dimensional structure of the protein product. They can also describe phases of process development, such as strain screening, different scales of the fermentation process and downstream operations, and in final production they can be used for the design or installation of a production plan, as well as for the control of the actual manufacturing process including its optimization.

The required advances in automation, process analytical technologies (PAT), and computer-aided tools for bioprocess monitoring and control are available [31]. The main challenge in building a functional digital twin is the difficulty in harmonizing these existing technologies through standardized communication protocols and data management systems. Such a digital framework tightly embedded into the highly automated experimental systems and production facilities through PAT and advanced mathematical modeling tools can build the path for knowledge transfer between the whole bioprocess development workflow.

Scale-up and scale-down present arguably the most descriptive examples for the challenges of knowledge transfer as well as its relevance in bioprocess development [32, 33]. Scaling is basically an effort to transfer the information generated in one stage to another aiming to maximize the generation of relevant knowledge for the industrial process [17].

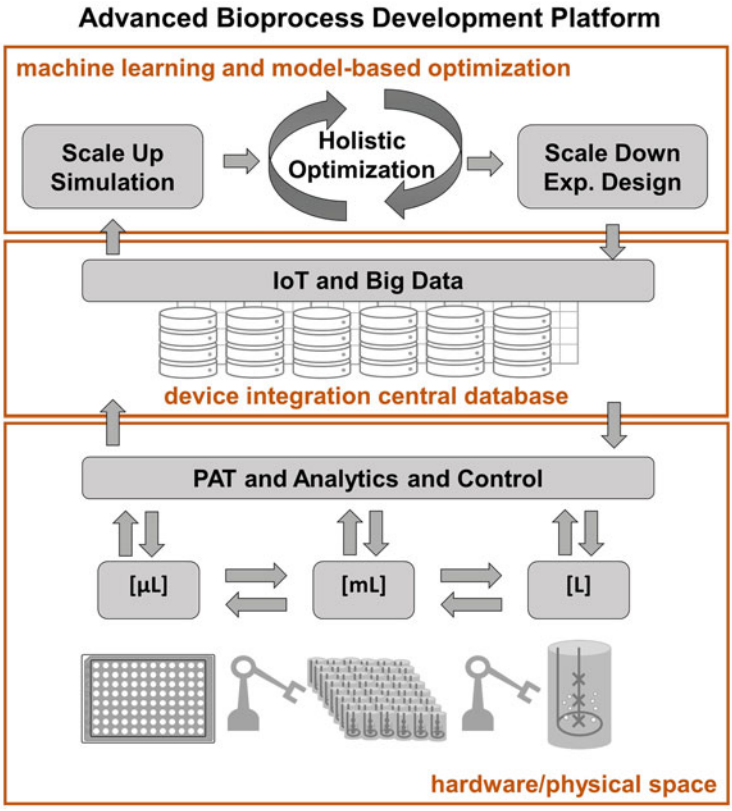



**Fig. 1** The role of the digital twin in advanced bioprocess development. Integration of scales, units, and disciplines. From bottom to top, automated hardware (the physical system), device integration, data transfer and handling, and model-based optimization tools for scale-up and scale-down

In biotechnology, scale-down is eminent to assure the generation of relevant-to-process knowledge in lab scale. As discussed further below, the variation to the cellular response caused by the bioreactor stresses cannot be predicted without extensive experimental information under proper conditions. That is, the experiments at lab scale must create the proper environment, to emulate industrial conditions which are unfortunately difficult to predict beforehand due to the highly complex interaction between the organism(s) and the bioreactor. Scale-up is the art of extracting knowledge from experimental data to translate it into an efficient manufacturing strategy. This process is as challenging as scale-down for the same reasons: insufficient data on the underlying dynamics of the bioprocess and the lack of a proper mathematical translation of the information throughout scales. In both cases (scale-up and scale-down), purely data-driven methods fail to understand the complex interactions between the organism(s) and the distinct bioreactor environment, generating unrealistic predictions for the next scales and corroborating the risks of extrapolating black-box models [34, 35]. On the other side, highly complex

mechanistic models offer important insights into the system [36, 37], but require process data that is very expensive, if at all obtainable.

There is no doubt that the high complexity of living organisms, especially in connection with their flexible information networks, which have been developed in millions of years of evolution, is the main challenge in modeling biological processes, and hence the reason why purely in silico development methodologies are doomed to fail [17]. Natural organisms have been so successful in surviving unfavorable conditions as they are characterized by genetic and metabolic heterogeneity between the individuals in a population. This increases especially under stress conditions in an effort to maximize the chance for some to survive [38]. Bioprocesses on an industrial scale (1) contain highly heterogeneous cell populations and (2) induce a constantly changing environment that sets the organisms under a high stress due to the inhomogeneities in the reactor. It is for this reason that advanced experimental facilities, sensor technologies, and mathematical modeling must be tightly integrated into a digital twin framework to finally achieve development times comparable to other industries [39–43]. The predictions at the digital level require a continuous re-calibration and evolving mathematical description to cope with the unpredictable behavior of living systems as well as efficient strategies to design and operate the experimental campaigns [29]. Most importantly, these models also reflect the dynamic development of the heterogeneity of the population and take statistical uncertainties into account in order to account for possible batch-to-batch variations on an industrial scale and be robust in terms of their prediction.

## 3 Inhomogeneities in Industrial-Scale Bioreactors and Their Influence on the Biological System

Bioprocess development is usually started in shaken cultures, traditionally in shake flasks, or more recently in parallel microwell plates where only a few endpoint measurements are possible [44]. Nowadays there is a great interest in implementing the final strategy, the fed-batch method, very early in the development process. For this purpose, various methods have been developed in the past years, which make this possible despite the small volumes [45].

Although this represents a significant advance on the traditional approach, the methodologies can vary greatly depending on the applied system. Substrate feeding can be either continuous or intermittent (pulse-based), and the controllers for the continuously measured parameters (e.g., pH value and dissolved oxygen) can be set differently, if these parameters are adjusted at all. If, as discussed above, we assume that the culture is highly sensitive to the process conditions – and thus the product formation is influenced accordingly – it is necessary to set these parameters so that the conditions are similar to those on an industrial scale.

This is important, especially for the development of biopharmaceuticals, where drug substance for clinical trials may be produced as early as the laboratory development phase, or in early pilot plant phase under cGMP guidelines. Once regulatory clearance is obtained, the process is then scaled up to industrial conditions. Two important points must be considered here: (1) the product quality characteristics affecting the efficacy of the biopharmaceutical candidate must comply with the specifications defined during clinical trials and (2) the process must be economically feasible whilst producing sufficient quantities to supply the market. Unfortunately, the increase in size has many implications for the process conditions inside the bioreactor. That is, if an 80 L bioreactor is scaled up to 10,000 L whilst maintaining a constant power input per unit volume, the mixing time increases 3 times, the impeller tip speed doubles, and the shear forces increase almost 10 times [46]. Oldshue showed that a scale-up design to satisfy mass transfer (constant $K_l a$ criterion) from a 75 L pilot scale process to a 95,000 L production scale would increase the shear rate by 180%, whereas maintaining a constant shear rate between the two scales could only produce 40% of the mass transfer requirements of the culture in the large scale [47]. The most common consequence is an inevitable increase in mixing times of up to 200 s in larger-scale bioreactors (since scale-up is mostly based on $K_l a$, P/V, impeller tip speed) [48].

In addition to the increased mixing times, fed-batch processes are fed with concentrated substrates at localized feeding points, which are mechanically fixed. The longer mixing times and the localized addition of highly concentrated viscous substrates lead to the formation of concentration gradients in the bioreactor [49, 50]. Cells that are traversing these gradients respond in many ways, by the varied distribution of metabolic fluxes due to the changed uptake rates in different positions of the bioreactor, and by specific gene expression profiles which include both specific responses and general stress adaptation. The specific reaction of an individual cell depends not only on its metabolic state and the current phase in the cell cycle, but also on its specific historical situation, i.e. what conditions it has experienced previously in the dynamic course of time [51]. This is currently being investigated using fluid dynamic models by simulating cell lifelines. The sum of all of these affects the fermentation efficiency in terms of yields and overall process robustness.

When the characteristic time of relevant cellular processes (translation, cell division) is close to the mixing time in large-scale bioreactors, there is a measurable influence of gradients on the growth and metabolic behavior of the culture [46, 52]. The inefficient mixing in large-scale bioreactors leads to the creation of spatial concentration pockets of relevant process parameters, such as substrate (glucose), dissolved oxygen, acidity, and temperature. Furthermore, GMP manufacturing processes suffer from the rigidity of the process due to the difficulty in using validated equipment for such studies, especially when the characterization study requires minor retrofitting of the bioreactor, such as installing extra sensors. In cases where bioreactor characterization has been done, companies consider the data as confidential; therefore, the information is not available to the scientific research community.

