Evolution of the activated sludge process: the first 50 years

Derin Orhon\textsuperscript{a,b}\textsuperscript{*}

Abstract

The paper presents a critical overview of the first 50 years in the evolution of the activated sludge process. Recognition of the role of aeration and microbial activity in the purification of sewage in the early studies established the basis for the accidental discovery of the process, which was immediately adapted into practice. The problems encountered during operation started a period of empirical expansion with many process modifications. As scientific support lagged behind practice, efforts were then directed towards exploring and understanding the fundamentals of the system related to substrate removal mechanisms, microbiology, process kinetics and stoichiometry, nitrogen transformations and, then, translating accumulated information into design. Extensive research generated remarkable findings which should be regarded as major milestones for future progress in many areas such as fractionation of substrate and biomass, particle size distribution, and substrate storage. A rational foundation of the activated sludge process for the removal of organic carbon could then be established based on mass balance. This approach also defined essential parameters that could relate microbial reactions involved with plant design and operation. Scientific ideas and discoveries have also enabled conceptual development of many emerging technologies such as the sequencing batch reactor, enhanced biological phosphorus removal, the oxic-settling anaerobic process, and the super-fast membrane bioreactor.

© 2014 Society of Chemical Industry

Keywords: activated sludge; historical evolution; system modifications; substrate removal; process kinetics and stoichiometry; microbial storage; microbiology; particle size distribution; emerging technology

INTRODUCTION

The activated sludge process is perhaps one of the most complex microbial systems ever engineered for a specific purpose. Several different groups of heterotrophs and autotrophs can now be sustained together to perform different functions, such as removal of organic carbon, nitrification and denitrification, enhanced biological phosphorus removal, etc. Today, their metabolic activities may be predicted by modelling, and individually controlled by elaborate manipulation of system parameters, to ensure optimum process performance. Considerable activity is now directed towards developing novel processes to meet new waste management objectives; however, each new idea brings together several significant unanswered questions concerning the microbial mechanisms involved. So the continual progress of activated sludge relies on a never-ending quest for more science to face and tackle new problems.

However, the discovery and the initial development of the activated sludge process was more a work of art than science. In fact, its early phases passed in total darkness concerning the two main ingredients: (1) the substrate to be treated and (2) the biomass that treats it. Therefore, the process established itself largely on the basis of empirical experience. In 1915, only a year after its discovery, Fuller\textsuperscript{1} noted:

It is pleasant to be able to record that the art of sewage disposal is in a state of active progression. Many works in the field introduce changes and improvements, some in matters of detail or design, others in the matter of affecting the fundamental principles of operation.

The rapid initial incentive has to be praised because of the common sense and ingenuity that always accompanied successful applications. Indeed, the expanding practice could not wait upon the findings of science, when it was required to start something that would work for sewage purification, regardless of the state of knowledge at the moment. The limited experience available attracted substantial engineering interest to expand it by means of pilot plant operations and, also, to develop more scientific information through laboratory experiments. Porter\textsuperscript{2} lists 200 references related to the activated sludge process in 1914, which rapidly increased to 800 by 1920.

Only a few scientists actively working on various aspects of the activated sludge process are aware of its colourful and instructive development in the past. We should all begin to benefit from it by first paying respect to the outstanding scientists such as Buswell, Heukelekian, Sawyer, McKinney, etc. for having prepared the scientific ground from which the process flourished. So, basically, the early phases should be regarded as the necessary scientific gestation period for all the major achievements that the process enjoys today.

In this context, the objective of this paper is to provide a review on the evolution of the activated sludge process during the first

\textsuperscript{*} Correspondence to: Derin Orhon, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey. E-mail: orhon@itu.edu.tr

\textsuperscript{a} Istanbul Technical University, 34469 Maslak, Istanbul, Turkey

\textsuperscript{b} The Science Academy, 34353 Beşiktaş, Istanbul, Turkey
50 years after its discovery. It was prepared not as a historical account of events but more as an overview of pioneering ideas and findings filtered and purified through much research, which greatly contributed to the understanding of major issues, such as substrate storage, particle size distribution, process kinetics and stoichiometry, reactor hydraulics and mass balance, etc., also referring to related emerging technologies. The period was slightly extended by a few years to include major research work that should be considered as milestones for future progress. The text has avoided mathematical expressions to the extent that makes it possible to better focus on emerging ideas and has adopted a currently used uniform nomenclature for the selected equations, where needed. Direct quotations are also included in the text; while some may appear to be wrong, simple or even naive, it is believed that they all provide a colourful richness in reflecting the amazing evolution in the conceptual understanding of the process.

EARLY PRACTICE AND PIONEERING WORK
Historical approach to sewage disposal
The concept of sanitation and hygiene first flourished during ancient civilisations. Archeological studies have revealed amazing sewage disposal systems from the early periods in Mesopotamian, Indus, Egyptian, Greek and Roman history. Unfortunately, the art and practice of wastewater management gradually faded away and could not survive the dark years of the medieval era. As noted by Angelakis et al., ‘in the light of history it is a matter of astonishment, if not chagrin, that man in this respect has progressed so very little, if at all, in some 4000 years’. In fact, at the turn of the 19th century, most villages and towns did not have a proper sewage disposal system. Liquid wastes and garbage were generally poured onto the unpaved streets, which were soon covered with mud, including a good portion of sewage sludge, particularly after heavy rain. For a long period, houses with no toilets in the courtyards kept on emptying their ‘chamber pots’ onto the street or in the nearest creek. After 1830, small open channels were gradually constructed in the middle of the streets or directly near sidewalks. These drains soon became a part of city life, discharging human and household wastes waste along with stormwaters to the nearest water course. They were frequently clogged.

After the first half of the 19th century, towns in developed countries showed signs of awareness and responsibility for the sanitary problems of urban life, safe water supply and proper disposal of sewage and refuse. Hamlin provides an excellent account of this period: in England, Edwin Chadwick, an English lawyer, credited with encouraging the ‘sanitary movement’, argued that people could have access to ample supply of safe water if sewage was collected in sewers that would convey it to suburban market gardens and farms for use as fertiliser. This idea was later modified in the mid-1860s into building sewers that would take sewage downstream along the River Thames. This way, dilution and irrigation emerged as two major alternatives of sewage disposal into water or upon land.

Two different approaches were adopted for sewage disposal through irrigation: sewage farming and sub-surface irrigation. Sewage farming dates back to the earliest days of sewerage systems. In this process, sewage was spread over cultivated fields, seeped through the ground until it reached the ground-water table or flowed into drainage channels and finally into creeks or rivers. In sub-surface irrigation, sewage was distributed below ground level and penetrated into the soil from open jointed pipes. Sewage farming was carried out for two main purposes: (1) proper disposal of sewage, and (2) watering and, to some extent, fertilising crops as part of a commercial process. In sub-surface irrigation, the commercial aspect was, at best, incidental. The sewage farm method of sewage disposal was used almost exclusively in England from 1870 to 1882; the operation was supervised and controlled by The Royal Commission of Sewage Disposal. It was also practised to a large extent around large cities in Europe. In 1910, sewage farms in Berlin were receiving 290 000 m$^3$ day$^{-1}$ (77 MGD; where MGD is million gallons per day) of sewage. In Paris, around 450 000 m$^3$ day$^{-1}$ (120 MGD) of sewage, representing 60% of the daily load in 1923, was used for irrigation and the remaining fraction was discharged into the River Seine. The historical sewage disposal system in Paris may best be recognised today by the famous outcry of the author Victor Hugo, urging that the sewage in Paris be used for irrigating farms around the city. Unfortunately, the practice did not prove as fruitful as he expected.

Do you know what those piles of ordure are, collected at the corners of streets, those carts of mud carried off at night from the streets, the frightful barrels of the night-man, and the fetid streams of subterranean mud which the pavement conceals from you? All this is a flowering field, it is green grass, it is mint, thyme and sage, ... it is cattle, it is gilded wheat, it is bread on your table, it is health, it is joy, it is life. (Les Miserables, Book II, Chapter I, Valjean).

Sewage farming was later abandoned to a large extent, mainly because farmers were more interested in newly available synthetic fertilisers ‘rather than sewage that kept coming whether it was needed or not’, and also because it was costly, inefficient and often created a nuisance.

During the course of the 19th century, it became common practice to discharge wastewaters into the rivers and large water bodies. In 1915, 42.3% of 99 million people in the US were served by sewerage, 26.7% discharging untreated wastes into lakes and streams and 8.6% into the sea; 7% of the sewage was discharged into waters, after receiving some degree of treatment. Despite the growing practice, the concept of dilution through discharge did not meet with immediate unanimous approval among experts, who also argued that sewers had been a needless expenditure as their source of water supply. Later, rapid urban and industrial development created a growing concern for untreated wastewater discharges. In short, sewerage systems served a good purpose but discharge of collected wastewaters was no solution to the problem, if the discharge flow was too high compared to the capacity and the flow rate of the river. The pollution problems encountered helped to recognise the significance of biological factors related to sewage discharges and to explore scientific basis of these factors. In 1863, the discovery by Pasteur revealing that fermentation and ‘putrefaction’ could not take place without living microorganisms was the major step towards the scientific explanation of observed pollution caused by sewage discharges. The 19th century was also marked by major epidemics, and nobody could suggest a realistic explanation about the causative agents, until Robert Koch
discovered the bacteria which caused the illness. It was also argued that a similar cholera epidemic started in Hamburg in 1892.

Role of aeration and microbial activity
The problems observed related to the putrefaction of organic matter due to sewage disposal onto land or into the rivers initiated ideas on different treatment strategies. In 1882, The Royal Commission recommended chemical settling of suspended solids in sewage with the objective of ‘sterilization of the effluents’. It is interesting to note that in this approach, purification by chemical treatment was conceived as prevention of decomposition. In an effort to ‘deodorise sewage’, Dibdin tested chemical treatments with three most common precipitants – lime, alum and iron sulfate – and obtained limited organic matter removal of around 20% regardless of the selected chemical. He claimed that the treated effluent would not putrefy in the river, although admitting that on hot days when the oxygen solubility was low, an ‘artificial source of oxygen’ might be necessary; in that case, he proposed to further treat the effluent with permanganate of soda, a powerful oxidising agent. Later, Roscoe continued the ‘deodorisation’ application using sulfuric acid and chloride of lime, but with no success. He admitted his failure but chose to blame dry weather and low water, rather than his choice of chemicals.

The same period was also a milestone, which was marked with amazing discoveries shaping the basic fundamentals of the activated process. Scientific observations during this period revealed the essential function of the microbial community and the vital role of oxygen in the removal of organic matter. This way, the conceptual understanding of purification or treatment was drastically changed and redefined as a system sustaining biological processes that would destroy organic matter in sewage. A detailed account of related observations is presented by Hamlin. In 1882, after a microscopic survey of mud samples from the Thames Estuary, Sorby concluded that ‘sewage was consumed by minute organisms, mostly crustacea’. Dupre argued that it was oxidized by aerobic bacteria and suggested that the same principle could be applied in a sewage treatment plant. Dibdin (quoted in Hamlin) reiterated the same argument in 1887: ‘the future is going to let the organisms in the rivers purify the sewage, for free’ and in 1988 ‘our treatment should be such as to avoid the killing of these organisms or even hampering them in their actions but rather to do everything to favour them in their beneficial work’.

Works of biological filtration
These observations inspired Dibdin to test his ‘biological considerations’ on a filter-type of treatment scheme where microbial action and aeration would be involved as basic ingredients for organic matter removal from sewage. In 1893, he started filtration experiments with different types of filter media: pea ballast, coke breeze, gravel, burnt clay and a combination of materials including polystyrene. He observed that bacteria could completely purify sewage or fail to do so, when overloaded. The performance of the filter could be maintained by adjusting the rates of several processes: organic load, aeration and growth of bacteria, as he noted that ‘the life of the filter, in fact appears to be practically without limit, providing always that the balance between aeration and food supply to the organisms be preserved’. This way, the contact bed filter was developed in the early 1890s, as the pioneering biological process for sewage treatment. While it was developed mainly to meet political ends, the success of the process ensured the public that the sewage problem was solved in an economical, effective and rational way. In 1896, the system was tested at the treatment plant in Sutton, near London. By the end of the first decade of the 20th century, several towns and cities adopted this new technology offering biological treatment for sewage. Later, Imhoff and Moore argued that for 20 years, the practice did not recognise that the contact bed application as a fill and draw system was wrong in principle; the contact bed would be filled with sewage and emptied after a suitable time had elapsed. Nevertheless, the experience paved the way for the adoption of still higher rates of treatment by allowing sewage to trickle downward through a bed of coarse filter material and, eventually, improved the process by converting it into continuously operated trickling filters.

Aeration experiments
Successful experience with biological filters also tempted scientists to conduct studies testing direct aeration of sewage. The work of Dr Angus Smith in 1882 is commonly cited as the original study where aeration was directly applied to sewage. Two years later, Dupre and Dibdin were recognized for their pioneering work on biological treatment, also attempting to oxidize sewage by aeration; they reported that the observed oxidation proceeded too slowly and therefore was not at all practicable. After a similar experiment in 1892, Mason and Hine also concluded that air had but little oxidising effect on sewage.

During the following few years, studies also focused on the suspended matter generated and its beneficial impact for improving odour. In 1893, Mather and Platt argued that the ‘precipitated impurities’ accumulated at the bottom of the aeration tank enhanced the level of treatment achieved. In 1897, Fowler aerated Manchester sewage and obtained a clear effluent with rapidly settleable particulate matter; the results were reported as a failure because Fowler personally believed that the particulate deposits must be rendered soluble or gasified for optimal treatment. In 1911, aeration was again tested by Black and Phelps with the sewage of New York City, for various periods up to 24 h in tanks filled with inclined wooden gratings. Only a slight oxidation could be observed based on nitrogen measurements; despite limited efficiency, the process was adopted for a large-scale installation.