**Substrate Gradients** In fed-batch cultures, the existence of excess substrate zones in the broth defeats the purpose of this tight control for the fraction of the culture that comes into contact with these zones. The exposure of the culture to zones of higher substrate concentrations has direct consequences on the uptake capacities of the cells for this substrate [51, 53]. As a result, the excess substrate zones may cause the cells to grow at the maximum specific growth rate, which may plunge organisms such as *E. coli* and *Saccharomyces cerevisiae* into overflow metabolic states as reported in numerous studies [51, 54–56]. The high metabolic flux of glucose through the glycolysis which is favored by high affinity uptake systems, i.e. low $K_S$ values, also leads to the accumulation of NADH-$H^+$, and thus to a higher rate of respiration.

As a consequence, the high metabolic activity in the feeding zone can also lead to oxygen limitation if the biochemical reduction of oxygen by the cells is faster than the limited diffusion of oxygen into the cultivation medium. It is likely that the uneven distribution of the substrate due to feeding is the main cause for the *dissolved oxygen gradients*, besides the uneven fluid-dynamic distribution of the gas bubbles. The dissolved oxygen problem which is basically caused by the inherently low solubility of oxygen in fermentation broths [49] becomes even greater in processes with pellet forming organisms (oxygen gradient in the pellet) or shear-sensitive cells (limited sparging to prevent shear stress caused by the bursting of gas bubbles) [57].

**Temperature Gradients** Temperature gradients are among the least studied scale-up effects in bioprocess development. Although it is clear from a microbiological point of view that small temperature fluctuations of a few degrees have a major impact on cellular reactions and that, from a process engineering perspective, precise temperature control in industrial bioreactors is a serious problem, to the best of our knowledge, there is no information about local temperature profiles in industrial bioreactors, nor have experiments been performed in scale-down simulators to simulate the effect of perturbing temperatures on a process.

**pH Gradients** pH gradients are recently gaining attention in the bioprocess research community. Simen et al. investigated the effect of ammonia pulses (shifts in pH) in *E. coli* cultivations and observed a higher maintenance energy and the activation of over 400 genes in response to the pH gradients [58]. pH gradients are also relevant in industrial-scale batch cultivations of lactic acid bacteria. This has been revealed by combined approach by the use of multiple pH probes and a computational fluid dynamic model coupled with a kinetic model for a process of *Streptococcus thermophilus* in a 700 L pilot scale bioreactor [59]. Recently, we also could demonstrate by two- and three-compartment bioreactor cultivations that such pH oscillations affect the cocci chain length and decrease the growth rate in *S. thermophilus* cultures (manuscript in preparation). Also in CHO fed-batch bioprocesses pH perturbations decrease the cell viability and increase lactate accumulation [60]. Also pH oscillations have been recently demonstrated to affect product accumulation in a cell line specific manner [61].

**Carbon Dioxide Gradients** In microbial cultivations, a recent study of $CO_2$/$HCO_3^-$ gradients in *Corynebacterium glutamicum* showed no significant impact

of these stresses in the physiological response of the organism, although there was a marked increase in the expression of certain genes, upon genomic analysis [62]. In a recent report, *E. coli* cells exposed to $CO_2$ levels above 70 mbar $CO_2$ partial pressure in the inlet gas led to reduced biomass yields and rapid accumulation of acetate, even under non-overflow and fully aerobic conditions [63].

**Interaction Between Multiple Gradients** Finally, the results of Limberg and colleagues show that when pH gradients are coupled to oxygen limitation, *C. glutamicum* loses its robustness against dissolved oxygen fluctuations [64], leading to yield losses of up to 40%. This implies that the study of concentration gradients in fermentation should be conducted in a multi-faceted manner, to consider all possible gradients and the necessary combinations among them to arrive at a more holistic conclusion for each strain. There is also a close correlation between $pCO_2$ levels, pH, base addition, and osmolality in large-scale CHO cell cultures which affect the metabolic lactate shift (transition from lactate production to lactate consumption) [65, 66].

## 4 Framework for Bioprocess Scale-Down Studies

### *4.1 Characterization of the Large Scale*

A good characterization of the large-scale bioprocess is important to conclude proper scale-down experiments which really imitate the large scale (see Fig. 2). Since the scale-down data is only as good as the environment it mimicked, it is absolutely necessary to characterize both the cellular state and specific heterogeneity (gradient profiles) in the larger scale. Standard analytical methods of the medium and gas composition and the derivation of cell specific rates need to be complemented by direct monitoring of the physiological state of the cells. A proper scale-down methodology should be based on the similarity of cellular responses, all at the level of metabolism, protein expression, and population heterogeneity between laboratory and industrial scale. In order to avoid false conclusions and to reduce the risk of scale-up, robustness analyses must be used to assess the final batch-to-batch variability. This complex problem can only be solved if digital approaches (digital twin) can be coupled with a large number of experiments.

In the past, there were a large number of approaches to simulate these gradients occurring in the industrial bioprocess in scale-down systems, see reviews by Neubauer and Junne [3], Lara et al. [46], Delvigne and Noorman [67]. All these systems achieve oscillating conditions regarding the specifically investigated parameters, i.e. the specific parameters which were the focus of the investigators. Different priorities were set depending on the specific approach. In multi-compartment reactors, the dominant parameter is the residence time distribution in different compartments where cells are located within a defined period of time. In more-compartment stirred tank systems, the zones are characterized by a previously defined state, e.g. in

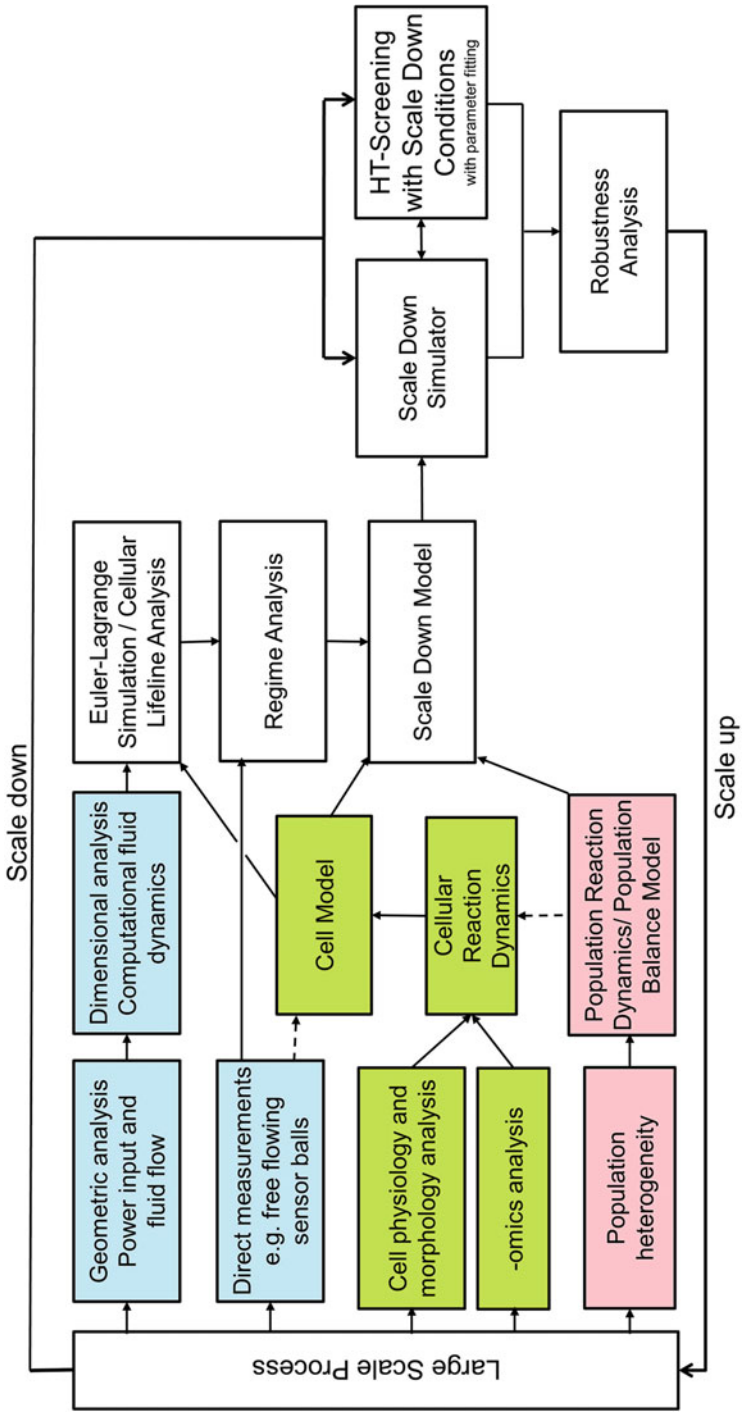



**Fig. 2** Schematic outline of state-of-the-art methods for bioprocess scale-down. Scale-down starts with an analysis of the industrial large-scale bioprocess. While current approaches dominate in fluid dynamic characterization, the generation of a reliable cell model (digital twin) and the characterization of the cellular reaction dynamics, i.e. parametrization of the model, are most important for scale-down experiments. The three pillars of bioprocess characterization are highlighted in colors: reactor (blue), cell (green), populations (red)

which a defined pH value is set in each of the reactors [68]. When using plug-flow reactors, a gradient is established within the system. Then, the sampling in different positions along the plug-flow reactor also allows an insight into the time course of the cellular reactions [69, 70]. In contrast to multi-compartment systems, scale-down simulators with a pulse-based feeding are easier to establish and to run in parallel. Parameter control, however, may be more difficult to achieve due to the restriction of the feeding profiles. Since the feeding profile is easy to change (e.g., distance between the feed pulses), pulse-based systems also seem to be well applicable for robustness analyses. Alternative approaches, in which installations (e.g., plates between the different stirrers to restrict the tangential flow [71]) are realized in a laboratory reactor to extend the mixing time to the order as it is measured in the large reactor, can, in individual cases, reproduce the industrial process quite well, but are technically more complex.

### 4.1.1 Monitoring of the Cellular State Across Different Scales

The most successful scale-down methodology will maintain the physiological state of cells across lab and industrial scales. Naturally, it is the most suitable pre-requisite to obtain similar results, and should be considered as scaling parameter, although the examination of the physiological cell status is not easy to quantify with suitable measures. The impact of gradient formation on physiology has to be investigated with the measurement of sensitive parameters, e.g. the energy charge, stress response factors, and the respiratory activity, among others [68, 72–74]. Additionally, the physiological state may vary from cell to cell, which demands the consideration of population heterogeneity. It has been observed several times that gradient formation in fed-batch cultivation mode has an impact on population heterogeneity [75]. It adds additional parameters that lead to different phenotypes in culture (Fig. 3).

In natural habitats, mainly the cell cycle, cell ageing, and epigenetic regulation are known to have a great impact on the evolvement of phenotype diversity [77]. Stochastically asymmetric growth and mutation events drive the formation of subpopulations, which might be even better adapted to a previous change in an environment. Nevertheless, these events usually lead to lower yields in processes, which are conducted in bioreactors [78]. The role of cell cycling on the development of subpopulations in industrial bioprocesses, however, is not clear yet, while it was found out that the dominant driver for different protein concentrations, and thus various metabolic activity, is the growth rate in *Pseudomonas putida* [79].

In particular, singe-cell based and sensitive volumetric measurement techniques can provide new information about the impact of gradients on the cellular viability and metabolic activity and the formation of subpopulations independently of the scale-down system. Suitable monitoring technologies in combination with a physiological understanding of stress responses support the identification of the suitable scale-down conditions, as it puts the cell in the center of the investigation of consequences of gradient formation in the liquid phase. Such technologies, including proper accompanying off-line measurements, allow one to properly model the stress

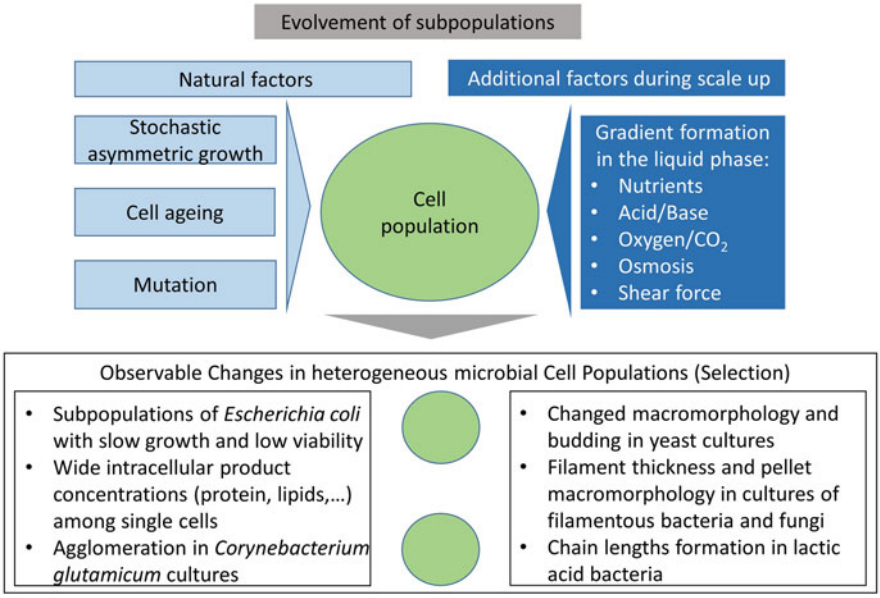



**Fig. 3** Parameters that are putative effectors on the formation of subpopulations. Examples for microbial cultivations are own observations, further described in [76]

response and provide a basis for systems biology interpretation, which deepens the methodological understanding of cellular responses in the large-scale environment.

So far, investigations have shown that the exposure of cells to gradients leads to a higher population heterogeneity under scale-down conditions. This was examined in particular for protein concentrations in *E. coli* [80]. It was found that the dynamics of glycolysis might play an important role in the development of non-growing subpopulations [81]. One way to observe this evolution of subpopulations with sufficient accuracy and time-resolution is the creation of a strain with a reporter protein that can be quantified by fluorescence, which enables the application of fluorescence-assisted characterization of single cells, and eventually cell sorting [38, 76]. For example, the green fluorescence protein can be used, if coupled to automated sampling, eventually coupled with a multiplexer, and spectroscopic methods like flow cytometry, for a statistically proven detection of subpopulations [82]. In *P. putida* cultivations, the change in DNA content in individual cells was investigated under different environmental growth conditions with flow cytometric analysis at various dilution rates in chemostat experiments. The impact of oxygen deprivation, solvent exposure, and iron availability on DNA replication was also investigated [83]. The application of flow cytometry and cell staining to characterize population subgroups was described in several other studies as well [76, 84]. Nevertheless, this is a challenging technique to apply in bacteria due to their size. In bacteria, the quantification based on fluorescence is subject to genetic noise [75], which in this case might not be predominantly affected by large-scale cultivation

conditions, but constantly present across scales. The results, however, would be biased by the noise, if the samples measured are not enough to ensure a statistically valid distribution. Nevertheless, the methodology is suitable to understand the effects of gradients on cell viability and vitality while a large amount of cells can be examined in a considerably short time. While the possibility of the application of fluorescent markers represents an approach mostly on the expression level, cell staining can be used to investigate several cellular components, including the quantification of metabolite concentrations. If accumulated intracellularly, metabolite concentrations can be used to assess population heterogeneity as well, if the metabolite can be quantified with sufficient accuracy in individual cells. Flow cytometry is able to quantify the accumulation of intracellular lipids in microbial cells, and thus identify subpopulations of different lipid contents [85], e.g. with Nile Red or Bodipy® stains. The accumulation of lipids and also other components might correlate with changes in macromorphology of organisms. The measurement of lipid content with optical methods can lead to conclusions about metabolic activity of individual cells. This has been shown for heterotrophic algae, which accumulate to large extent polyunsaturated fatty acids in lipid droplets. While using light microscopy and 3-dimensional holographic microscopy, the individual lipid storage in cells was measured based on their individual cell size [86]. With rapid image analysis using trained software, image acquisition can be performed in flow cells that are connected to a cultivation. Automated workflows that offer considerably fast analysis of populations similar to flow cytometry are feasible, without the requirement of staining.

Besides intracellular product accumulation, the macromorphology can provide suitable information about the cell status and the impact of gradient formation on it. It was examined that the cell size of *S. cerevisiae* cultures changed with the degree of environmental heterogeneity in a three-compartment scale-down reactor [87]. This happened in parallel to growth reduction and side metabolite accumulation with a concomitant change of the sterol content, in comparison with homogeneous growth conditions. Cells showed a diverse macromorphology under scale-down cultivations, which supports the hypothesis that population heterogeneity is rather increased under growth in gradients. A morphologic response of cells to scale-down cultivation conditions can also lead to agglomeration due to stress response. Although the macromorphology of individual cells may stay unchanged, the secretion of side products or proteins supports the agglomeration of cells. Observations with laser-light back-reflection for cell particle size measurement indicated the formation of clumps of *C. glutamicum*, when exposed to oscillatory oxygen supply, either in a scale-down reactor concept (three-compartment reactor) or in shake flask cultures with interrupted shaking [88]. Agglomeration under fluctuating oxygen availability was postulated to be a result of increased secretion of biofilm forming metabolites, e.g. in *Mycobacterium tuberculosis* and *E. coli* cultures in connection with oxidative stress response [89, 90].