In short, very little was obtained after 20 years of extensive work, mainly because the experiments mostly overlooked the magical clue – the role of humus-like material generated during the aeration process. Fuller thought that the accumulated information was only useful in defining a pretreatment step for successful filtration. The disappointment for the on-going studies was best expressed by Fowler when he stated that:

The scientific solution of the sewage disposal problems will not be attained until the following results can be guaranteed for any given case: (1) An effluent which will not deteriorate the stream into which it flows; (2) no nuisance in the course of sludge disposal; (3) no nuisance from smell or flies in connection with the filter beds; (4) an expenditure strictly proportionate to the sanitary and esthetic results achieved.

DISCOVERY AND INITIAL YEARS
Today, there is still debate on the original discovery of the activated sludge process. Its starting point is generally attributed to the work of Ardern and Lockett as presented to The Society of the Chemical Industry in 1914. The International Water Association (IWA) organised an international conference for this purpose in 2014 entitled 'Activated sludge – 100 years and counting'. On the
other hand, Jeanette Brown, a former president of the Water Environment Federation celebrated the birthday of activated sludge in 2013, referring to the studies of Clark and Gage reported in 1913. Therefore, it would only be fair to do justice to the efforts of both parties for this monumental step forward for the biological treatment of wastewaters.

The 20 years between 1910 and 1930, while largely ignored or overlooked in the related literature, were a period with remarkable achievements in the basic understanding, experimentation and full-scale implementation of the process. This period undoubtedly started with the outstanding research conducted at Lawrence Experiment Station in Massachusetts; reports and bulletins issued should be regarded as collectors' items for those interested in the evolution of the process as well as for active researchers in this field. Studies reported by Clark and Gage in 1912 primarily tested aeration of sewage as an aid to filtration. In the preliminary part of the experimental programme a perceptible clarification was obtained and when ‘green growths’ developed in the aerated sewage, it became saturated with oxygen; nitrification occurred in the aerated sewage when fed to a small sand filter. In the second part, an aeration tank containing slabs of slate was put into operation; it was observed that the slate and the sides of the tank became covered with a ‘compact brown growth’, which collected the suspended and colloidal matter in sewage after a few hours. The same study was continued for testing the effect of air supply on ‘bacterial filters’. Similar experiments were also conducted with the same success on four different sewages. The influence of sewage temperature was also investigated in the experiments, yielding good results between 12 and 20°C; significant deterioration of the removal rates was observed below 10°C. Sewage temperatures above 30°C caused slight deterioration in producing a clarified effluent. The authors also argued that while the experiments were carried out in fill and draw systems, equally good results would be obtained with less aeration period in a continuous-flow system. An original account of the experimental results leading to their discovery is displayed in Fig. 1. Years later, the rewarding sensation of this new discovery was best described by Jenkins.

Consider what must have passed through their minds as the content of their cylinders became darker and darker, more and more dirty looking; but imagine their delight when they found that these dirty grey–brown mixtures, when settled, were associated with higher and higher wastewater purification rates.

In 1930, Clark was still claiming that the activated sludge was discovered and put into operation at Lawrence Experiment Station in 1911 and 1912. However, the true credit for the discovery should go to the work of Ardern and Lockett, mainly because they were able to (1) define and operate a process fully based upon a suspended-growth microbial community, (2) recognise

Table II. Oxidation of Manchester sewage.

<table>
<thead>
<tr>
<th>Experiment 14</th>
<th>Experiment 22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw sewage</td>
<td>R.S.† + sludge</td>
</tr>
<tr>
<td>4 hours oxygen absorption</td>
<td>54-0-0</td>
</tr>
<tr>
<td>Per cent. reduction on raw sewage</td>
<td>---</td>
</tr>
<tr>
<td>Free and soluble ammonia</td>
<td>0.8</td>
</tr>
<tr>
<td>Per cent. reduction on raw sewage</td>
<td>---</td>
</tr>
<tr>
<td>Alkaloidal ammonia</td>
<td>0.0</td>
</tr>
<tr>
<td>Per cent. reduction on raw sewage</td>
<td>---</td>
</tr>
<tr>
<td>Nitrate and nitrite (in terms of N2O)</td>
<td>0.0</td>
</tr>
<tr>
<td>Per cent. reduction on raw sewage</td>
<td>---</td>
</tr>
<tr>
<td>R.S.†</td>
<td>Original settled.</td>
</tr>
<tr>
<td>1 vol. sludge to 10 vol. sewage</td>
<td>---</td>
</tr>
</tbody>
</table>

* Original sewage with 20 per cent. water added, to obtain dilution effect of sludge.

Figure 1. The original experimental data supporting the discovery of the activated sludge process as presented by Ardern and Lockett.
The activated sludge process


‘activated sludge’, i.e. biomass in the reactor, as the essential control parameter to ensure the desired process efficiency, and (3) introduce, indirectly, oxygen as a control parameter for the system, recommending an air supply of 7.0–10 m³ air m⁻³ sewage (1.0–1.5 cu ft of air per gallon of sewage), a second design value which has survived over time. Consequently, their work, well beyond experimental observations, provided most of the major ingredients of a novel sewage treatment process. In fact, in discussing the study presented before The Society of Chemical Industry, Fowler said that the authors of the paper had now ‘fired a bombshell into the camp’ which would start new ways of sewage treatment in the future.

The process was quite well received with only a few sceptical opinions. Naylor argued that one volume of sludge to four to five volumes of sewage would be almost ‘too solid’ to be dealt with, since conventional precipitation methods involve one part of sludge to 110 parts of sewage and require 6 h for settling. His objection was, of course, totally unfounded as it overlooked the fact that the sludge referred to in the experiment was also quite diluted so that the dry solids content of the mixed liquor would be significantly lower than 1%. He also indicated that the air used in the process may require additional treatment to solve the problem of odour taken up from sewage.

Only a year after the work of Ardern and Lockett, Fuller published a review where he first stated that the ‘art of sewage disposal’ passed through a stage of active progression and some of the improvements elucidated the fundamentals of operation. He specifically referred to the work of Clark and Adams underlining the ‘peculiar property of brown jelly-like bacterial growth on the slate bodies’ of coagulating the colloidal matter in sewage and reducing the putrefying characteristics of sewage. He also mentioned that the work of Ardern and Lockett promised interesting possibilities mainly because filters would no longer be needed as ‘the so called activated sludge’ would continually provide a compatible level of sewage purification.

The process as described by Ardern and Lockett was immediately adopted and tested in major research centres; enormous amount of experimental work was put in progress throughout the world in more than 80 experimental plants, and 17 activated sludge plants were either completed or in the process of construction around 1922. In a comprehensive bibliography of the subject published in 1921, Porter indicated the consensus on the definitions of activated sludge and activated sludge treatment as:

- **Activated sludge**: A flocculent sludge of medium brown colour enveloped by masses of aerobic organisms possessing the power of rapidly oxidising and nitrifying sewage, and which, though low specific gravity, settles rapidly.

- **Activated sludge treatment**: A biochemical process by which the purification of sewage is accomplished by passing through tanks in which sewage sludge is artificially agitated and intimately mixed with sewage and is supplied with the requisite oxygen for the optimum development of a countless number of nitrifying organisms incorporated in and adhering to the sludge, the final settlement of which causes a distinct clarification of the oxidised sewage.

Experimental investigations were mostly conducted to make sure that the same results could be obtained, as correctly diagnosed by Clark in 1930: ‘Much research of past years was rediscovered by new workers and given to us again’. Nevertheless, some of the continued research activities were significant, especially in the US, as they paved the way for the rapidly expanding practice on activated sludge plants. The Lawrence Experiment Station assumed again the primary role. In August 1914, Bartow saw the work in progress at Manchester and, upon his return, suggested that a comprehensive study be carried out involving a series of experiments on the purification of sewage by aeration in the presence of activated sludge at the University of Illinois; they basically followed the lines described by Ardern and Lockett, mainly to become familiar with the process. In the first part, sewage was aerated without sludge; the time for complete nitrification varied from 15 to 33 days. In the second part, the experiments involved aeration of sewage with sludge; the treated supernatant was siphoned off and a fresh portion of sewage was added onto the sludge. During the second sequence of experiments, the time for complete nitrification was reduced from 15 days to 4 days. In the third experiment, after the supernatant had been removed again nitrification was complete in 2 days, in the 12th sequence in less than 8 h. Finally, in the 31st sequence with sludge and sewage in a proportion of 1:5, purification was complete in less than 5 h. The following part of the experiments using larger scale tanks evaluated the effect of cyclic sewage loadings; furthermore, reducing the cyclic sewage feeding time to 6 h was investigated in terms of sludge accumulation and process efficiency. The air supply used in the experiment was controlled and calculated to be around 7.5 m³ air m⁻³ of sewage, as also prescribed by Ardern and Lockett. In 1916, Fowler underlined another significant aspect, indicating that it was not necessary to nitrify sewage while building up sludge. Bartow and Hatfield also conducted pot studies and gardening experiments, which indicated that sludge could be a valuable fertiliser, since its nitrogen content was in a quite available form. Later, an experimental station was established at Mount Vernon, NY, early in 1919 to treat a flow of 170 m³ day⁻¹ (45000 GPD; where GPD is gallons per day), producing equally successful results. Detailed evaluation of the plant performance led to remarkable pioneering studies on the biology and the biochemistry of the activated sludge process.

The new discovery found positive response in practice both in England and in the US. It was rapidly commercialised in the UK and Jones and Atwood received the patent for the process in 1914. They soon started manufacturing air supply equipment and shortly later, they undertook the construction of a full-scale continuous flow activated sludge unit that would treat 1000 m³ day⁻¹ of sewage, around 10% of the works total flow. Supporting experiments were carried out hand in hand with the construction and commissioning of new plants, mostly on site.

In England, the first activated sludge plant for sewage treatment with 300 m³ day⁻¹ capacity was started as a fill and draw system at Salford, in 1914. It was shortly followed by a similar fill and draw plant at Davyhulme in 1915. These early plants also served as experimental stations. At Salford sewage works, experiments conducted as described by Ardern and Lockett proved to be quite successful. Around 285 m³ day⁻¹ of sewage could be treated in an aeration tank of 130 m³ filled with 25% sludge, with 90% ‘oxygen adsorption’ [biochemical oxygen demand (BOD) reduction] and 76% ammonia oxidation. The system was operated in a sequencing batch reactor mode with three cycles a day; each cycle was adjusted to 7 h, with 1 h of fill, 3 h of aerations, 2 h of settling and 1 h to empty, and provided complete treatment of sewage received at the end of the cycle. Melling also referred to the experiments conducted at the Salford sewage works, reporting that it was possible to clarify and purify raw sewage in 3 h, with an effluent that satisfied local and international standards.
In the US, the most extensive series of experiments were carried out at Milwaukee, Wisconsin, using fill and draw and continuous systems and leading to the design and construction of an activated sludge plant for the entire city. The results obtained were found to be quite promising to warrant full-scale application, though many practical problems were yet to be solved.\textsuperscript{41} Copeland,\textsuperscript{42} also reporting an interesting aspect of the experiments, indicated that by increasing the air supply from 12 to 17 m$^3$ air m$^{-3}$ sewage,\textsuperscript{43} occasionally falling to 4.5\textdegree C, at the Milwaukee treatment plant.

After the first small activated plants at Salford and Davyhulme, a total of 10 plants with capacities in the range of 300–750 m$^3$ day$^{-1}$ were constructed and put into operation in England, until 1921. In 1916, San Marcos, Texas, was the first town in the US to use activated sludge treatment for its entire sewage flow of 600 m$^3$ day$^{-1}$.\textsuperscript{15} The same year, in a paper presented to the Indiana Engineers Society, Hatton\textsuperscript{43} pointed out that the activated sludge process should not be confused with the kind of artificial aeration of sewage, since the process depended on the presence of biological life in the sludge on the basis of accumulated experimental findings. Nine more activated sludge plants were commissioned between 1916 and 1927, including the large installations at Milwaukee, Indianapolis and Chicago. It is interesting to note that, despite the initial fill and draw experiments, only two plants in England operated with the same principle; all the others were installed as continuous-flow systems. The main difference between the two countries was in the adopted aeration period. In England, treatment plants were generally designed for longer aeration periods of 8 to 12 h, mostly with mechanical aeration, whereas a shorter aeration time of 3 to 6 h were implemented in the US, typically equipped with diffused aeration.\textsuperscript{15}

The performance of all the pioneer activated sludge plants in their early years of operation was quite satisfactory, as summarised in Table 1.\textsuperscript{44} This not only provided fairly basic criteria for the principal design features of new plants, but firmly established the activated sludge process.

### Table 1. Performance of early activated sludge plants in the USA\textsuperscript{44}

<table>
<thead>
<tr>
<th>Treatment plant</th>
<th>Year</th>
<th>Sewage flow* (1000 m$^3$ day$^{-1}$)</th>
<th>Aeration time (h)</th>
<th>Air used* (m$^3$ air m$^{-3}$ sewage)</th>
<th>Effluent SS (mg L$^{-1}$)</th>
<th>Reduction in BOD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Springfield, Virginia</td>
<td>1930–1940</td>
<td>24</td>
<td>5.4</td>
<td>6.4</td>
<td>13</td>
<td>92/90/94</td>
</tr>
<tr>
<td>North Side, Chicago</td>
<td>1930–1940</td>
<td>730</td>
<td>4.9</td>
<td>3.0</td>
<td>13</td>
<td>92/89/95</td>
</tr>
<tr>
<td>Indianapolis, Indiana</td>
<td>1930–1940</td>
<td>62</td>
<td>8.8</td>
<td>9.6</td>
<td>24</td>
<td>90/83/93</td>
</tr>
<tr>
<td>Peoria, Illinois</td>
<td>1931–1940</td>
<td>43</td>
<td>8.1</td>
<td>7.5</td>
<td>20</td>
<td>92/89/97</td>
</tr>
<tr>
<td>San Antonio, Texas</td>
<td>1932–1939</td>
<td>65</td>
<td>7.4</td>
<td>7.2</td>
<td>32</td>
<td>87/70/97</td>
</tr>
<tr>
<td>West Milwaukee, Wisconsin</td>
<td>1933–1940</td>
<td>297</td>
<td>6.6</td>
<td>9.0</td>
<td>19</td>
<td>95/94/96</td>
</tr>
</tbody>
</table>

\*Average values.