In case of filamentous organisms, macromorphological similarity across scales is often achieved only if the shear force regime is maintained. Mechanical shear forces as they appear close to the stirrer can lead to filament disruption, with consequences

on growth and secretion profiles. Up to now, however, the change of macromorphology due to scaling effects, e.g. an oscillating shear force regime, has not been investigated thoroughly. Nevertheless, it can be assumed that the macromorphology of filamentous organisms will change in comparison with the lab scale if the exposure time to high-shear forces is diminished, like it most probably is in large-scale cultivations at high cell densities and elevated viscosities, where large residence times exist in different compartments. The knowledge that exists so far about how a changed shear force regime influences the process performance [91, 92] leads to the assumption that macromorphology is an important parameter to consider while choosing a suitable scale-down system. Alternating shear forces can be achieved by interrupted stirring, which usually couples low-shear stress to oxygen limitation in stirred tank reactors, or in multi-compartment reactors, in which low-shear and high-shear regimes are applied at similar gas mass transfer rates. The application of other reactor systems beyond stirred tanks can support the investigation of consequences of low-shear forces on the macromorphology, physiology, and overall process performance as recently described for clavulanic acid production with *Streptomyces clavuligerus* in shaken bioreactors. Secretion of clavulanic acid was strongly diminished while thicker filaments were observed [93]. Consequently, morphological monitoring in an automated manner [94] is a promising technique to identify crucial characteristics for growth and product formation under specific environmental regimes.

Finally, macromorphological heterogeneity can be modeled to describe the response of a cell to environmental perturbations. In silico prediction of physiological population heterogeneity was conducted by a combination of computational fluid dynamics (CFD) and a cell cycle model of *P. putida* [95]. It was observed that 72% of the cells switched between standard and multifork replication and 52.9% showed higher than average adenosine triphosphate (ATP) maintenance demands (12.2%, up to 1.5 fold). Such an approach, however, requires sufficient knowledge of the interaction between gradient formation and consequences for the macromorphology of a population. This still represents a bottleneck as the time frequency with which morphological changes are measured might be inadequate to achieve a sufficient accuracy while correlating the response to specific regulatory events in a cell. In recent years, however, many more techniques like in situ microscopy and others are being developed rapidly. It is hoped that they become more applicable in biotechnological processes operating at elevated cell densities. The impact of gradients in the liquid phase on the formation of phenotypic heterogeneity can be investigated also if spectroscopic methods are coupled to microfluidic devices, when the growth of single cells can be monitored constantly [96]. The aforementioned methodologies will increase the possibility for the consideration of heterogeneity in population balance models and their integration in the description of consequences of gradient formation. So far, the few attempts rely on physiological measures, e.g. the adaptation to substrate excess [97, 98], but investigations will benefit from the additional consideration of macromorphological characterization data in future.

### 4.1.2 Combination of CFD Approaches with Mechanistic Models (Euler–Lagrange) to Describe the Large Scale

Computational fluid dynamics has been traditionally used to describe the flow under defined conditions, as a way of characterizing the heterogeneous conditions in large-scale bioreactors. With the increasing computational power it is possible to implement in such CFD approaches cellular reaction kinetics. For the first time this interaction between the intracellular state of the individual cells of the population and the turbulent flow field in the bioreactor has been realized by Lapin et al. [99]. This so-called Euler–Euler approach considers gas, liquid, and biophase as a continuum and is an answer to the very complex simulation, which is dependent on the definition and resolution of the reactor into flow cells. Later, the pioneering paper by Lapin et al. [100] was the first approach to couple a CFD model of a bioreactor with a Lagrangian approach for the combined solution of flow patterns and cellular kinetics.

In this Lagrangian–Euler approach the liquid phase is treated as a continuum (Euler) and the dispersed phase is tracked using Lagrangian representation. While this modeling approach was first used in gas–liquid simulations, here cells with their specific metabolic reaction network were described as discrete entities. With these models, individual cells are monitored with respect to their experience of local environments in relation to the fluid-dynamic distribution and pathways within the bioreactor. This kind of structured-segregated approach realizes that the individual history of a biological entity determines their reaction. With a big computational power this can be realized also in a three-dimensional turbulent field. This approach allows one to get an indication of the heterogeneity in the biotic and abiotic phases of the reactor and it considers the individual history of the cells as important for its final response. By tracking the pathway of a single particle over time it is possible to derive lifelines of a big number of cells and thus draw conclusions for regimes, i.e. conditions which should be represented in scale-down experiments.

In summary, the use of mathematical descriptions of large-scale bioreactors by the combination of Euler–Lagrangian approaches is very illustrative and has made major progress during the last years.

However, there are some limitations:

1. Due to calculation expense it is not possible to consider realistic amounts of cells. Current computational approaches consider approx. 100,000 cells, which is enough to see and follow the population dynamics. Currently this number of cells is fixed. However, it would be interesting to consider growth and an increase in cell number over time.
2. The cellular models and the parameters used in these studies are mostly derived from continuous experiments (mostly chemostats), i.e. from experimental conditions which do not reflect the large scale. Thus, as discussed above, the reactions in a real reactor may be totally different. Therefore urgent approaches and methods which describe how a cellular model can be derived and parameterized

are needed. In our opinion, there is a need to look for alternative approaches which are easily applicable and better reflect the large-scale situation.

3. Due to the very high computing power, which is needed to solve these Euler–Lagrangian models with a reasonable resolution, the current models cannot describe the process dynamics over time, but only represent a very narrow time point of the cultivation. Nevertheless, this kind of simulations, e.g. if they are performed at different time points of a cultivation can provide important information to plan scale-down experiments.
4. Current models only consider the liquid phase. The implementation of the gas phase would additionally need much computing power and in our opinion it would be very laborious to validate these models.
5. An important characteristic of living systems in connection with their adaptation to the environment is the heterogeneity in a population. Physiological (i.e., metabolic) and genetic heterogeneity ensure the survival of a population of cells if environmental changes occur (stress phenomenon) and has been described historically as the survival of the fittest. As growth in a bioreactor is related to different stresses in different phases of a process and as additional perturbations that occur in large-scale bioreactors add a further stress layer, the population heterogeneity in a bioprocess is an important feature, which needs to be monitored and can be used for a validation of similarity between different process scales.

As a consequence, every scale-down approach needs to start with a good understanding of the large scale especially in view of the cellular response dynamics. Furthermore, these cellular response dynamics must be reduced in mathematical models, the so-called digital twins, to the characteristics essential for the process scale-down. Finally, methods need to be implemented which help to validate the quality of the scale-down – and this is only possible by measurements.

## *4.2 Execution of Scale-Down Experiments*

### 4.2.1 Combination of Scale-Down Experiments with Model-Based Approaches

As shown in the scale-down scheme in Fig. 1, the characterization of the larger-scale bioreactor environment is followed by transferring the environmental blueprint to the laboratory-scale simulator in which the actual scale-down experiments are executed. To achieve this transfer, the digital twin of the bioprocess should contain model units that adequately describe the physiology of the cells, as well as the cultivation process and geometric analysis of the bioreactors involved. These details should be digitally embedded into the definition of the scale-down model (Fig. 1). Thereafter, the application of the modeling framework (digital twin) in the context of scale-down experimentation can take two forms: (1) the design of the scale-down

experiments and (2) the interpretation of the data from scale-down experiments. These two branches of application are presented in the following sections, with respect to the state of the art, as reported in current literature and future perspectives.

## Model-Based Design of Scale-Down Experiments

The concept of scale-down experiments developed in the past 3 decades usually involves creating one type of stress (e.g., dissolved oxygen limitation, excess substrate, excess metabolite concentrations, and perturbations in pH) in the scale-down simulator. However, considering an actual larger-scale bioreactor, concentration gradients arise from mixing effects in a 3-dimensional space, combined with the uptake of substrates and release of metabolites by cells. Moreover, the type of gradients are always coupled and may co-exist (e.g., pH-oxygen-substrate gradients) in the larger bioreactor. Therefore, at best, the scale-down simulators are only a gross estimate of the actual environments in large bioreactors. Additionally, the fraction of cells exposed to a given gradient in a large-scale bioreactor has been variable in the definition of the scale-down model. In multi-compartment scale-down simulators, this has been in the range of 10%–30%, whereas in pulse-based single-compartment simulators, the total population is subjected to the stresses, without population subgroups. Notwithstanding these challenges, important physiological responses have been reported by researchers using these physical approximations of larger bioreactors. In this light, experimental set-ups for scale-down studies can improve greatly when they are combined with the ideas of digitalization in the industry 4.0 era.

The scale-down model, defined on the blueprint of the environmental heterogeneity of the larger scale, will contain process specifications such as the gradient profiles, zone definitions with defined boundaries, residence time distributions, magnitudes of gradients, frequencies and other experimental inputs that should be implemented in the simulator. With such a scale-down model in hand, the scale-down experiment can be designed on a computer, and results of the design sent to intelligent equipment (pumps, pipettes, agitators, liquid handlers, etc.) to control a small-scale bioreactor to mimic the blueprint of the larger-scale bioreactor. That is, mixed-gradient zones, containing excess substrate with acidic pH, or limited oxygen with high hydrodynamic stress, or any desired combination of gradients can be created with the model, and the size of the zone and its frequency/duration can be varied randomly to more closely resemble actual gradient dynamics in the larger scale. A step in this direction is the work of Anane et al. [29] who used a mechanistic model of *E. coli* [101], in combination with a mechanistic description of the gradient profiles of a multi-compartment scale-down bioreactor [1] to calculate glucose pulses in high-throughput scale-down experiments. The outputs of the two models were integrated into the operation scheme of the high-throughput system to reproduce heterogenous glucose conditions in minibioreactors. The calculated glucose pulses were commensurate with the physiology of the strain, as the pulse sizes were derived from uptake capacities and physiological limits of the strain, as well as

mixing effects in the scale-down bioreactor. The authors reported significant yield losses, incorporation of non-conventional amino acids into the recombinant protein product and accumulation of metabolites in response to the calculated gradients.

Although these results were similar to observations in other non-model-based scale-down approaches [69, 102, 103], the added advantage of using the model was the flexibility of stress definition, in which one parallel experimental set-up was used to implement six different stress zones in the scale-down system, which otherwise could not be done manually.

A further advancement of this concept is the application of models to design intelligent scale-down experiments. In all the previously discussed scale-down methods, the researcher sets out a prior design space, within which the stresses and gradient profiles are pre-defined (manually or using a model) and executed during the experiment. The use of digital twins to advance scale-down design should involve designing the stresses as the experiment runs. In other words, the nature of gradients to be imposed on the culture at time point t2 will depend on its state at time point t1. This will prevent overestimating or underestimating the magnitude of gradients, as well as the exposure time of the cells to a given gradient. Such an adaptive re-design technique was employed by Cruz Bournazou and colleges in designing optimal experiments to maximize information content for model parameter identification [40, 104]. Although no scale-down efforts were made in these works, the authors demonstrated the ability to re-define feed regimes based on the current state of the culture and model predictions for a given time window. Such a model-based adaptive system can easily be employed to execute dynamic scale-down experiments.

The addition of digital twin concepts to the definition of the scale-down model offers flexibility of stress definition, automation, and a high turnover in experimental throughput, to drive the digital revolution in bioprocess engineering, as discussed by Neubauer et al. [30]. The few pioneering works published so far point in the future direction where mathematical methods, in the form of digital twins, will help to design more informative and smart (scale-down) experiments, to move away from the traditional, commonly used static design of experiment (DoE) paradigm.

### Model-Based Interpretation of Scale-Down Data

Scale-down bioreactors offer important insights into cellular behavior under heterogeneous fermentation conditions of larger-scale bioreactors. The data is usually interpreted at the macroscopic level, by comparing metabolite, substrate, and growth profiles to cultivations under homogeneous conditions. In a few studies, derivative indices (e.g., specific uptake rates, yield coefficients) have been calculated from the raw scale-down data to support the interpretation of the data, e.g. in [105]. A few extensions of the data space in scale-down experiments involve molecular level analysis. For instance, Simen and co-workers used transcriptomics data from an STR-PFR scale-down bioreactor to monitor different gene expression levels under short-term and long-term substrate fluctuations in *E. coli* culture [58]. In another

study, quantitative metabolomics was used to monitor the consumption of amino acids in scale-down cultivations of *Bacillus megaterium* expressing green fluorescent protein. The metabolomics results from the scale-down conditions were then used to design a better feed composition for the process [106].

It is important to note that a strain's response to heterogenous environments in a scale-down bioreactor is the total sum of its molecular level responses. What if a particular response characteristic, such as accumulation of non-conventional amino acids or de-activation of acetate cycling (accumulation of acetate) in *E. coli* could be traced to particular metabolic pathways and mechanisms? What if data from a scale-down bioreactor could be used to trace specific metabolic fluxes of a clone as it is exposed to various concentration gradients? Such information would be useful, not only in strain engineering, but also in designing efficient processes at industrial scale. Advanced modeling of scale-down data, i.e. fitting mechanistic and dynamic metabolic models to data from scale-down cultivations can reveal specific pathways that are active under given heterogeneous conditions. This is offcourse assuming that the parameterization of the model includes such metabolic and physiological indices. The flux terms to be fitted to the scale-down data should be an integral part of the building of the cell model (Fig. 1). Again, Anane et al. [29] fitted data from parallel scale-down cultivations of *E. coli* under multiple glucose gradient conditions to a mechanistic model describing the process and the strain. The authors found that the different responses of the strain to the gradients translated directly into different values of the parameters of the model. Therefore, the model expanded on the primary data of the scale-down experiment, and expanded the interpretation of the available data for better process and strain design. In a similar study, Janakiraman and co-workers used multi-variate data analysis techniques to interpret scale-down data [107]. Their aim was to establish comparability between scale-down cultivations in Ambr15® minibioreactors and cultivations in 15,000 L manufacturing scale, by applying principal component analysis to the scale-down datasets. By employing this model-based approach, they were able to clearly identify that the runs in both scales were statistically similar to each other, a conclusion that would have been difficult to draw by looking at the raw scale-down data.

For digital twins to be applicable in this sense, there are a few pre-requisites the model of the bioprocess must fulfill: (1) the parameter estimates in the cell model must be subjected to rigorous validity and uncertainty tests, as presented by Anane et al. [18]. The reported parameter values should always be accompanied by confidence intervals at valid significance levels, to be able to derive biological meaning from the model results. (2) The model should be just as detailed as is necessary for its application. As pointed out by Gábor and Banga [108], the parsimony principle should always be applied in building the model: i.e., the number of parameters should not be more than those required to describe the process in its simplest form [108, 109], and (3) the model should be constantly updated to include the most recent research findings in cell physiology and metabolism. The physiological accuracy of the model should be ascertained by subject matter experts in the field, which may not necessarily be the modeler (mathematician).

### 4.2.2 High-Throughput Execution of Scale-Down Experiments in Parallel Cultivation Systems

High-throughput experiments in parallel cultivation platforms have become common in bioprocess development laboratories. In the past decade, there has been an exponential increase in the adoption of these systems for early bioprocess development [45, 110, 111]. At the same time, due to Quality-by-Design (QbD) guidelines, there has been an increasing demand to fully characterize bioprocesses at the development phase, to forestall unforeseen consequences of the final process, upon scale-up [112, 113]. This requirement demands that all conditions, including actual large-scale process conditions are considered and tested in the early development phases of the process. Therefore, the question of whether cultivations in minibioreactors are adaptable to mimic concentration gradients and the heterogeneous environments that exist in large-scale bioreactors has become very important, and should be addressed.

A few studies conducted in high-throughput cultivation systems that consider the heterogeneous conditions of larger bioreactors are reported in the literature. As described above, Janakiraman et al. [107] matched the volumetric aeration rates (vvm) between parallel Ambr15® cultivations of CHO cells and a 15,000 L production-scale bioreactor. They used this criterion to mimic the carbon dioxide profile of the production bioreactor in the minibioreactor cultivations, which led to similar productivity and product quality profiles in both the 15 ml bioreactors and the 15,000 L scale. In another study, Velez-Suberbie et al. [114] used the power per unit volume (P/V) as a scale-down criterion to compare Ambr15 cultivations of *E. coli* with 20 L bioreactor cultivations [114]. Perhaps the most comprehensive work in this regard was reported by Anane et al. [29], who used model-calculated glucose pulses to induce both dissolved oxygen and glucose gradient zones in 15 ml parallel minibioreactors. A key aspect of their work was the use of robotic liquid handling stations and mechanistic models in the operation of the scale-down set-up. These smart equipment were interphased with the minibioreactors, such that model outputs describing specific gradient conditions could be implemented in selected minibioreactors by the robotic system. Their results in *E. coli* fermentation development showed significant accumulation of non-conventional amino acids in the recombinant protein product, as well as accumulation of acetate in the scale-down cultivations, when compared to cultivations under homogeneous conditions.

The results of scale-down cultivations as performed in high-throughput minibioreactor systems so far show that it is possible to mimic large-scale environmental conditions in miniaturized bioreactors. Particularly, the physiological responses of both *E. coli* and CHO cells to the induced heterogeneous conditions in minibioreactors, as discussed above, is a proof of concept that gradient profiles that are relevant in industrial-scale cultivations can be reproduced in milliliter scale for scale-down studies. However, the adoption of enabling technological methods, such as robotic liquid handling stations and mechanistic modeling is fundamental for the successful operation of such minibioreactor facilities as scale-down platforms.