BOD, biochemical oxygen demand; SS, suspended solids.

EMPIRICAL DEVELOPMENT

Although widely accepted, successful experience with the activated sludge plants did not last long. The magnitude and composition of sewage loads exhibited significant changes due to rapid population expansion and industrial development. Higher flows and stronger sewage gradually affected system performance, mostly in activated sludge plants built for smaller communities and towns with significant industrial activities. Significant problems developed and escalated, affecting both the effluent quality and the general attitude toward the process.

The most visible problem was the significant deterioration experienced in the settling properties of activated sludge, often leading to the escape of suspended solids in the effluent and even the partial loss of biomass in the aeration tank. The first experience with this phenomenon, described as ‘sludge bulking’,\textsuperscript{45} was at the Des Plaines River plant in Chicago.\textsuperscript{46} The incident was associated with the predominance of a filamentous bacterium \textit{Sphaerotilus} in the biomass and the source of this microorganism was traced to Melrose Park where a considerable amount of sugar was dumped into the sewer. Another interesting incident of bulking was observed at the Glenelg plant in Australia\textsuperscript{47} receiving a mixture of winery wastewaters with sewage. The high carbohydrate content of untreated winery wastewaters reaching at one point 1\% in volume of the sewage flow converted activated sludge to something closely approaching a pure culture of filamentous organisms’ within 1 or 2 weeks; the sludge indices went as high as 500–600 mL g$^{-1}$, leading to complete breakdown of the plant. Sludge bulking soon became a widespread problem, often related to substrate imbalance due to high carbohydrate ratio levels in sewage,\textsuperscript{48,49} or to low dissolved levels in the aeration tank.\textsuperscript{50} However, the cause and nature of the problem has never been fully understood.

The ever-increasing sewage loads also resulted in oxygen shortage, mainly at the head end of the aeration tanks. This was related to high initial ‘biochemical oxygen demand’, based on studies on the quantitative measurements of oxygen utilisation by activated sludge; it was concluded that provision must be made to secure and maintain oxygen in the entire volume of the aeration tank.\textsuperscript{51,52}

Problems encountered in maintaining the stability of operation and performance in existing activated sludge plants, under dynamic conditions in the nature of sewage to be treated, required urgent corrective measures without a full understanding of the problems observed. Therefore, extensive efforts devoted to plant improvements and modifications had to rely on experience and good judgement; they were mostly implemented on a trial and error basis. Nevertheless, some of the new configurations defined as process modifications, not only solved some of the major problems, but also persisted as original concepts for the activated sludge treatment practice today. Evaluation of the empirical developments achieved should definitely start with the excellent overview by Chase,\textsuperscript{53} providing not only a short historical coverage but also the fundamentals of the process as conceived in 1944.
Systems with improved aeration

Studies directed towards problems observed in the operation of activated plants indicated significant decrease in the oxygen utilisation as a function of aeration time. In experiments conducted with non-nitrifying sludge, it was observed that 40–50% of the total oxygen demand was utilised by the end of the first 2 h, 28–31% in the second 2 h, and 20–29% in the last 2 h, within a 6 h aeration period. Accordingly, tapered aeration was developed as a modified activated sludge configuration. It involved an aeration scheme which varied as a function of the anticipated oxygen demand in the process; in a diffused air system, for instance, the number of diffusers at the head end of the aeration tank would be higher than those near the outlet. This new scheme was soon incorporated into a number of major plants experiencing aeration problems.

Activated aeration emerged as another process modification, using the idea that organic matter removal potential of excess sludge generated in one of the units, which would ordinarily be stored under aeration at higher concentrations toward the inlet and the solids concentration of the tank effluent would be reduced. The performance of step aeration, where implemented, was satisfactory, mainly due to reductions in volume of the aeration tank. A survey of activated sludge plants in New York City revealed that the step aeration scheme provided a 50% reduction in the aeration volume compared with conventional systems, without an increase in air requirements or deterioration in the effluent quality.

Modified aeration

Modified aeration was promoted as a new process configuration which was operated with aeration tanks having a hydraulic retention time of around 2 h with 10% sludge return based on average sewage flow. Gould argued that in modified aeration:

the active sludge grows in the presence of more food material than it can assimilate and it must be removed from the system before it has time to acquire poor settling qualities.

This paper is also quite interesting for detailed accounts of a number of treatment plants implementing process modifications. The principles of the modified aeration were applied for the first time at the Jamaica treatment plant in New York, in 1943. The plant treated a flow of 3800 m$^3$ day$^{-1}$ with only 17% sludge return, which maintained average suspended solids concentration of 375 mg L$^{-1}$ in the aeration tank; the 5-day value for the biochemical oxygen demand (BOD$_5$) and removal of suspended solids in the range of 75–85% could be achieved with an air supply of only 2.5 m$^3$ m$^{-3}$ of sewage. Similar experience with modified aeration was also reported for a number of treatment plants in New York, Miami and Philadelphia.

Activated aeration

Activated aeration emerged as another process modification, using the idea that organic matter removal potential of excess sludge could be utilised: This new operation scheme required an activated sludge plant where the existing aeration tanks could be separated and operated as two independent units. The excess sludge generated in one of the units, which would ordinarily be wasted to sludge thickening tanks, was diverted into the other unit called ‘activated aeration’ sustained without recycling its own settled sludge. Part of the treatment facilities at the Wards Island plant in New York, was converted into activated aeration system, which was able to achieve 75% BOD reduction, with an air supply of 1.5 m$^3$ m$^{-2}$ of sewage and an average solids concentration of at only 225 mg L$^{-1}$ in the aeration tank. Despite favourable results, this process modification was not widely applied in locations other than the New York City.

Reduction of aeration volume

Two directions were essentially followed in attempts to reduce the aeration volume. The first direction focused on lowering aeration, i.e. reducing the air supply, at the expense of a moderate performance, as implemented in modified aeration or activated aeration systems. The totally opposite direction, however, was to increase the rate of air feeding per unit volume of the reactor by means of high-capacity aeration devices so that the conventional aeration period could be drastically reduced. Experiments indicated that it was possible to achieve a substantial increase in the sewage loads to be treated with a high oxygenation capacity, which would...
improve oxygen diffusion into smaller flocs.65 This was the basis of high-rate activated sludge, basically characterised by short aeration periods of 1–3 h, high solids concentrations in the range of 2800–3900 mg L⁻¹ and high return sludge rate of around 50% of the incoming sewage flow. This process scheme was observed to be as stable as conventionally operated plants, while achieving BOD reductions of at least 80–85%.66 Benedek67 used the term ‘high performance’ for systems where the efficiency of the conventional process is obtained in significant shortened periods; he listed characteristic operational data of 12 high performance activated sludge plants with aeration times between 0.4 and 2.16 h. The process was rapidly marketed under such names as ‘rapid bloc’, ‘aero accelerator’ or ‘supra activation’.68

Extended aeration
Extending the aeration period was a typical example of the modifications of the activated sludge process, which was explored through local experience and found beneficial under diverse conditions. It was developed to reduce or eliminate the problem of excess sludge handling, while securing a highly stabilised effluent. Extended aeration mainly intended to avoid separate aerobic digestion by decreasing the organic fraction of the activated sludge maintained in the system. The stabilisation provided not only oxidation of sewage organics, but also auto-oxidation of the microbial community, so that sludge produced would mainly contain inert organic residues.69 The initial idea was to provide a hydraulic retention time of 5 days with no sludge wasting,62 but the aeration time was reduced to 18–48 h in practice.

Extended aeration was first implemented in 1947, to modify and improve an under-loaded conventional activated sludge plant in East Palestine, Ohio; the plant was operated with 100% return sludge without sludge wasting and excellent effluent quality was obtained.70,71 Its popularity rapidly increased, especially for small sewage flows, because it simplified technology, reduced operation requirements and increased system stability against shock loads. Observations also indicated, however, that the quality of the effluent was susceptible to significant fluctuations because of sludge accumulation and rising problems in the settling tanks. Benedek67 argued that simultaneous sewage treatment and sludge oxidation taking place in the same unit often resulted in over-dimensioning of the plant, which could be avoided by separate sludge stabilisation; this way ‘oxidative sludge treatment’ would be more economical.

Sludge re-aeration
Sludge re-aeration was originally defined as ‘the continued aeration of sludge after its initial aeration in the activated sludge process’.72 Its practice dates back to the early plants; in fact, the first full-scale activated sludge plant installed in the US included facilities for re-aeration. In initial systems, less than 10% of the total aeration volume was generally allotted to sludge re-aeration (Table 2). Despite wide application, the benefit of re-aeration remained debatable; it was generally felt that while it seemed to bring a saving in the total aeration volume, it increased the complexity of the process and required a higher level of plant supervision. In 1938, despite the increase in the number of plants involving sludge re-aeration, Mohlman was quoted to state that ‘no convincing evidence has ever been presented to show that re-aeration is either more effective than an equal tank capacity devoted to straight aeration’.72 Similarly, Ruchhoft and Smith73 found that prolonged re-aeration altered the characteristics of activated sludge to a point where it entirely lost its ability to oxidise sewage. In spite of common unfavourable opinion, several researchers suggested process modifications that included sludge re-aeration in a different context from previous practice. They were successful in operating existing plants far above the original design load by devoting a part of mixed liquor aeration to separate sludge re-aeration.74

<table>
<thead>
<tr>
<th>Plant Year</th>
<th>Re-aeration</th>
</tr>
</thead>
<tbody>
<tr>
<td>North and South Plants, Houston, Texas 1917</td>
<td>5–33</td>
</tr>
<tr>
<td>Des Plaines, Chicago 1923</td>
<td>10</td>
</tr>
<tr>
<td>Passadena, California 1924</td>
<td>10</td>
</tr>
<tr>
<td>Mamaroneck, New York 1925</td>
<td>4</td>
</tr>
<tr>
<td>Pomona, California 1926</td>
<td>3</td>
</tr>
<tr>
<td>Waco, Texas 1927</td>
<td>6</td>
</tr>
<tr>
<td>San Antonio, Texas 1935–1936</td>
<td>7.5</td>
</tr>
<tr>
<td>Topeka, Kansas 1935–1936</td>
<td>14</td>
</tr>
</tbody>
</table>

*Results for re-aeration are given as % of the total aeration volume.

The Hatfield process
The Hatfield process was implemented in about 1942 at Decatur, Illinois, in an effort to improve the performance of the activated sludge plant seriously affected by chronic bulking problems. Accordingly, one third of the existing aeration tank volume was converted to provide 10–14 h re-aeration of the return sludge. The plant was also modified to add supernatant or sludge from the digester into the re-aeration tank, either continuously or intermittently.75

The Kraus process
The Kraus process was developed under similar circumstances at Peoria, Illinois, in 1946, which mainly consisted of re-aerating the activated sludge—digested sludge mixture until an acceptably low sludge volume index was obtained.76 For this purpose, provisions were made for re-aerating about 25% of the total aeration volume; the return sludge flow was split and a minor fraction of 10–20% was diverted to the re-aeration tanks. As in the Hatfield process, digested supernatant or digested sludge was introduced at a low but continuous rate into the re-aeration tank.77

Although resulting from trial and error modifications in the field and ‘not from the research laboratory’,78 both the Hatfield system and the Kraus system were quite successful in solving the specific problems prior to their application. Observed improvements were explained on the basis of the acquired ability of the modified process to sustain a much higher activated sludge mass as compared with existing conventional plants and also, the improvement of sludge conditioning by the ammonia addition contained in the digester into the re-aeration tank.76,77

Contact stabilisation
Contact stabilisation was another significant modification of activated sludge also relying on sludge re-aeration. It differed significantly from other re-aeration processes in that only a small fraction of volume was utilised for direct aeration of sewage. Basically, it involved two aerated reactors separated by a settling tank: Either raw or settled sewage was aerated in the contact tank for a short period of time, usually less than 1 h; afterwards, the settled activated sludge was directed first to a stabilisation tank and
re-aerated for several hours prior to return to the contact basin. The flow scheme was structured by Ullrich and Smith based on three mechanisms believed to occur during the removal of organic matter from sewage: (1) sorption of the majority of suspended solids and dissolved organic matter during the contact period, (2) improvement of settling characteristics by the sorption of sewage solids on the activated sludge, and (3) ‘restoration’ of the sorption capacity of the settled sludge in the re-aeration tank so that it can continuously be recycled into the contact tank. This explains why the system was initially called ‘the biosorption process’. This new configuration was first tried at Austin, Texas, in 1954; the modified flow scheme could treat a much higher sewage flow with the existing treatment volume and improve effluent quality by ending previously experienced chronic bulking conditions. In the following years, successful application of contact stabilisation was reported in many treatment plants in the US, treating both sewage and industrial wastewater. Table 3 summarises operating characteristics of the contact stabilisation process together with other activated sludge configurations.

**Table 3. Operating characteristics of different activated sludge configurations**

<table>
<thead>
<tr>
<th>Process configuration</th>
<th>Hydraulic residence time, ( t_h ) (h)</th>
<th>Mixed liquor solids, ( (\text{mg SS L}^{-1}) )</th>
<th>Volumetric loading, ( L_{x^*} (\text{kg BOD}_5 \text{ m}^{-3} \text{ day}^{-1}) )</th>
<th>Air supply, ( \theta_x ) <strong>(days)</strong></th>
<th>Sludge age, ( \theta_s ) <strong>(days)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional AS</td>
<td>6–9</td>
<td>2000–4000</td>
<td>0.7–1.0</td>
<td>6–9</td>
<td>6–8</td>
</tr>
<tr>
<td>Step aeration</td>
<td>1.5–4</td>
<td>1000–2000</td>
<td>1.5–3.5</td>
<td>3–4</td>
<td>4–6</td>
</tr>
<tr>
<td>Modified aeration</td>
<td>With primary settling</td>
<td>1.5–3</td>
<td>150–500</td>
<td>2.0–4</td>
<td>2–4</td>
</tr>
<tr>
<td></td>
<td>Without primary settling</td>
<td>1.7–2.1</td>
<td>300–1000</td>
<td>3.0–3.6</td>
<td>2–4</td>
</tr>
<tr>
<td>High rate AS</td>
<td>1–3</td>
<td>3000–5000</td>
<td>2.0–6.0</td>
<td>8–15</td>
<td>2–4</td>
</tr>
<tr>
<td>Extended aeration</td>
<td>18–48</td>
<td>4000–5000</td>
<td>0.1–0.4</td>
<td>10–15</td>
<td>20–25</td>
</tr>
<tr>
<td>Hatfield &amp; Krauss processes</td>
<td>6–9</td>
<td>5000–8000</td>
<td>0.7–1.0</td>
<td>6–9</td>
<td>10–15</td>
</tr>
<tr>
<td>Contact stabilisation</td>
<td>Contact tank</td>
<td>0.5–2</td>
<td>2000–4000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Stabilisation tank</td>
<td>2–6</td>
<td>6000–10 000</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Based on average \( \text{BOD}_5 \) of 250 mg L\(^{-1}\).**

**Estimated from existing data.**

\( \text{BOD}_5 \) is the 5-day value of the test for biochemical oxygen demand.

AS, activated sludge; SS, suspended solids.

Two-stage activated sludge

Babitt and Bauman also suggested a different conceptual approach for what they called a ‘stage aeration’ system, based on the stipulation that processes in the aeration tank occur in two stages: (1) flocculation and clarification taking place within the first hour of aeration, and (2) nitrification, starting toward the end of this period and tapering off after many hours of aeration. Accordingly a flow scheme with two stages was proposed, incorporating an aeration tank for the first stage and a tricking filter for the second.

**EXPLORING AND UNDERSTANDING THE PROCESS**

Although rapidly recognised and implemented after its discovery, the actual mechanism of the activated process was neither understood nor properly clarified in the following years. It was generally accepted that physical and biological mechanisms might be responsible to a different extent for the purification of sewage, although attempts to identify their existence or their relative importance remained largely inconclusive.

In 1965, McKinney argued that research and development played an important role in the acceptance and improvement of the process; he made the following statement: ‘one of the oddities in biological waste treatment is that there is very little use of the research data by the practicing engineering profession’, while still expecting that significant clarification in basic concepts will occur in future periods. In fact, understanding activated sludge has been a slow, exhaustive uphill process. Climbing each step
Mechanism of organic matter removal

The work of Buswell and Long^88 should be recognised as a monumental step for the proper interpretation of organic matter removal by activated sludge. On the basis of their observations, they suggested that purification of sewage was accomplished by ‘ingestion and accumulation’ of organic matter for synthesis into viable microbial fraction of the floc. Essentially, biological action was regarded as the process responsible for changing colloidal and soluble organics to a state that would settle out.

The biological mechanism, while verified by subsequent observations and accepted as the ultimate oxidation step, remained long debatable in explaining the initial phases of the process. Rudolphs et al.^89 stated that observations presented no evidence to exclude either physical or biological factors, which may both be responsible for the rapid initial removal rates. Several investigators suggested the existence of other initial steps in the process.

Baly^90 postulated that the charge difference between colloids in sewage and the activated sludge floc would induce an electrostatic attraction and result in consequent clarification. Theriault and McNamee^91 suggested a ‘bi-zeolite theory’ which removed particulate organics with a series of exchange reactions with activated sludge. It was also stipulated that iron compounds played a dominant role in the oxidation process.^15 Rudolphs et al.^89 also referred to the the study of Lumb who conducted an experiment investigating the balance of solids removal by activated sludge under laboratory conditions; it was concluded that removal was basically a physical process since 100% of the suspended and colloidal matter removed was recovered as activated sludge. A number of studies advocated that the reduction of organic matter was mediated by enzymatic reactions.^92,93 Ingols^94 argued that the purification potential of the activated sludge depended upon the concentration of enzymes; the activity of these enzymes could be affected by the type of organic matter, amount of oxygen supplied, pH and temperature.

These speculations were soon abandoned in the absence of additional experimental support in favour of the ‘adsorption mechanism’ which, curiously enough, also lacking similar support, has persisted over the years. The first and perhaps ‘the most noticeable’ mechanism of the activated sludge process was coagulation and flocculation of the suspended and colloidal matters in sewage. According to O’Shaughnessy^95:

Sewage contains sludge-forming material not in solution and other material which is in true solution. Of the sludge-forming material, 78 to 80% is readily settled out. The settled sewage contains non-settling colloidal matter which may choke the trickling filter, but which may be flocculated very rapidly by the activated sludge. It also contains material which may be dealt with only slowly by the expensive activated sludge technique.

Metcalf and Eddy^96 defined the floc as a ‘sponge-like mass’, which would be formed by enveloping and entrapping colloidal matter and bacteria; it is fed upon this matter and restore again its power of adsorption, when introduced into the incoming sewage. Based on the information available in 1935, three stages of activated sludge activity were identified: namely, clarification of the incoming sewage, reactivation of the sludge, and nitrification of the ‘sewage and sludge’. The clarification of sewage was confined to the first 1–2 h of aeration; it was believed to be chiefly physical in nature and produced the bulk of the improvement in sewage treatment. Restoration of the clarification potential of sludge and nitrification could be accomplished with additional aeration in the same tank or in separate ‘re-activation or re-aeration’ tanks. Heukelekian and Littman^97 found that the uptake of suspended and colloidal solids was followed by the removal of soluble organic matter. Activated sludge was increased by the amount of particulate organics and by the synthesis of bacterial cell. Ammonia nitrogen was assimilated by the activated sludge but nitrification did not occur.

After the discovery of the biological nature of activated sludge activity, recognition of the role of dissolved oxygen should perhaps be considered as the second major landmark in the understanding of the process. In this context, the adsorption mechanism was seriously challenged by studies focusing on the utilisation of dissolved oxygen. The adsorption mechanism, if admissible, should involve activated sludge as an adsorbent, and this potential should not be altered in the absence of oxygen. However, exclusion of air/oxygen impaired the removal ability of activated sludge;^98 studies performed to visualise the influence of air, mechanical stirring and nitrogen feeding, indicated that the removal of organic matter substantially decreased when aeration was substituted by mechanical aerations, and it was completely
stopped with nitrogen gas. Similar studies also reported that initial organic matter removals were always accompanied by oxygen uptake. The oxygen utilisation rate was observed to reach the highest levels during the first 30–60 min of contact; afterwards, it gradually decreased to a constant low level.

Later, Gellman and Heukelekian noted that organic matter removal did not proceed parallel to the oxygen demand, which continued at a fairly high rate for some time after the substrate had been depleted. Gaudy and Engelbrecht reported a variable oxygen utilisation under growth conditions: The initial uptake was more rapid and exhibited an inflexion point corresponding to the point where maximum uptake occurred, i.e. for the organic matter loading used in the experiments, there was an initial period of rapidly increasing oxygen utilisation rate, followed by a pronounced decline.

The accumulated experimental findings provided clear indication that, if adsorption did indeed exist, it had to be accompanied by biochemical oxidation. McKee et al. stated that organic matter removal from sewage included two inter-related processes: (1) a rapid adsorption and/or flocculation step, and (2) a biological step consuming organic matter for energy and cellular synthesis, which is started simultaneously with the first step but proceeding more slowly. Heukelekian and Ingols, however, argued that activated sludge incorporated adequate hydrolytic capacity to make particulate matter readily available for metabolic activities, in view of the fact that the highest rate of adsorption/particulate matter removal coincided with the period of highest oxidation, i.e. oxygen utilisation.

In 1955, Sawyer summarised the general consensus, attributing the removal of organic pollutants primarily to bacterial activities: rapid removal could be related to ‘unrestricted’ growth of bacteria using a fraction of the available organic matter for energy. He recognised for the first time that activated sludge was a ‘mixed culture’, where microorganisms capable of utilising the available organic matter most efficiently, i.e. reproduce more rapidly, will predominate, providing that adequate sources of mineral elements are present in the culture. In the paper, it was further stated that as the food supply increased above the required energy level for endogenous respiration, organic matter would become available for growth; the growth rate will increase with food supply until the maximum growth rate establishes and becomes the limiting factor.

Biological stabilisation of organic matter, as conceived by McKinney in 1963, consisted of two processes, synthesis and ‘respiration’. These two processes were inter-related and could not be considered as separate functions, while they should be evaluated individually to better interpret the overall mechanism. The respiration was also defined as the oxidation of organic matter to CO₂ and water, in a fully aerobic environment containing all the necessary nutrients for the metabolic reactions. An important remark was also made on the composition of activated sludge: regardless of the chemical nature of the organic matter utilised, the end product will be cellular material with a constant structure. From a practical point of view, this means that activated sludge generated from sewage will be the same as that formed by different industrial wastewaters.

**Microbiology**

It is quite noteworthy that only 4 years after his discovery of what was going to become the most ingenious biological treatment process ever implemented, Ardern explained the purification action of activated sludge as ‘primarily and essentially a desolution effect brought about purely by physical causes; any bacterial or biological action is definitely ancillary’. Microbiological studies, however, played an essential role to prove otherwise. Johnson, in his pioneering study involving microscopic examination of activated sludge, expressed his opinion as follows: ‘It seems possible that zooglaea, assisted by other minute organisms chiefly of animal origin, may be responsible for the rapid purification’.

Another important finding on this issue was the significant difference observed between the microbiology of the activated sludge and raw sewage. In a experimental study conducted to assess the role of the microbiological features of activated sludge on sewage purification, Russel and Bartow concluded that organic matter stabilisation was due to a typical aerobic flora ‘which gains almost complete ascendancy’; they also reported a large and consistent reduction of the total number of bacteria in sewage. Similar experimental work on the development of activated sludge from raw sewage by means of aeration also indicated a progressive change from the microbiology of raw sewage; the base of the sludge was composed of a zooglaeal mass intermixed with filamentous bacteria and occasional zooglaea ramigera. It was observed that filamentous bacteria overwhelmingly predominated the sludge.

Buswell and Long attributed biological growth in the activated sludge to the germination of microorganisms whose spores are always present in sewage. Examination of activated sludge under the microscope revealed no adsorbed, precipitated or coagulated amorphous matter but only active-growing microorganisms including a wide range of bacteria with a variety of free swimming or attached protozoa. Based on available experimental data, they defined the activated sludge floc as ‘a gelatinous matrix similar to that of Nostoc or Merismopedia, in which filamentous and unicellular bacteria are imbedded and on which various protozoa crawl and feed’.

Agersborg and Hatfield carried out an extensive study on the characterisation of microbial community in the activated sludge process and reported the following: (1) during pre-aeration, the sludge consisted entirely of Zooglaea and Sphaeritolus; (2) increased aeration favoured the succession of small flagellates and peritrichous ciliates; (3) the sludge was relatively free of protozoa.

Buswell referred to two groups of reactions taking place during the removal of organic matter from sewage, namely bio-precipitation and oxidation; he stated that bio-precipitation was accomplished by the Chlamyobacteriaceae, the fungi, the larger protozoa and ‘certain higher animals’, representing the bulk of the growth in the floc. Partial oxidation could also be attributed to these organisms since they generate carbon dioxide.

Heukelekian and Littman studied carbon and nitrogen transformations in sewage and came to the conclusion that the zooglaeal bacterium, Zooglaea ramigera, was of primary importance in the activated sludge; nitrifying bacteria, protozoa and filamentous organisms had secondary roles. Later, detailed taxonomic studies indicated that in normal activated sludge with good settling properties, the microbial flora mainly included members of the families Achromobacteraceae, Pseudomonadaceae and Corynebacteriaceae.

In 1956, McKinney provided a summary of the general consensus on the microbiology of activated sludge as follows: activated sludge was formed as a direct result of slime excretion by the zooglaeal bacteria and that these organisms were required for the best operation of the aerobic biological treatment processes. Microscopic examinations of the non-flocculating bacteria showed that all were extremely motile, indicating that active
metabolism of substrate was still in progress. Bacteria, protozoa, inorganic/organic colloids all contributed to form the activated sludge floc. Sludge wastage controlled the active/inactive ratio in the floc. High wastage resulted in a very active floc with small inactive portion; wasting only a small fraction of the total floc would result in a large inactive floc fraction with a high adsorptive capacity acting as a check valve on the final equilibrium between the substrate and the active floc.

These observations, although informative, did not contribute much to the understanding or the operation of activated sludge plants. Since good settling sludge was identified as the key issue in the satisfactory performance of the process, studies were more focused more on microbiological characteristics and changes likely to cause bulking problems.

Based on microscopic studies on activated sludge sustained in plants experiencing problems of poor settling and clarification, both Morgan and Beck and Ruchhoft and Watkins concluded that the observed bulking condition was due to the growth of *Sphaerotilus* stimulated by carbohydrates. In a similar study, the addition of 8 ppm of dextrose in the sewage encouraged the abundant growth of *Sphaerotilus* in the aeration tanks. According to Haseltine, the settleability of sludge would deteriorate and bulking would occur, if oxidation lagged behind adsorption, although clarification would still be excellent.