The adoption of such parallel cultivation systems and their combination with robotic liquid handling stations will ensure that a large number of gradient profiles, defined in the scale-down model, can be tested in a single parallel run. Additionally, such high-throughput systems can be used for strain screening under conditions that are amenable to the larger scale, to select the most robust strain for further development and scale-up.

## 5 General Conclusions and Perspectives

The lead times of biotechnological products, especially biopharmaceuticals, from discovery to market, can be up to 15 years [30]. Although other issues such as clinical trials may contribute to this time, bioprocess development and troubleshooting scale-up problems are key contributors to the lengthy lead times. The use of parallel cultivation systems and robotics has, undoubtedly, reduced these process development times significantly [45, 110]. Prior to screening, the development of strains is nowadays performed in a high-throughput manner, e.g. with the use of standardized genetic methods [115] and non-targeted high-throughput strain engineering [116]. Thus, the bottleneck of a faster overall bioprocess development is shifted from strain engineering to screening and cultivation development. The use of parallelized minibioreactor systems for both screening and upstream process development, as demonstrated in different studies [29, 107, 114], will greatly relieve this bottleneck, and ensure that a potential bioprocess reaches production within the earliest possible times. Additionally, the framework of screening under scale-down conditions and the associated methods will not only facilitate rapid bioprocess development, but also ensure a consistent and efficient cultivation process development by taking into account all the possible cultivation conditions that would be encountered upon process scale-up.

Digital twins have become an integral part of bioprocess development and process control. Particularly, the high degree of parallelization and automation of the development process, the integration of PAT and the requirements for a higher robustness of the processes in connection with an improved process control could only be realized through the comprehensive implementation of mathematical and statistical methods. Thus, the current challenge lies especially in the fusion of the individual tools into a uniform overall system.

The application of new possibilities that arose from the ongoing development of sensor technology and the corresponding data processing allows a stronger consideration of cell-to-cell variation and cellular features as scaling parameters. Such technologies, including proper accompanying off-line measurements, allow one to properly model stress responses and provide a basis for the integration of systems biology knowledge to deepen the methodological understanding of cellular responses in a large-scale environment. This can support the identification of suitable scale-down systems with the cell status as scaling factor as it represents the central

location for the product synthesis. It fosters the application of population balances and their integration into model-based descriptions of scale-up effects.

## References


1. Anane E, Sawatzki A, Neubauer P, Cruz Bournazou MN (2019) Modelling concentration gradients in fed-batch cultivations of *E. coli* – towards the flexible design of scale-down experiments. J Chem Technol Biotechnol 94:516–526
2. Neubauer P, Cruz N, Glauche F, Junne S, Knepper A, Raven M (2013) Consistent development of bioprocesses from microliter cultures to the industrial scale. Eng Life Sci 13:224–238
3. Neubauer P, Junne S (2016) Scale-up and scale-down methodologies for bioreactors. In: Mandenius CF (ed) Bioreactors: design, operation and novel applications. Wiley-VCH Verlag GmbH, Weinheim, pp 323–354
4. Reitz C, Fan Q, Neubauer P (2018) Synthesis of non-canonical branched-chain amino acids in *Escherichia coli* and approaches to avoid their incorporation into recombinant proteins. Curr Opin Biotechnol 53:248–253
5. Wang G, Haringa C, Tang W, Noorman H, Chu J, Zhuang Y, Zhang S (2020) Coupled metabolic-hydrodynamic modeling enabling rational scale-up of industrial bioprocesses. Biotechnol Bioeng 117:844–867
6. Grieves M, Vickers J (2016) Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Transdisciplinary perspectives on complex systems: new findings and approaches, pp 85–113
7. Grossmann I (2005) Enterprise-wide optimization: a new frontier in process systems engineering. AIChE J:1846–1857
8. Qi Q, Tao F (2018) Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. IEEE Access 6:3585–3593
9. Batstone DJ, Keller J, Angelidaki I, Kalyuzhnyi SV, Pavlostathis SG, Rozzi A, Sanders WTM, Siegrist H, Vavilin VA (2002) The IWA anaerobic digestion model no 1(ADM 1). Water Sci Technol 45:65–73
10. Tsugawa H (2018) Advances in computational metabolomics and databases deepen the understanding of metabolisms. Curr Opin Biotechnol 54:10
11. Kitano H (2002) Computational systems biology. Nature 420:206–210
12. Stephanopoulos GN, Aristidou AA, Nielsen J (1998) Review of cellular metabolism. In: Metabolic engineering. Academic Press, San Diego, pp 21–79
13. Varma A, Palsson BO (1994) Metabolic flux balancing: basic concepts, scientific and practical use. Bio/Technology 12:994–998
14. Marchisio MA, Stelling J (2009) Computational design tools for synthetic biology. Curr Opin Biotechnol 20:479–485
15. Saeys Y, Inza I, Larrañaga P (2007) A review of feature selection techniques in bioinformatics. Bioinformatics 23:2507–2517
16. Bailey JE (1998) Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. Biotechnol Prog 14:8–20
17. Koutinas M, Kiparissides A, Pistikopoulos EN, Mantalaris A (2012) Bioprocess systems engineering: transferring traditional process engineering principles to industrial biotechnology. Comput Struct Biotechnol J 3:e201210022
18. Anane E, López CDC, Barz T, Sin G, Gernaey KV, Neubauer P, Cruz Bournazou MN (2019) Output uncertainty of dynamic growth models: effect of uncertain parameter estimates on model reliability. Biochem Eng J 150:107247

19. Muñoz-Tamayo R, Puillet L, Daniel JB, Sauvant D, Martin O, Taghipoor M, Blavy P (2018) Review: to be or not to be an identifiable model. Is this a relevant question in animal science modelling? Animal 12:701–712
20. Villaverde AF, Barreiro A, Papachristodoulou A (2016) Structural identifiability of dynamic systems biology models. PLoS Comput Biol 12:1–22
21. Brubaker TA (1979) Nonlinear parameter estimation. Anal Chem 51:1385A
22. Brun R, Kühni M, Siegrist H, Gujer W, Reichert P (2002) Practical identifiability of ASM2d parameters – systematic selection and tuning of parameter subsets. Water Res 36:4113–4127
23. Kravaris C, Hahn J, Chu Y (2013) Advances and selected recent developments in state and parameter estimation. Comput Chem Eng 51:111–123
24. Vajda S, Rabitz H, Walter E, Lecourtier Y (1989) Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models. Chem Eng Commun 83:191–219
25. Bellman R, Astrom KJ (1970) On structural identifiability. Math Biosci 7:329–339
26. Cobelli C, DiStefano JJ (1980) Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. Am J Phys 239:R7–R24
27. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25:1923–1929
28. Neubauer P, Cruz-Bournazou MN (2017) Continuous bioprocess development: methods for control and characterization of the biological system. In: Subramanian G (ed) Continuous biomanufacturing – innovative technologies and methods. Wiley, Hoboken, pp 1–30
29. Anane E, García ÁC, Haby B, Hans S, Krausch N, Krewinkel M, Hauptmann P, Neubauer P, Cruz Bournazou MN (2019) A model-based framework for parallel scale-down fed-batch cultivations in mini-bioreactors for accelerated phenotyping. Biotechnol Bioeng 116:2906–2918
30. Neubauer P, Glauche F, Cruz-Bournazou MN (2017) Editorial: bioprocess development in the era of digitalization. Eng Life Sci 17:1140–1141
31. Narayanan H, Luna MF, von Stosch M, Cruz Bournazou MN, Polotti G, Morbidelli M, Butté A, Sokolov M (2020) Bioprocessing in the digital age: the role of process models. Biotechnol J 15:1–10
32. Noorman H (2011) An industrial perspective on bioreactor scale-down: what we can learn from combined large-scale bioprocess and model fluid studies. Biotechnol J 6:934–943
33. Petsagkourakis P, Sandoval IO, Bradford E, Zhang D, del Rio-Chanona EA (2020) Reinforcement learning for batch bioprocess optimization. Comput Chem Eng 133:106649
34. Mandenius CF, Brundin A (2008) Bioprocess optimization using design-of-experiments methodology. Biotechnol Prog 24:1191–1203
35. Wechselberger P, Sagmeister P, Herwig C (2013) Model-based analysis on the extractability of information from data in dynamic fed-batch experiments. Biotechnol Prog 29:285–296
36. Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. Bioinformatics 24:2044–2050
37. Haringa C, Deshmukh AT, Mudde RF, Noorman HJ (2017) Euler-Lagrange analysis towards representative down-scaling of a 22 $m^3$ aerobic *S. cerevisiae* fermentation. Chem Eng Sci 170:653–669
38. Delvigne F, Goffin P (2014) Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol J 9:61–72
39. Barz T, Lopez Cardenas DC, Cruz Bournazou MN, Körkel S, Walter SF (2016) Real-time adaptive input design for the determination of competitive adsorption isotherms in liquid chromatography. Comput Chem Eng 94:104–116
40. Cruz Bournazou MN, Barz T, Nickel DB, Lopez Cárdenas DC, Glauche F, Knepper A, Neubauer P (2017) Online optimal experimental re-design in robotic parallel fed-batch cultivation facilities. Biotechnol Bioeng 114:610–619