Later, Heukelckian and Ingols and Ingols and Heukelckian provided a detail account of the subject. They were able to produce bulking in a experiment by diluting the oxygen concentration in the air supply with nitrogen. With constant sewage load and air supply, bulking was intensified as solids concentration in the aeration tank increased. A highly oxidised activated sludge bulked less readily than one ‘poorly oxidised’ and the addition of 10–20 ppm of nitrate prevented bulking caused by oxygen deficiency. It was also observed that bulking occurred when activated sludge fed with sugar, calcium butyrate, peptone, glycerine or calcium propionate is agitated with limited amount of air. According to Lackey and Wattie, it was possible to produce experimental bulking of activated sludge by heavy doses of sugars or mixtures of nitrogenous and carbonaceous compounds. In such bulking systems, the predominant organism was the filamentous bacterium *Sphaerotilus natans*. Chlorine was the cheapest and most readily available toxic substance for controlling its growth. Littman observed that sludge containing *Sphaerotilus* did not grow well in sterile sewage devoid of carbohydrates.

The basic understanding of the activated microbiology in the late 1960s was best described by Jenkins as ‘all floc-forming bacteria became Zooglea ramigera and all filamentous organisms became Sphaerotilus natans’. An interesting exemption was proposed by Pasveer on the basis of studies conducted on oxidation ditch systems: The causative agent of bulking in the activated sludge examined was an unknown feature of *E. coli* which showed filamentous growth in the pH range of 5–6, very similar to that of *Sphaerotilus* under the microscope.

**Assessment of organic substrate and biomass**

In the early days of sewage treatment practice, it was believed that nuisance conditions in the receiving streams could only be avoided with highly purified effluents. Therefore, the major criterion for the efficiency of these plants was the degree of nitrification achieved, which would also ensure full oxidation of organic matter. During this period, wastewater was mostly characterised with measurements of different nitrogenous compounds, aside from physical parameters such as turbidity, colour and smell. Other inorganic constituents like potassium, magnesium were sometimes measured to assess the potential of sewage as a fertiliser. Specific analysis of important parameters including solids as well as dissolved oxygen, pH, conductivity, alkalinity, etc. basically started after the discovery of the activated sludge process. The main problem was the lack of standard methods for reliable characterisation. The experiments of Ardern and Lockett only included ‘free and saline ammonia’, ‘albuminoid ammonia’, nitrite and nitrate, aside from ‘4 hour oxygen absorption’, as illustrated in Fig. 1, which displays related historical data. Similarly, Fig. 2 shows a historical document on the characteristics of screened sewage treated at the experimental station at Mount Vernon, NY, for a period of around 5 months in 1921. The analyses covered the same parameters as that of Ardern and Lockett, together with turbidity, alkalinity and chlorides.

**Assessment of organic substrate**

The oxygen consumption measurements reported in early analysis presumably relied upon the measurement of potassium permanganate consumption for ‘partial’ oxidation of organic compounds, a method available since 1846. Soon after the implementation of the activated sludge technology, it was improved to what is generally known as the biochemical oxygen demand (BOD) and the 5-day value of the test (BOD₅) has been traditionally accepted and utilised as a collective substrate parameter for characterising the great variety of different organics in sewage and industrial wastewaters. While extensively used during the first 50 years of research and practice on activated sludge, the merit of BOD₅ always remained highly debatable. The confusion about BOD was perhaps best described by Hoover et al. in 1953:

The BOD test is paradoxical. It has been the basis of all regulatory agencies, and is used routinely in almost all control and research studies on sewage and industrial waste treatment. It has been the subject of a tremendous amount of research, yet no one appears to consider it adequately understood or well adapted to his own work.

In essence, BOD₅ values could not do better than reflecting a variable portion of the available organic matter depending on poorly defined biochemical reactions during the test, i.e. in the BOD bottle. However, accurate calculations of excess sludge generation and oxygen utilisation require the full amount of organic matter removed in the system. This requirement was satisfied with the test for chemical oxygen demand (COD), which oxidises all organic matter by potassium dichromate in boiling concentrated sulfuric acid and in the presence of a silver catalyst. The strong oxidising environment in the COD test does not oxidise ammonia. This makes COD an acceptable substrate parameter, reflecting only the carbonaceous organic matter in biochemical reactions. This way, COD, unlike BOD, was introduced as a direct parameter of organic substrate for the stoichiometric calculations of excess sludge generation and oxygen utilisation in studies reported in the early sixties. Later, it served as the basic substrate parameter in the works of Pearson and Jenkins and Garrison, which established the fundamentals of process kinetics and stoichiometry of the activated sludge process based on mass balance. A major drawback of the COD parameter is that it includes, aside from biodegradable organics, all non-biodegradable fractions also oxidised in the test. It was argued that soluble COD remaining in the effluent after activated sludge treatment could be used as
an approximation of the residual soluble COD initially present in the wastewater. While this assumption may be acceptable, it certainly does not cover residual particulate matter and the release of soluble residual metabolic products (soluble microbial products, S MPs) which would interfere with the calculations related to substrate removal.129

Assessment of biomass
Initially, the parameter ‘total suspended solids (TSS)’ was adopted as a measure of ‘sludge’ in activated sludge systems. Jenkins30 reminded that the allusion of Ardern and Lockett to the importance of the deposited solids was probably the first reference to TSS as a measure of biomass. In fact, TSS has been a convenient parameter for the plant operation in determining the quantity of sludge to be treated and disposed. It also assumed quite a useful role in changing the volumetric loading parameters into the organic loading rate, i.e. kg BOD5 kg−1 TSS day−1, for system design. The parameter ‘volatile suspended solids (VSS)’ was later adopted as a more refined biomass parameter, mainly to exclude the inorganic matter content of the activated sludge. Assessing biomass by means of VSS analysis survived through the years, despite the fact that it is also a crude measurement covering, aside from active microorganisms, non-viable particulate matter either of influent origin or generated as residual particulate metabolic products during the course of biochemical reactions.

The weakness of VSS in measuring only the concentration of activated sludge and not its activity has stimulated many studies to explore new parameters that would define active biomass130–132. The dehydrogenase activity as a measure of general biological activity was first proposed by Lenhard et al.130 for soil and bottom sediments. Bucksteeg131 proposed the use of the same parameter for the evaluation of wastes and sludges; they showed a workable correspondence between dehydrogenase activity of sludges and counts of viable bacteria. The experiments led them to conclude that the dehydrogenase activity would be a valuable parameter as a routine check of activated sludge activity. However, aside from a few experimental studies, specific viability parameters could not find acceptance in research and practice of activated sludge.

This period was also marked by pioneering studies on the chemical composition of the activated sludge, mainly because they assumed an essential role in establishing the basic stoichiometry of the process. Ardern and Lockett27,28 reported in their experiments an activated sludge composition of 65% VSS, 4.6% nitrogen and 2.6% phosphorus. Bartow and Mohlman133 indicated that 75% VSS, 4.0% fat, 6.3% nitrogen and 1.44% phosphorus characterised biomass during their study at the University of Illinois in 1915. Later a slightly lower nitrogen content of 5.63% was obtained.15 A more detailed study conducted in 1940 on sewage solids, digested sludge and activated sludge yielded similar values of 6.2% and 2.5% for the nitrogen and phosphorus (expressed as PO4 3−) fractions of biomass.19 The discrepancies among values reported were largely due to differences in the characteristics of sewage and operation conditions of experiments. Helmers et al.133,134 provided data to show an average level of 8.0% nitrogen and 1.02% phosphorus in the activated sludges developed on various industrial wastes. Based on these values, an empirical formula of C118H170O51N17P was suggested for the average composition of activated sludge; this formula was simplified to C7H10NO3 when the phosphorus content was ignored. In another study conducted by Hoover and Porges135 on the biodegradation of skim milk, which eliminated impurities and interference associated with sewage, the nitrogen content of the generated activated sludge was obtained as 11.27%, along with an ash fraction of 8.6%, leading to a suggested empirical formula of C5H7NO2; although it did not include the phosphorus fraction, this formula found universal acceptance in the stoichiometric evaluation of substrate removal and endogenous respiration mechanisms.

Process kinetics
Microbial growth and substrate removal
The quest for a better understanding of the process was continued with studies suggesting mathematical expressions for the quantitative evaluation of organic matter removal. These expressions were largely derived on an empirical basis and included variety of coefficients of unknown nature which were difficult to determine;
they could not be evaluated or extrapolated for providing a general basis for design and operation.

There was also a great deal of inconsistency among the rate expressions proposed. The observed variance was partly the result of the adopted notations. Each study proposed its own nomenclature with no attempt to generate a uniform notation system so that different removal expressions could be evaluated on the same basis. The essence of the inconsistency, however, was related to basic assumptions and simplifications used in deriving a particular mathematical expression. As rightly pointed out by McKinney78 Failure to understand the basic biochemical relationships ruins any mathematical analysis, however rigorous the mathematics may be.

The major shortcomings commonly noticed in most efforts on the modelling of organic matter removal may be summarised as follows. (1) Although two different mechanisms were identified for the removal of suspended/colloidal and soluble fractions of organic matter, kinetic expressions were formulated using a single parameter, BOD, assumedly characterising the entire organic matter content. (2) Most expressions did not include biomass concentration as a model component, while defining a biochemical process. (3) The bacterial growth curve was adopted as the basis of some kinetic studies; undoubtedly the curve has nothing in common with a continuously operated activated sludge unit, sustained with the lowest possible substrate concentration in order to achieve the desired removal performance. (4) Attempts were made to use proposed rate equations for calculating the substrate concentration in the effluent without any consideration for reactor hydraulics. Consequently, these efforts yielded misleading results and generally failed to establish a rational foundation of the process in terms of microbial kinetics and materials balance.

The work of Monod136 should be recognised as a major breakthrough for the kinetic evaluation of the activated sludge process. Results of batch experiments conducted on different pure cultures fed with single and complex organic compounds were verified with an empirical rate expression for microbial growth. It is important to note that the rate expression introduced for the first time the concept of a rate limiting substrate, and suggested an hyperbolic relationship between the substrate concentration, S, and the concentration of microbial culture, X; it also defined a maximum specific growth rate, \( \mu_H \), for the culture, a half saturation coefficient, \( K_S \), indicating the substrate concentration that would sustain half the maximum specific growth rate:

\[
\frac{dX}{dt} = \mu_H = \frac{S}{K_S + S}
\]  

Later in a comprehensive review of the subject, Monod137 stated that quantitative aspects of bacterial growth would also set the basis for understanding bacterial physiology and biochemistry. This review also described major phases in the growth of a bacterial culture and defined the observed yield coefficient, \( Y_{OBS} \) – currently defined as the net heterotrophic yield, \( Y_{NH} \) – as follows:

\[
Y_{OBS} = \frac{\text{amount of bacterial substance formed}}{\text{amount of limiting substrate utilised}}
\]  

Garret and Sawyer138 were first to conceive that the organic matter removal in the activated sludge process could be interpreted by adopting the kinetics of microbial growth with a similar empirical approach suggested by Monod. Based on experiments conducted with a synthetic organic mixture of peptone and glucose, they concluded that the mathematical relation between the growth rate and soluble substrate, which was represented by BOD, \( S_B \), could be defined using a two phase-discontinuous function. This approach stipulates a constant growth rate equivalent to the maximum specific growth rate, \( \mu_H \), for high \( S_B \) levels:

\[
\frac{dX_{SS}}{dt} = \mu_H X_{SS}
\]  

where \( X_{SS} \) is the biomass concentration expressed in terms of total suspended solids (TSS). At low concentrations, the growth rate becomes directly proportional to the remaining soluble BOD concentration in solution:

\[
\frac{dX_{SS}}{dt} = kS_B
\]  

Apparently, \( k \) is an overall rate coefficient incorporating the heterotrophic yield coefficient, \( Y_H \), and the biomass concentration, \( X_{SS} \) in the mixed liquor:

\[
k = k_Y Y_H X_{SS}
\]  

where \( k_Y \) is the true rate coefficient. A \( \mu_H \) value of 4.8 day\(^{-1}\) (0.20 h\(^{-1}\)) was calculated for the synthetic mixture used in the experiments at 20 °C, changing between 1.9 day\(^{-1}\) and 7.2 day\(^{-1}\) (0.08 – 0.30 h\(^{-1}\)) within a temperature range of 10 – 30 °C. It should be noted that a substrate mixture with peptone as the main ingredient is extensively used today in many experimental studies on the modelling of process kinetics, mainly because it exhibits a similar COD fractionation and biodegradation characteristics with sewage.139,140

Since most wastewaters include a wide array of organics with different biodegradation characteristics, a gradually decreasing rate was generally observed when the organic matter content was monitored in terms of BOD. Fair et al.141 suggested the following expression accounting for the observed ‘retardation effect’ on BOD removal:

\[
\frac{dC_B}{dt} = \frac{k_s}{1 + mt} C_B
\]  

Later, Fair and Geyer142 related the decrease in the removal rate to the gradual reduction of the available substrate:

\[
\frac{dC_B}{dt} = k_3 \left( \frac{C_B}{C_{BO}} \right)^n C_B
\]  

where \( m \) and \( n \) are empirical coefficients.

This conceptual approach proposed by Garret and Sawyer served as the stepping stone for most similar studies conducted in the following period. In this context, it should be mentioned that a good survey of this period cannot be made without consulting proceedings of the two sets of events, namely Conferences on Biological Waste treatment sponsored by Manhattan College, and International Conferences on Advances in Water Pollution Research, which included almost all the major contributions in process kinetics as well as on other major aspects of the activated sludge process. The latter should also be recognised for the outstanding discussions which sometimes proved more valuable that the paper itself.

In a paper presented at the Manhattan College conference in 1955, Eckenfelder and Weston143 also confirmed that the maximum rate of organic matter removal would be associated with the log growth phase, characterised by a maximum specific growth...
rate as expressed in Eqn (3); they suggested, however, variable \( \mu_{B} \) values as a function of sludge loading, much lower than that reported for pure culture studies. They postulated a BOD removal mechanism which occurred concurrently with synthesis and followed ‘the log phase law of cell growth’.

According to McKinney,\(^{111} \) however, the initial food to microorganism ratio in conventional in conventional activated sludge systems remains below that for logarithmic growth and the oxygen utilisation curve reflects the declining growth and extended period of endogenous respiration. In high rate activated sludge plants the initial food to microorganism ratio becomes sufficient ‘to cause the microorganisms to grow unrestricted’.

The concept of a discontinued growth rate with different phases, inspired from the bacterial growth curve that would be observed in a batch system, was later pursued in many studies.\(^{144−147} \) McCabe\(^{145} \) included a third ‘auto-oxidation’ phase defining the endogenous decay process and expressed the declining growth phases. Based on experimental results, \( \mu_{B} \) values were calculated as \( 2.7 \text{ day}^{-1} (0.115 \text{ h}^{-1}) \) for peptone, \( 2.9 \text{ day}^{-1} (0.122 \text{ h}^{-1}) \) for glucose and \( 1.6 \text{ day}^{-1} (0.069 \text{ h}^{-1}) \) for a pharmaceutical wastewater.\(^{146,147} \) Similarly, corresponding \( C_{BT} \) values were reported as \( 170−526 \text{ mg L}^{-1} \) for glucose, \( 58−116 \text{ mg L}^{-1} \) for peptone and \( 980−1580 \text{ mg L}^{-1} \) for the pharmaceutical wastewater, all significantly higher than half the initial BOD\(_{5} \) levels at the start of the experiments.

It is now quite difficult to understand why kinetic evaluations insisted on a discontinued growth rate approach in view of the fact that activated sludge plants have to reduce the organic matter in the effluent much lower than the reported \( C_{BT} \) levels. In fact, Tench and Morton\(^{148} \) confirmed the validity of a Monod-type rate expression for several organic compounds and for settled sewage. Their equation was based on the assumption that enzyme kinetics would also be applicable to describe the substrate removal mechanism by activated sludge. Today, an enzyme analogy is similarly adapted to interpret the inhibitory impact of different chemicals on the metabolic activity of microbial community.\(^{149} \) Wilson\(^{150} \) also recognised the validity of Monod’s equation in describing the kinetics of microbial growth in activated sludge and provided a comparative evaluation with the two-phase growth mechanism. Experiments conducted using respirometric analysis revealed no direct correlation between \( C_{BT} \) and the half saturation coefficient, \( K_{S} \), in the Monod expression. The lowest \( K_{S} \) value was determined as \( 1.4 \text{ mg BOD}_{5} \text{ L}^{-1} \) for acetic acid and the highest \( K_{S} \) value of \( 187 \text{ mg BOD}_{5} \text{ L}^{-1} \) was associated with peptone. Values of the maximum specific growth rate ranged between \( 0.5 \text{ day}^{-1} (0.023 \text{ h}^{-1}) \) and \( 2.9 \text{ day}^{-1} (0.122 \text{ h}^{-1}) \) for various organic compounds. It was concluded that the experimental data offered no significant support for the two-phase mechanism. McGauhey\(^{151} \) confirmed Wilson’s conclusive remarks on the grounds that the \( K_{S} \) values of 2.0, 4.0 and 20 mg L\(^{-1} \) reported by Monod for lactose, glucose and mannitol were significantly smaller than \( 1/C_{BT} \). Tench\(^{152} \) conducted a series of biodegradation tests using oxygen uptake rates for different simple organics and settled sewage. Evaluation of results with the linearised Lineweaver–Burk plots indicated no match with the two-phase discontinuous approach as illustrated in Fig. 3, showing that the Monod equation was a better representation of respective growth rates. Evaluations for settled sewage from three different treatment plants yielded relatively high \( K_{S} \) values in the range of \( 171−189 \text{ mg BOD}_{5} \text{ L}^{-1} \).

The high \( K_{S} \) values obtained when BOD\(_{5} \) is used to characterise the organic content of sewage led Eckenfelder\(^{147} \) to formulate the substrate removal mechanism according to first-order kinetic both with respect to substrate and biomass:

\[
\frac{dC_{B}}{dt} = -k_{S}X_{SS}C_{B} \tag{9}
\]

A similar mechanism was also adopted by Eckhoff and Jenkins\(^{153} \) suggesting a pseudo-first-order expression for the removal of soluble COD, \( S_{t} \), at concentrations commonly encountered in conventional activated sludge systems. This approach is also in good agreement with the Monod equation, which may be modified to a pseudo first-order rate expression when \( K_{S} > > S \):

\[
\frac{dS_{t}}{dt} = \frac{\mu_{B}}{Y_{B}K_{S}} X_{SS}S_{t} \tag{10}
\]
The significant discrepancy between the reported \( K_s \) values for simple substrates and sewage provided the basis of zero-order sequential removals proposed in a number of studies for single substrates with pure and mixed cultures.\(^{154-156}\) Wuhrman\(^{157}\) was first to suggest that nonlinear removal associated with complex wastes are actually a summation of a number of simple compounds. Later, Tischler and Eckenfelder\(^{158}\) also reported zero-order removals of sequential or simultaneous feedings of lactose, phenol and aniline to a biomass acclimated to the mixture, whereas the removal trend of the mixture as assessed by COD measurements was not linear. They claimed that the resulting removal rate of the complex substrate mixture would not necessarily be first order; any type of a relationship could be obtained depending on the removal characteristics of individual compounds in the mixture. Chudoba\(^{159}\) expressed zero-order kinetics as follows:

\[
\frac{dC_r}{dt} = -k_XSS
\]

(11)

where \( C_r \) is the COD concentration and \( k_s \) is defined as the specific rate of COD removal (mg COD mg\(^{-1}\) MLSS h\(^{-1}\)) (where MLSS is mixed liquor suspended solids). The \( k_s \) value for a simple substrate such as glucose, aniline, etc., while supposedly a constant in the rate expression, was reported to vary as a function of the biomass concentration, \( X_{SS} \), both in the same study and also in different studies. The observed variations were attributed to differences in the activity of activated sludge in batch tests.

It is important to note that various mathematical models summarised above included empirical rate coefficients – each assigned a different number in this review – with no relation with the fundamentals of the microbial mechanisms they represent. In this context, despite debatable arguments on the validity of different models, the Monod expression, verified by all subsequent observations, has found universal acceptance and remained virtually unchallenged in describing process kinetics for sewage as well as industrial wastes. Kinetic experiments on phenolic wastes reported by Graham\(^{160}\) are interesting because a linearised version of the Monod equation was used for the assessment of corresponding \( \mu_s \) and \( K_s \) values. A set of five different experiments were characterised with \( K_s \) of 21–95 mg L\(^{-1}\) and \( \mu_s \) of 4.3–9.3 day\(^{-1}\) (0.18–0.39 h\(^{-1}\)).

**Endogenous respiration**

In most experiments involving extended aeration periods, oxygen consumption was observed to continue at a lower rate together with a decrease in the biomass concentration. This observation was generally interpreted as an endogenous respiration mechanism. The microbial basis of the mechanism has long remained debatable, mostly under the influence of the bacterial growth curve representing a batch system, despite the fact that all practical applications involved continuous flow activated sludge plants.

The initial explanation for endogenous metabolism offered by Lamanna and Mallette\(^{161}\) was the oxidation of various organics accumulated in the presence of exogenous food. Porges *et al.*\(^{162}\) chose to define the process as a slow and continuous oxygen uptake for ‘the utilisation of protoplasm’ in the absence external organic matter. Wilson and Harrison\(^{163}\) argued that the true endogenous phase was not reached until intracellular storage products were exhausted. According to Eckenfelder and Weston,\(^{145}\) endogenous respiration would occur under all conditions; sludge is destroyed when the organic loading is insufficient to support active growth. McKinney\(^{111}\) defined endogenous decay as a phase of ‘declining cell mass in which bacteria must draw their own food reserves’. McCabe\(^{145}\) associated the ‘auto-oxidation phase’ of the bacterial growth curve with endogenous decay with the following rate expression:

\[
\frac{dX_S}{dt} = -k_0XSS
\]

(12)

The paper by McWhorter and Heukelekian\(^{125}\) and the subsequent discussions presented at the International Conference on Advances in Water Pollution Research held in London in 1962 deserve specific emphasis for the understanding of the endogenous respiration mechanism at that period. The paper summarised an experimental study using batch systems where the aeration period was extended beyond that required for substrate removal. After the exhaustion of exogenous food, a rapid weight loss of around 25–30% in the activated sludge was observed during the first day, which was gradually decreased to reach 40% at the end of 25 days. Carbonaceous storage materials were initially oxidised and above 35% exertion of the theoretical oxygen uptake, nitrogen leached out from the cell in the form of ammonia.

The scientifically interesting feature of the study was the part that monitored and interpreted changes in the nitrogen content of the biomass. Cellular nitrogen remained constant at 8–9% in the presence of exogenous food, below the 12% level associated with the theoretical biomass composition.\(^{135,164}\) This was attributed to the storage of carbohydrates. Upon depletion of exogenous substrate, utilisation of stored material increased the nitrogen content to the stoichiometric level of 12%. Continued aeration induced ammonia leakage from biomass as a result of protein oxidation during the endogenous respiration phase. In this framework, the authors suggested for the starting point of endogenous respiration the depletion of exogenous substrate and utilisation of stored organics, marked with substantial decrease in the rate of oxygen utilisation. The suggested mechanism confirmed the metabolic stoichiometry in a total oxidation activated sludge system presented earlier by Kountz and Forney.\(^{165}\)

In discussing the paper, however, McKinney and Symons,\(^{166}\) defined endogenous respiration as a process that takes place simultaneously with microbial growth, as it is still conceived today. Their arguments may be regarded as the consensus on the subject. In aerobic systems, oxygen uptake is in direct proportion to the energy expended. Microorganisms also require energy for cellular maintenance along with microbial growth. Therefore, the observed oxygen uptake reflects the sum of oxygen consumed for energy requirements of synthesis and maintenance.

**Process stoichiometry**

In the activated sludge process, estimation of the magnitude of excess sludge produced and the amount of oxygen utilisation is quite significant for successful plant operation; accurate assessment is closely related to process stoichiometry and particularly to the concept of yield. In the early practice, both aeration requirements and sludge production were defined as a result of empirical guesswork derived from practical experience. Consequently, yield was basically conceived as an observed value highly variable as a function of selected operating conditions.\(^{105,135,167}\)

In this context, pioneering studies by Porges *et al.*\(^{168}\) deserve special recognition, not because they supplied the necessary information for system operation, but mainly because they were able to define process stoichiometry as a material balance between all reactants and products separately for both microbial growth...
and endogenous respiration. This way, the concept of true yield, $Y$, was introduced as the quantitative relationship between the important components, i.e., biomass generated and substrate utilised. The approach adopted C$_6$H$_5$NO$_2$ as the simplified empirical formula representing biomass among different alternatives suggested in many studies on the chemical composition of activated sludge.$^{133,135,164}$ Using CH$_3$O as the simplest expression of a carbohydrate substrate, the stoichiometry for the assimilation and biosynthesis reactions were established as follows:

$$\text{CH}_3\text{O} + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O} \quad (13)$$

$$5\text{CH}_3\text{O} + \text{NH}_3 \rightarrow \text{C}_6\text{H}_5\text{NO}_2 + 3\text{H}_2\text{O} \quad (14)$$

In the experiment conducted with dairy wastewaters (lactose), the manometric data showed that 37% of the theoretical amount of oxygen was used during assimilation; hence 3/8 of sugar (CH$_3$O) was oxidised, leading to the overall process stoichiometry given below:

$$8\text{CH}_3\text{O} + 3\text{O}_2 + \text{NH}_3 \rightarrow \text{C}_6\text{H}_5\text{NO}_2 + 3\text{CO}_2 + 6\text{H}_2\text{O} \quad (15)$$

Similar overall reactions were also derived for casein (C$_8$H$_{12}$N$_2$O$_3$) and for the lactose–casein mixture. The experiment repeatedly indicated a higher utilisation of oxygen during the assimilation phase, followed by a much slower oxygen uptake during the following endogenous phase, which was also defined as follows:

$$\text{C}_6\text{H}_5\text{NO}_2 + 5\text{O}_2 \rightarrow 5\text{CO}_2 + \text{NH}_3 + 2\text{H}_2\text{O} \quad (16)$$

These stoichiometric relations are particularly significant as they enable calculation of $Y$, as 0.45 mg biomass mg$^{-1}$ COD or 0.63 mg cell COD mg$^{-1}$ COD, accounting for the COD equivalent of biomass as $f_x = 1.42$ mg COD mg$^{-1}$ biomass. It should be noted that these values closely approximate the corresponding default values in current activated sludge models.$^{169−171}$

Eckenfelder and Weston$^{143}$ essentially suggested the same reactions in a more general form, associating organic matter oxidation with a ‘combustion-type of a reaction’ which will also oxidise nitrogen to nitrate, and sulfur to sulfate. It was argued that if sufficient time is allowed to convert all organic matter into biomass and oxidise it during endogenous decay, ‘BOD can be completely oxidised to CO$_2$ and water’. The ultimate oxygen demand was confirmed as 1.42 mg g$^{-1}$ VSS of activated sludge. This was higher than the value of 1.25 mg COD g$^{-1}$ dry weight of solids found for activated sludge from dairy waste oxidation, including the ash content of the sludge.$^{135}$

The study by McWhorter and Heukelekian$^{125}$ is quite interesting, particularly with discussions providing new insights on the possible magnitude of the observed yield in cases where storage occurs concurrently with growth and when nitrate is used instead of ammonia for the nitrogen requirement of biosynthesis. In a series of batch experiments using glucose as the organic carbon source and nitrate as the nitrogen source to prevent nitrification, the authors found an average value of 0.315 mg VSS mg$^{-1}$ COD for the observed yield; the variation of yield values was sufficiently small to conclude that that yield was independent of substrate and sludge seeding. This value was appreciably lower than that obtained in studies similarly investigating glucose removal: Ruchhoft et al.$^{172}$ reported yields in the range of 63–80% which would be reduced to 52–59% when expressed as mg VSS mg$^{-1}$ COD. In the same way, the 61% yield on glucose reported by Gaudy and Engelbrecht$^{102}$ corresponded to still a high value of 0.57 mg VSS mg$^{-1}$ COD. In discussing the paper, Pipes$^{173}$ pointed out that glucose removal may sometimes involve substrate storage, which may be approximated by the following stoichiometry:

$$4\text{C}_6\text{H}_12\text{O}_6 + 9\text{O}_2 + 2\text{NH}_3 \rightarrow 5\text{CH}_3\text{O} + 2\text{C}_6\text{H}_5\text{NO}_2 + 9\text{CO}_2 + 15\text{H}_2\text{O} \quad (17)$$

Storage included in this stoichiometry may partly be the reason why the yield values obtained by McWhorter and Heukelekian$^{125}$ remained lower than other reported yield levels. Pipes also emphasised the use of nitrate as the nitrogen source for microbial in the experiments. During the course of the experiments, nitrate had to be reduced to amino (organic) nitrogen to be part of the cellular material; this reduction would require energy that consumes part of the available substrate COD for this purpose. It was concluded that the growth yield would be expected to vary for different substrates as microorganisms are capable to assimilate some substrates more efficiently than others.