41. Dörr M, Fibinger MPC, Last D, Schmidt S, Santos-Aberturas J, Böttcher D, Hummel A, Vickers C, Voss M, Bornscheuer UT (2016) Fully automatized high-throughput enzyme library screening using a robotic platform. Biotechnol Bioeng 113:1421–1432
42. Haby B, Hans S, Anane E, Sawatzki A, Krausch N, Neubauer P, Cruz Bournazou MN (2019) Integrated robotic mini bioreactor platform for automated, parallel microbial cultivation with online data handling and process control. SLAS Technol 24:569–582
43. Unthan S, Radek A, Wiechert W, Oldiges M, Noack S (2015) Bioprocess automation on a mini pilot plant enables fast quantitative microbial phenotyping. Microb Cell Factories 14:32
44. Lattermann C, Büchs J (2015) Microscale and miniscale fermentation and screening. Curr Opin Biotechnol 35:1–6
45. Hemmerich J, Noack S, Wiechert W, Oldiges M (2018) Microbioreactor systems for accelerated bioprocess development. Biotechnol J 13:1–9
46. Lara AR, Galindo E, Ramírez OT, Palomares LA (2006) Living with heterogeneities in bioreactors: understanding the effects of environmental gradients on cells. Mol Biotechnol 34:355–382
47. Oldshue JY (1966) Fermentation mixing scale-up techniques. Biotechnol Bioeng 8:3–24
48. Oosterhuis NMG (1984) Scale-up of bioreactors. TU Delft 162
49. Enfors SO, Jahic M, Rozkov A, Xu B, Hecker M, Jürgen B, Krüger E, Schweder T, Hamer G, O'Beirne D, Noisommit-Rizzi N, Reuss M, Boone L, Hewitt C, McFarlane C, Nienow A, Kovacs T, Trägårdh C, Fuchs L, Revstedt J, Friberg PC, Hjertager B, Blomsten G, Skogman H, Hjort S, Hoeks F, Lin HY, Neubauer P, Van der Lans R, Luyben K, Vrabel P, Manelius A, Manelius Å (2001) Physiological responses to mixing in large scale bioreactors. J Biotechnol 85:175–185
50. Larsson G, Törnkvist M, Ståhl Wernersson E, Trägårdh C, Noorman H, Enfors SO (1996) Substrate gradients in bioreactors: origin and consequences. Bioprocess Eng 14:281–289
51. Brand E, Junne S, Anane E, Cruz-Bournazou MN, Neubauer P (2018) Importance of the cultivation history for the response of *Escherichia coli* to oscillations in scale-down experiments. Bioprocess Biosyst Eng 41:1305–1313
52. Sweere APJ, Luyben KCAM, Kossen NWF (1987) Regime analysis and scale-down: tools to investigate the performance of bioreactors. Enzym Microb Technol 9:386–398
53. Lin HY, Mathiszik B, Xu B, Enfors SO, Neubauer P (2001) Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of *Escherichia coli*. Biotechnol Bioeng 73:347–357
54. Bylund F, Collet E, Larsson G, Enfors SO, Larsson G (1998) Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. Bioprocess Eng 18:171–180
55. Neubauer P, Häggström L, Enfors SO (1995) Influence of substrate oscillations on acetate formation and growth yield in *Escherichia coli* glucose limited fed-batch cultivations. Biotechnol Bioeng 47:139–146
56. Xu B, Jahic M, Blomsten G, Enfors SO (1999) Glucose overflow metabolism and mixed-acid fermentation in aerobic large-scale fed-batch processes with *Escherichia coli*. Appl Microbiol Biotechnol 51:564–571
57. Nienow AW (2006) Reactor engineering in large scale animal cell culture. Cytotechnology 50:9–33
58. Simen JD, Löffler M, Jäger G, Schäferhoff K, Freund A, Matthes J, Müller J, Takors R, Feuer R, von Wulffen J, Lischke J, Ederer M, Knies D, Kunz S, Sawodny O, Riess O, Sprenger G, Trachtmann N, Nieß A, Broicher A (2017) Transcriptional response of *Escherichia coli* to ammonia and glucose fluctuations. Microb Biotechnol 10:858–872
59. Spann R, Glibstrup J, Pellicer-Alborch K, Junne S, Neubauer P, Roca C, Kold D, Lantz AE, Sin G, Gernaey KV, Krühne U (2019) CFD predicted pH gradients in lactic acid bacteria cultivations. Biotechnol Bioeng 116:769–780

60. Paul K, Böttinger K, Mitic BM, Scherfler G, Posch C, Behrens D, Huber CG, Herwig C (2020) Development, characterization, and application of a 2-compartment system to investigate the impact of pH inhomogeneities in large-scale CHO-based processes. Eng Life Sci 20:368–378
61. Paul K, Hartmann T, Posch C, Behrens D, Herwig C (2020) Investigation of cell line specific responses to pH inhomogeneity and consequences for process design. Eng Life Sci 20:412–421
62. Buchholz J, Graf M, Freund A, Busche T, Kalinowski J, Blombach B, Takors R (2014) CO 2/HCO 3− perturbations of simulated large scale gradients in a scale-down device cause fast transcriptional responses in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 98:8563–8572
63. Spadiut O, Rittmann S, Dietzsch C, Herwig C (2013) Dynamic process conditions in bioprocess development. Eng Life Sci 13:88–101
64. Limberg MH, Joachim M, Klein B, Wiechert W, Oldiges M (2017) pH fluctuations imperil the robustness of *C. glutamicum* to short term oxygen limitation. J Biotechnol 259:248–260
65. Xu S, Jiang R, Mueller R, Hoesli N, Kretz T, Bowers J, Chen H (2018) Probing lactate metabolism variations in large-scale bioreactors. Biotechnol Prog 34:756–766
66. Brunner M, Doppler P, Klein T, Herwig C, Fricke J (2018) Elevated pCO2 affects the lactate metabolic shift in CHO cell culture processes. Eng Life Sci 18:204–214
67. Delvigne F, Noorman H (2017) Scale-up/scale-down of microbial bioprocesses: a modern light on an old issue. Microb Biotechnol 10:685–687
68. Cortés JT, Flores N, Bolívar F, Lara AR, Ramírez OT (2016) Physiological effects of pH gradients on *Escherichia coli* during plasmid DNA production. Biotechnol Bioeng 113:598–611
69. Junne S, Klingner A, Kabisch J, Schweder T, Neubauer P (2011) A two-compartment bioreactor system made of commercial parts for bioprocess scale-down studies: impact of oscillations on *Bacillus subtilis* fed-batch cultivations. Biotechnol J 6:1009–1017
70. Käß F, Hariskos I, Michel A, Brandt HJ, Spann R, Junne S, Wiechert W, Neubauer P, Oldiges M (2014) Assessment of robustness against dissolved oxygen/substrate oscillations for *C. glutamicum* DM1933 in two-compartment bioreactor. Bioprocess Biosyst Eng 37:1151–1162
71. Schilling BM, Pfefferle W, Bachmann B, Leuchtenberger W, Deckwer W-DD (1999) A special reactor design for investigations of mixing time effects in a scaled-down industrial L-lysine fed-batch fermentation process. Biotechnol Bioeng 64:599–606
72. Delvigne F, Boxus M, Ingels S, Thonart P (2009) Bioreactor mixing efficiency modulates the activity of a *prpoS::GFP* reporter gene in *E. coli*. Microb Cell Fact 8:15
73. Junne S, Neubauer P (2018) How scalable and suitable are single-use bioreactors? Curr Opin Biotechnol 53:240–247
74. Löffler M, Simen JD, Jäger G, Schäferhoff K, Freund A, Takors R (2016) Engineering *E. coli* for large-scale production – strategies considering ATP expenses and transcriptional responses. Metab Eng 38:73–85
75. Delvigne F, Baert J, Sassi H, Fickers P, Grünberger A, Dusny C (2017) Taking control over microbial populations: current approaches for exploiting biological noise in bioprocesses. Biotechnol J 12:1600549
76. Lemoine A, Delvigne F, Bockisch A, Neubauer P, Junne S (2017) Tools for the determination of population heterogeneity caused by inhomogeneous cultivation conditions. J Biotechnol 251:84–93
77. Avery SV (2006) Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Microbiol 4:577–587
78. Binder D, Drepper T, Jaeger KE, Delvigne F, Wiechert W, Kohlheyer D, Grünberger A (2017) Homogenizing bacterial cell factories: analysis and engineering of phenotypic heterogeneity. Metab Eng 42:145–156