This issue was further discussed by McKinney and Symons$^{174}$ who first underlined that the basic stoichiometry for glucose utilisation with ammonia nitrogen indicates 33% of the theoretical oxygen uptake and 47% of the theoretical oxygen demand (COD) as dry weight of cell mass ($Y = 0.47$ mg VSS mg$^{-1}$ COD):

$$5\text{C}_6\text{H}_12\text{O}_6 + 4\text{NH}_3 + 10\text{O}_2 \rightarrow 4\text{C}_6\text{H}_5\text{NO}_2 + 10\text{CO}_2 + 22\text{H}_2\text{O} \quad (18)$$

in the case where nitrate was used as the nitrogen source, the above stoichiometry would be modified into:

$$5\text{C}_6\text{H}_12\text{O}_6 + 3.6\text{NO}_3 + 4.8\text{O}_2 \rightarrow 3.6\text{C}_6\text{H}_5\text{NO}_2 + 12\text{CO}_2 + 19.2\text{H}_2\text{O} \quad (19)$$

The use of nitrates reduces the initial oxygen requirements as well as the quantity of cell mass produced. Only 16% of the theoretical oxygen demand would be exerted and 42.4% of the COD would be converted into cellular mass. The lower yield essentially results from the energy loss for the reduction of nitrates. This issue still retains its scientific merit today for further elaboration, especially for activated systems performing simultaneous carbon and nitrogen removal.$^{175,176}$

Oxidation of ammonia: nitrification

A review on biological oxidation of ammonia is certainly beyond the scope of this paper since nitrification is a well-studied microbial process, which would also deserve a separate full review, but mainly for the fact that 50 years of activated sludge research and practice was mainly devoted to the removal of organic carbon; nitrification was generally regarded as an auxiliary process providing security for sewage purification. Observations on nitrification date back to the 19th century, before the discovery of activated as a sewage treatment process. In Martin’s book$^{18}$ The Activated Sludge Process, published in 1927, reference was made to the first study on sewage aeration by Angus Smith in 1882, who noted that: ‘with aeration, ammonia nitrogen was oxidized to nitrate nitrogen and the putrescible organic matter was also oxidized’. Hamlin$^8$ also mentioned the studies by Schloesing and Muntz in 1897 on bacterial nitrification.

Studies conducted on activated sludge at the University of Illinois in 1922 provided the first numerical clues about nitrification in relation to the aeration pattern.$^{19}$ Nitrification took place when the aeration was sustained at around 10 m$^3$ air m$^{-3}$ sewage
Figure 4. Historical concept of the nitrogen cycle. The original hand drawing by the authors is shown.

(1.5 cu ft/gallon) and reduction of nitrate was observed when the aeration rate was lowered to 7.5 m³ air m⁻³ sewage. It was concluded that the nitrification phase was not functional, with the significant implication that nitrification was not essential for the success – efficient performance – of the process. The following comment is particularly important as the first statement of substrate removal under anoxic conditions:

It is possible under some conditions to produce a clarification and reasonable stabilization of sewage with so little air that nitrate oxygen in the raw sewage is actually consumed by microorganisms of the sludge.

Buswell and Neave provided what should be considered as the most significant contribution on the subject covering many aspects of biological nitrogen transformations. It is interesting to note that the concept of process biochemistry defined in the study for activated sludge was limited to nitrogenous processes, including a historical interpretation of the nitrogen cycle. Figure 4 and Fig. 5 reproduce this cycle and the nitrogen transformations presumably triggered by activated sludge, mainly to show the evolution in the understanding of the biochemical reactions involved, as original hand drawings by the authors. Ammonification was defined as the decomposition of proteins down to amino acids, ammonia, organic acids and CO₂ as a result of microbial activity; 65 years later, the same concept was incorporated into the first multi-component activated sludge model (ASM1); recently, it was also studied in terms of its impact on the kinetics of nitrification as a rate-limiting step.

The nature of nitrification was confirmed as a two-step biological reaction, where ammonia was first converted to nitrite and then to nitrate by *Nitrosomonas* and *Nitrobacter* spp., respectively. The sensitivity of these microorganisms to ‘changes in acidity’ was highlighted with the possible beneficial buffer effect of carbonates. Similarly, denitrification was defined as reduction of nitrates and nitrites triggered by bacteria; ‘the presence of nitrates and easily oxidisable organic matter were essential’. Payne indicates that the term ‘denitrification’ was also used in 1882 to describe gas production from sewage by means of ‘nitrate-using bacteria under conditions devoid of dissolved oxygen’.

Assimilation of nitrate and ammonia to form insoluble bacterial protein was generally defined as nitrogen fixation. Buswell and Neave referred to two earlier studies in particular: In summarising an extensive review on the conservation of nitrogen with specific reference to the activated sludge process, Fowler argued that biomass could also retrieve nitrogen from air when he stated: ‘there is little doubt that not only does the activated sludge process recover the nitrogen present in the faecal matter of sewage but through fixation from air as an actual increase takes place over what can be recovered from sewage’. However, Richards and Sawyer observed no evidence for fixation of atmospheric nitrogen and concluded that: (1) if activated sludge is aerated for a short period in an ammoniacal solution ‘recovery of nitrogen is quantitative’; (2) if aeration continues, loss of oxygen occurs; (3) with sewage, ammonia decreases and sludge gains nitrogen with a loss of N on the whole balance after 16 days; (4) extra nitrogen in activated sludge ‘over and above old type sludges’ is derived from ammonia of the sewage. Concerning the loss of gaseous nitrogen, two alternative mechanisms were mentioned from purely chemical considerations: first, by direct oxidation of ammonia to N₂ and H₂O; and second, by the reduction of nitrates and nitrites by organic matter with the formation of nitrogen gas and CO₂. The concluding remarks of the authors are important for the way they reflect the understanding of nitrogen transformations in the early years of activated sludge practice: (1) an effluent of reasonable...
quality can be obtained without using sufficient air to produce nitrification; (2) denitrification results in protein formation rather than in loss of nitrogen; (3) there is apparently no loss of nitrogen when using a minimum amount of air for aeration; (4) there is no evidence of nitrogen fixation.

The concept of a reaction rate for nitrification was first conceived in 1940 by Sawyer,181 who stated that the loss of nitrifying power by activated sludge at low temperature depended upon the relative rates at which the microorganisms were capable of stabilising both carbonaceous and nitrogenous matter. The growth of nitrifiers was hindered at low temperatures by the slow rate of oxidation of ammonia nitrogen and by the inability of sludge to 'adsorb and hold available, appreciable quantities of ammonia nitrogen'. However, development of a parameter for separate growth power by activated sludge at low temperature depended upon the temperature and pH dependencies of the process. Observed relations between organic load and nitrification yielded an average growth constant (net growth rate) of 0.31 day\(^{-1}\) provided better match with the experimental results of British activated sludge plants:

\[
\mu_A = \frac{A}{K_{NH} + S_{NH}} \tag{20}
\]

where \(\mu_A\) and \(\hat{\mu}_A\) is the growth rate and maximum growth rate of nitrifiers; \(K_{NH}\) is the half saturation coefficient for ammonia, and \(S_{NH}\) is the ammonia nitrogen concentration in the carbon tank. Downing and Kowles186 derived empirical expressions for temperature and pH dependencies of the process. Observed relations between organic load and nitrification yielded an average growth constant (net growth rate) of 0.31 day\(^{-1}\) at 20 °C; they indicated that a growth constant of 0.4 day\(^{-1}\) provided better match with the experimental results of British activated sludge plants:

\[
\mu_{A,T} = 0.18 \exp [0.12 (T - 15)] \tag{21}
\]

\[
\mu_{A,NH} = \mu_A [1 - 0.83 (7.2 - pH)] \tag{22}
\]

The information available for ammonia oxidation could only be used together with the conceptual interpretation of denitrification in setting the basis of a process modification, which established itself as a nitrogen removal activated sludge process in the 1980s, using an appropriate sequence of anoxic and aerobic phases.187,188

Towards process design

The first 50 years of activated sludge has been a period of amazing discoveries on the various aspects of the microbial processes involved. However, very few of these research efforts, if any, were reflected in practice. System design evolved on a trial and error basis, mostly relying on the results of extensive pilot plant operations. Successful process performance was largely due to ‘to the ingenuity of the operators than to the skill of the designer’.74 A clear effluent could be generally obtained by selecting an aeration period of 4–8 h and 3.5–15 m\(^2\) air m\(^{-3}\) of sewage (0.5–2.0 cu ft of air/gallon). Aeration periods generally included sewage flow and return sludge, which was usually specified as 25% of the sewage flow. An interesting explanation was offered on the magnitude of return sludge by Metcalf and Eddy:8 the quantity of 98% of the sludge required for the treatment of moderately strong sewage would be equivalent to about 25% of the volume of sewage treated. For partial treatment of sewage, the volume of return sludge would be small. At Birmingham, England, 2–5% return sludge flow was found to be sufficient in the partial treatment of settled sewage containing 80 ppm of colloidal and finely divided solids.

Haseltine’s paper74 should be consulted as a good reference material on the subject because it outlines detailed operation characteristics of a total of 45 conventional and high rate activated sludge plants. The aeration period ranged between 3.3 and 9.3 h for conventional plants and 1.7 to 7.9 h for modified/high rate plants; similarly, the air supply varied in the range of 4–14 m\(^3\) air m\(^{-3}\) of sewage (0.48–1.90 cu ft/gallon) for conventional plants and in a narrower range of 2.5–5.0 m\(^3\) air m\(^{-3}\) (0.33–0.69 cu ft/gallon) of sewage for high rate plants.

Metcalf and Eddy8 argued in 1930 that the character of sewage would have a market influence upon the required period of contact. A dilute fresh sewage may become ‘well coagulated’ in 3–4 h, whereas a strong sewage containing industrial wastes may require 6–8 h. Greeley189 also recommended a hydraulic retention time of 6 h based on 125% of the average sewage flow while shorter retention times would also be satisfactory, referring to good clarification with 30–40% BOD reduction with 1.0 h aeration at Decatur, and 75% BOD reduction at Ward Islands with 2.0 h aeration.

In activated sludge systems involving re-aeration, an initial contact period of 60–90 min was found advisable in view of the fact that 20–40 min would be required for the completion of the adsorption phase. This offered a logical approach for the partition of the total aeration time between the contact tank and sludge re-aeration.

Organic matter loading was initially calculated on a volumetric basis, disregarding the level of biomass sustained in the aeration tank; the adopted value for plant operation was around 0.3–0.8 kg BOD\(_5\), m\(^{-3}\) (20–50 lb BOD\(_5\) per 1000 cu ft) of aeration volume.189,190 The concept of solids loading, better known as the ‘Gould sludge age’ has also been implemented as a design index. This parameter, calculated as the ratio of the total weight of suspended solids in the mixed liquor to the weight of suspended sewage solids introduced daily into the aeration tank, implies that biomass is basically an accumulation of particulate matter removed from the incoming sewage.191 It was recommended that a sludge age of 4 days be selected for a conventional process; a range of 0.2–0.4 days was suggested for high rate processes.192

The use of an overall loading parameter based on biomass was first made by Ruchhoff and Smith73 in 1939; they observed the change in the oxidation characteristics of the sludge in a small plant as a function of organic loading rates which varied in the range of 0.1–1.0 kg BOD kg\(^{-1}\) SS day\(^{-1}\). Later, Heukelekian et al.193 found that the rate of activated sludge accumulation was primarily a function of the BOD input to the process; this way, they confirmed the observations of Ruchhoff and Smith correlating the total biomass sustained in the plant with the daily BOD load imposed on the process. Experimental results indicated that the activated sludge process could be expected to yield about 90% BOD reduction as long as the daily BOD\(_5\)
load did not exceed 0.5 kg BOD₅ kg⁻¹ of suspended solids in the aeration tanks. Haseltine, after reviewing published operation data, suggested that moderate to large activated sludge plants be designed for 0.30–0.40 kg BOD₅ kg⁻¹ SS day⁻¹; a lower loading range of 0.2–0.30 kg BOD₅ kg⁻¹ SS day⁻¹ could be selected for smaller plants with little or no skilled operation.