79. Lieder S, Jahn M, Seifert J, von Bergen M, Müller S, Takors R (2014) Subpopulation-proteomics reveal growth rate, but not cell cycling, as a major impact on protein composition in Pseudomonas putida KT2440. AMB Express 4:1–10
80. Delvigne F, Zune Q, Lara AR, Al-Soud W, Sørensen SJ (2014) Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. Trends Biotechnol 32:608–616
81. Van Heerden JH, Wortel MT, Bruggeman FJ, Heijnen JJ, Bollen YJM, Planqué R, Hulshof J, O'Toole TG, Wahl SA, Teusink B (2014) Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. Science 343:1245114
82. Brognaux A, Han S, Sørensen SJ, Lebeau F, Thonart P, Delvigne F (2013) A low-cost, multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors. Microb Cell Fact 12
83. Lieder S, Jahn M, Koepff J, Muller S, Takors R (2016) Environmental stress speeds up DNA replication in *Pseudomonas putida* in chemostat cultivations. Biotechnol J 11:155–163
84. Hewitt CJ, von Caron GN, Axelsson B, McFarlane CM, Nienow AW (2000) Studies related to the scale-up of high-cell-density *E. coli* fed-batch fermentations using multiparameter flow cytometry: effect of a changing microenvironment with respect to glucose and dissolved oxygen concentration. Biotechnol Bioeng 70:381–390
85. Patel A, Antonopoulou I, Enman J, Rova U, Christakopoulos P, Matsakas L (2019) Lipids detection and quantification in oleaginous microorganisms: an overview of the current state of the art. BMC Chem Eng 1
86. Marbà-Ardébol AM, Emmerich J, Neubauer P, Junne S (2017) Single-cell-based monitoring of fatty acid accumulation in *Crypthecodinium cohnii* with three-dimensional holographic and in situ microscopy. Process Biochem 52:223–232
87. Marbà-Ardébol AM, Bockisch A, Neubauer P, Junne S (2018) Sterol synthesis and cell size distribution under oscillatory growth conditions in *Saccharomyces cerevisiae* scale-down cultivations. Yeast 35:213–223
88. Lemoine A, Limberg MHMH, Kästner S, Oldiges M, Neubauer P, Junne S (2016) Performance loss of *Corynebacterium glutamicum* cultivations under scale-down conditions using complex media. Eng Life Sci 16:620–632
89. Nachin L, Nannmark U, Nyström T (2005) Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. J Bacteriol 187:6265–6272
90. Trivedi A, Mavi PS, Bhatt D, Kumar A (2016) Thiol reductive stress induces cellulose-anchored biofilm formation in *Mycobacterium tuberculosis*. Nat Commun 7
91. Kurt T, Marbà-Ardébol AM, Turan Z, Neubauer P, Junne S, Meyer V (2018) Rocking Aspergillus: morphology-controlled cultivation of *Aspergillus niger* in a wave-mixed bioreactor for the production of secondary metabolites. Microb Cell Factories 17:128
92. Lin PJ, Scholz A, Krull R (2010) Effect of volumetric power input by aeration and agitation on pellet morphology and product formation of *Aspergillus niger*. Biochem Eng J 49:213–220
93. Gómez-Ríos D, Junne S, Neubauer P, Ochoa S, Ríos-Estepa R, Ramírez-Malule H (2019) Characterization of the metabolic response of *Streptomyces clavuligerus* to shear stress in stirred tanks and single-use 2D rocking motion bioreactors for clavulanic acid production. Antibiotics 8
94. Hardy N, Moreaud M, Guillaume D, Augier F, Nienow A, Béal C, Chaabane FB (2017) Advanced digital image analysis method dedicated to the characterization of the morphology of filamentous fungus. J Microsc 266:126–140
95. Kuschel M, Siebler F, Takors R (2017) Lagrangian trajectories to predict the formation of population heterogeneity in large-scale bioreactors. Bioengineering 4:27
96. Ladner T, Grünberger A, Probst C, Kohlheyer D, Büchs J, Delvigne F (2017) Application of mini- and micro-bioreactors for microbial bioprocesses. In: Current developments in biotechnology and bioengineering: bioprocesses, bioreactors and controls. Elsevier, Amsterdam, pp 433–461

97. Morchain J, Gabelle J-C, Cockx A (2013) Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. AICHE J 59:369–379
98. Pigou M, Morchain J (2015) Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. Chem Eng Sci 126:267–282
99. Lapin A, Müller D, Reuss M (2004) Dynamic behavior of microbial populations in stirred bioreactors simulated with Euler−Lagrange methods: traveling along the lifelines of single cells. Ind Eng Chem Res 43:4647–4656
100. Lapin A, Klann M, Reuss M (2010) Multi-scale spatio-temporal modeling: lifelines of microorganisms in bioreactors and tracking molecules in cells. Adv Biochem Eng Biotechnol 121:23–43
101. Anane E, López DC, Neubauer P, Cruz Bournazou MN (2017) Modelling overflow metabolism in *Escherichia coli* by acetate cycling. Biochem Eng J 125:23–30
102. Lara AR, Taymaz-Nikerel H, Mashego MR, Van Gulik WM, Heijnen JJ, Ramirez OT, van Winden WA, Van Gulik WM, Heijnen JJ, Van Winden WA, Ramírez OT, van Winden WA (2009) Fast dynamic response of the fermentative metabolism of *Escherichia coli* to aerobic and anaerobic glucose pulses. Biotechnol Bioeng 104:1153–1161
103. Soini J, Ukkonen K, Neubauer P (2011) Accumulation of amino acids deriving from pyruvate in *Escherichia coli* W3110 during fed-batch cultivation in a two-compartment scale-down bioreactor. Adv Biosci Biotechnol 02:336–339
104. Barz T, Sommer A, Wilms T, Neubauer P, Cruz Bournazou MN (2018) Adaptive optimal operation of a parallel robotic liquid handling station. IFAC-PapersOnLine 51:765–770
105. Lemoine A, Martnez-Iturralde NM, Spann R, Neubauer P, Junne S (2015) Response of Corynebacterium glutamicum exposed to oscillating cultivation conditions in a two- and a novel three-compartment scale-down bioreactor. Biotechnol Bioeng 112:1220–1231
106. Korneli C, Bolten CJ, Godard T, Franco-Lara E, Wittmann C, Universita T (2012) Debottlenecking recombinant protein production in Bacillus megaterium under large-scale conditions—targeted precursor feeding designed from metabolomics. Biotechnol Bioeng 109:1538–1550
107. Janakiraman V, Kwiatkowski C, Kshirsagar R, Ryll T, Huang YM (2015) Application of high-throughput mini-bioreactor system for systematic scale-down modeling, process characterization, and control strategy development. Biotechnol Prog 31:1623–1632
108. Gábor A, Banga JR (2015) Robust and efficient parameter estimation in dynamic models of biological systems. BMC Syst Biol 9:74
109. Rollié S, Mangold M, Sundmacher K (2012) Designing biological systems: systems engineering meets synthetic biology. Chem Eng Sci 69:1–29
110. Bareither R, Pollard D (2011) A review of advanced small-scale parallel bioreactor technology for accelerated process development: current state and future need. Biotechnol Prog 27:2–14
111. Rameez S, Mostafa SS, Miller C, Shukla AA (2014) High-throughput miniaturized bioreactors for cell culture process development: reproducibility, scalability, and control. Biotechnol Prog 30:718–727
112. Herwig C, Garcia-Aponte OF, Golabgir A, Rathore AS (2015) Knowledge management in the QbD paradigm: manufacturing of biotech therapeutics. Trends Biotechnol 33:381–387
113. Rathore AS (2009) Roadmap for implementation of quality by design (QbD) for biotechnology products. Trends Biotechnol 27:546–553
114. Velez-Suberbie ML, Betts JPJ, Walker KL, Robinson C, Zoro B, Keshavarz-Moore E (2017) High-throughput automated microbial bioreactor system used for clone selection and rapid scale-down process optimization. Biotechnol Prog 15:1–11
115. de Lorenzo V, Schmidt M (2018) Biological standards for the knowledge-based BioEconomy: what is at stake. New Biotechnol 40:170–180
116. Schallmey M, Frunzke J, Eggeling L, Marienhagen J (2014) Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Curr Opin Biotechnol 26:148–154