A similar modification was also suggested for the air supply in a way to account for the amount of BOD removal in the suggested aeration rate. Haseltine also analysed the correlation between air consumption and organic loading and observed that the air consumption rate remained relatively constant as 30 m³ air kg⁻¹ BOD₅ (500 cu ft/lb) removed as long as the organic loading exceeded 0.35 kg BOD₅ kg⁻¹ SS day⁻¹; it increased rapidly however, up to 95 m³ air kg⁻¹ BOD₅ (1600 cu ft/lb) when the applied loading was below 25 kg BOD₅ kg⁻¹ SS day⁻¹. Greely suggested a more conservative value of 60 m³ of air per kg (1000 cu ft/lb) of applied BOD for the conventional activated sludge operation. Two observations still remain valid in evaluating the aeration process: In comparing volumes of air entering and gases escaping from the aeration tanks, Metcalf and Eddy estimated that only about 5% of the oxygen in the air introduced during the aeration process was consumed as an index for excess sludge generation: 193

It was usually observed to remain around 2000 mg SSL as higher than to increase the cost. 196

The general practice concerning the return sludge was summarised in terms of 0.25–1.0 for the volume ratio between return sludge and raw sewage, equivalent to 16.7–50% of the return sludge in the mixed liquor. The range of 25–35% was indicated as common practice.

The following formula, originally recommended by Kraus, indicated calculating the suspended solids concentration in the aeration tank, Xₜ, as a function of the return sludge ratio, R, and the sludge volume index (Mohlman index):

\[
X_t = \frac{R \times 10^{-6}}{SVI \times (1 + R)}
\]

It is interesting to note that Babitt and Bauman still mentioned the Gould sludge age as an important parameter in the design and operation of activated sludge plants. The empirical nature and the trial and error approach in process design mostly persisted during this period, until significant conceptual tools such as process stoichiometry and modelling could be made available for translating relevant microbial mechanisms into design practice during the early 1980s.

**MILESTONES FOR FUTURE PROGRESS**

**Fractionation of substrate and biomass**

Early experiments indicated the presence of a mechanism generating residual soluble organics. Most of the experiments used simple substrates; monitoring the specific substrate together with COD measurements revealed the amount of residual soluble organic matter. Experimental data also showed a direct relationship between the level of microbial products and influent COD. In a study of the oxidation of acetate at high initial concentrations, Gaffney and Heukelekian showed that a residual COD remained in filtered mixed liquor samples taken after complete depletion of oxygen consumption in the reactor; the level of the soluble residual COD was found to be around 10% of the COD in the feed stream. In a similar study, McWhorter and Heukelekian
observed complete COD removal at 30% of the theoretical oxygen uptake; however, a soluble residual COD remained in the supernatant representing 5–15% of the initial COD in all samples. Rao and Gaudy201 conducted batch experiments to show that soluble COD measurements in the mixed liquor did not correspond to the remaining glucose levels; the difference was attributed to the generation of soluble metabolic intermediates and products. The kinetic description for the release of soluble residual products was first provided by Eckhoff and Jenkins.202 They postulated a growth-associated mechanism where soluble residual COD could be expressed as a fraction of the influent COD concentration.

The most decisive experimental support for the residual soluble COD was provided by Chudoba203,204 in the experiments conducted with pure substrates under batch and continuous conditions. Tests started with biomass acclimated to a sludge age of 10 days and continued until the completion of COD removal. The reactors were further aerated for 2–4 h; the mixed liquor was filtered for measuring the remaining COD. The results indicated the presence of residual COD in the range of 1–4% of the initial substrate concentration. The lower range may be attributed to the fact that Chudoba’s experiments were stopped after the consumption of biodegradable substrate, therefore only characterising the growth phase whereas former experiments were extended to also cover endogenous metabolism, indicating that microbial decay is mainly responsible for the generation of microbial products.

The significance of particulate inert products was also recognised in many studies. Symons and McKinney164 conducted perhaps what must be considered as the pioneering study on the subject, noting that total oxidation of activated sludge was not possible and a small amount of non-oxidisable polysaccharide material remained. Based on experiments conducted with skim milk, Kountz and Forney165 also observed that 58% of substrate reappeared as new activated sludge and 23% by weight of biomass was non-oxidisable. In a similar study where the long-term carbon balance was monitored, Washington and Symons205 observed a VSS accumulation in the range of 11.5–14.5% of the ultimate BOD removed, corresponding to a ratio of 20–25% by weight of the activated sludge produced, adopting the same yield value of 0.58 g VSS g−1 ultimate BOD as suggested by Kountz and Forney.165 In a comprehensive work still maintaining its significance for the modelling of the activated sludge process as it is implemented today, McKinney206 defined an endogenous decay coefficient, bE, of 0.17 day−1 at 20 °C and concluded that a non-biodegradable VSS fraction would be generated at a rate of 0.036 day−1 during endogenous metabolism, yielding a value of 0.21 for the endogenous residue fraction fE. Later, the study by Washington and Hettling207 provided a similar experimental support for the generation of particulate residual products, where the VSS accumulation was calculated as 10.3% of organic matter removed, corresponding to a lower fE level of 0.15.

McKinney and Symons166 suggested a first-order rate expression for the generation particulate metabolic products in a way that still remains valid today; they basically modelled the rate of loss of total biomass, XVSS, as the sum of the decay rate of active biomass, XH, and the generation rate of the residual particulate matter, Xp:

\[
\frac{dX_{VSS}}{dt} = \frac{dX_{H}}{dt} - \frac{dX_{p}}{dt}
\]

(26)

where the generation rate for Xp was defined as:

\[
\frac{dX_{p}}{dt} = f_{E} \frac{dX_{H}}{dt}
\]

(27)

Substituting Eqn (27) into Eqn (26), the following equation was obtained:

\[
\frac{dX_{VSS}}{dt} = (1 - f_{E}) \frac{dX_{H}}{dt}
\]

(28)

the corresponding oxygen consumption rate was expressed as:

\[
\frac{dS_{O}}{dt} = f_{E} (1 - f_{E}) \frac{dX_{H}}{dt}
\]

(29)

where, S0 is the dissolved oxygen concentration and fE is the oxygen (COD) equivalent of biomass. The work by McKinney is also quite important for suggesting the hydrolysis mechanism in substrate removal.208 It was suggested that with complex organic matter, the bacteria must first break it down to simple molecules which can pass through the cell wall. Complex organic matter which must be broken down outside bacteria can be divided into two groups depending on the method of degradation: (1) those which can be hydrolysed, such as proteins, polysaccharides and fats; and (2) those which cannot be hydrolysed, such as long fatty acids, fatty alcohols, etc. Hydrolysis of complex organic matter has been shown to be brought about by extracellular enzymes. It was also argued that activated sludge systems tend to built up inert volatile solids to an equilibrium depending upon the solids wasting schedule and solids lost in the effluent. The addition of inert solids in the raw waste would also increase the build-up of solids.

Early studies have provided the scientific basis for defining a more accurate picture of substrate and biomass in the evaluation of the activated sludge process. Current activated sludge models now include soluble and particulate microbial products, S0, Xp, together with soluble and particulate inert COD coming from wastewaters, S, X as model components, aside from biodegradable COD fractions.209,210 Methods have been developed for detailed COD fractionation211 and the quantitative assessment of soluble and particulate residual fractions.212–215 Particular emphasis is now placed upon soluble microbial products (SMPs) as the crucially significant parameter for the performance of membrane bioreactors.216

**Particle size distribution**

The early interpretation of substrate removal involved different mechanisms for particulate and soluble components: Utilisation of particulate and colloidal organic matter first required their incorporation into the microbial floc by means of different physical and biological phenomena, whereas soluble matter could be directly oxidised for microbial growth. This interpretation certainly implied different sizes and biodegradation characteristics for organic matter fractions. Since substrate removal was loosely defined in terms of a first-order reaction as given by Eqn (4), different values for the rate coefficient, k, would apply for the sequential removal of substrate fractions differentiated with respect to size.

This issue, while quite significant, did not initially receive the attention it deserved; it was first tackled by Rudolfs and Gehm,217 in 1936, who analysed the BOD content of sewage in terms of ‘coarse, finely divided and soluble’ solids and indicated that each fraction had a different BOD to volatile matter ratio. Much later, Balmat218 conducted the major work on the subject in that period. Settleable solids were separated and the supernatant was further segregated into, supracolloidal, colloidal and dissolved solids with centrifugation under controlled conditions. As shown in Table 4 the rate coefficient, k, exhibited a significant variation between...
The concept of particle size was basically abandoned in favour of a single size threshold, i.e. the filtration size of 0.45 μm, which is still in use for the arbitrary differentiation of particulate fractions from ‘soluble’ components. This size threshold, quite different than that suggested by Balmat, is currently prescribed for defining suspended solids; it is also used as a basis for COD fractionation in current activated sludge models.\(^{209}\)

The structure of activated sludge model did not discourage research on the use of particle size information for a better understanding of biological processes.\(^{226–229}\) A procedure for particle size distribution (PSD) analysis was developed, involving a sequence of different sizes in the filtration and ultra-filtration ranges from 1600 nm (1.6 μm) down to 2 nm; this analysis enabled a specific fingerprint of the selected wastewater for different COD fractions to be obtained as well as other parameters such as total organic carbon (TOC), colour, etc.\(^{230,231}\) So far, the PSD analysis has been successfully implemented for the characterisation of a wide range of different wastewaters.\(^{232–234}\)

Physical size distribution of COD by means of sequential filtration/ultrafiltration with parallel oxygen uptake rate measurements is now integrated into modelling for a better interpretation of biodegradability. PSD measurements have a significant potential of complementing respirometric methods in better interpreting COD fractionation as related to biodegradation. This way, PSD provides the necessary information for selecting the optimum treatment technology, as it indicates which size fractions are significant within the overall organic matter content and how they relate to the biodegradation characteristics of the corresponding COD fractions. This information is of vital importance for membrane bioreactor applications.\(^{235,236}\)

### Substrate storage

Storage as an auxiliary biological mechanism for the removal of organic matter was first discovered in the 1950s. In a study conducted on seven different industrial wastewaters, Gelman and Heukelekian\(^ {101}\) observed that removal of organic matter occurred faster than the corresponding rates of oxidation. Eckenfelder and Weston\(^ {140}\) defined storage as COD removal from solution by activated sludge but not immediately metabolised.

Porges et al.\(^ {168}\) conducted experiments on the purification of skim milk by activated sludge. Almost 80% of the available substrate was removed from solution in 1 h; during the same time, the oxygen utilised corresponded to only 50% of the substrate removed. Storage increased to 60% in 2 h and it was reduced to 50% in the next 2 h. The authors concluded that storage occurred principally as glycogen since many microorganisms contained glycogen as a storage carbohydrate. Rapid removal of carbohydrate involved more than a simple adsorption mechanism; a portion was oxidised to CO\(_2\), another synthesised into new cells and the remainder was converted into and stored as insoluble ‘glycogen-like substance’. This way, the oxygen demand of the sludge would continue at a high rate while the stored carbohydrate was oxidised, decreasing afterwards to that of ‘true endogenous respiration’.

In a study conducted with glucose as the organic substrate, Nemerow and Ray\(^ {237}\) also indicated a rapid glucose removal which could not be accounted for by the corresponding CO\(_2\) production: Glucose must have been incompletely oxidised and stored as ‘some other compound’. They also provided an ingenious support to their observation. The ratio of CO\(_2\) production to COD reduction increased from 1.3/1.0 at the end of 3.0 h of aeration to 1.8/1.0 after 24 h; the complete oxidation of glucose to CO\(_2\) would give a ratio

<table>
<thead>
<tr>
<th>Sewage fraction</th>
<th>BOD rate constant, (k) (d(^{-1}))</th>
<th>(\text{BOD}_5) fractions (%), (k) (mg L(^{-1}))</th>
<th>(\text{BOD}_5/VSS) ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw sewage</td>
<td>0.22</td>
<td>315</td>
<td>0.85</td>
</tr>
<tr>
<td>Settleable</td>
<td>0.08</td>
<td>48</td>
<td>0.95</td>
</tr>
<tr>
<td>Non-settleable</td>
<td>–</td>
<td>269</td>
<td>0.85</td>
</tr>
<tr>
<td>Supracolloidal</td>
<td>0.09</td>
<td>89</td>
<td>0.97</td>
</tr>
<tr>
<td>Colloidal–soluble</td>
<td>–</td>
<td>117</td>
<td>0.84</td>
</tr>
<tr>
<td>Colloidal</td>
<td>0.22</td>
<td>49</td>
<td>1.26</td>
</tr>
<tr>
<td>Soluble</td>
<td>0.39</td>
<td>127</td>
<td>0.75</td>
</tr>
</tbody>
</table>

\(\text{BOD}_5\) is the 5-day value of the test for biochemical oxygen demand. VSS, volatile suspended solids.
of 1.33/1.0, while the complete oxidation of a storage material, such as glycogen, would give a ratio of 1.62/1.0. Consequently, the increased COD/CO₂ ratio obtained in these experiments as the aeration progressed indicated that ‘compounds richer in carbon were being oxidised’. Kountz and Forney165 offered a graphical explanation illustrating the role of storage in substrate removal (Fig. 6).

Gaudy and Engelbrecht102 continued to explore the storage mechanism in a study involving two parallel chemostats fed with glucose as the organic substrate; the second set was operated as a ‘respiring system’, due to lack of nitrogen in the feeding solution. The ‘growth’ unit was characterised by a rapid increase both in solids and carbohydrates peaking approximately after 4.0 h; the amount of protein continued to increase and eventually peaked approximately after 6.5 h. It would appear that ‘carbohydrate may be bled off into protein synthesis.’ In the ‘respiring’ unit, an increase in solids concentration was also observed, the same way as in the growth system. However, as the unit had been prevented from synthesising protein due to lack of nitrogen, the increase in biological solids was caused by gradual carbohydrate accumulation in the sludge, i.e. oxidative metabolism of the synthesised carbohydrate rather than that of protein. It was concluded that in the growth system, the increase in solids corresponded to an increase in protein, which was associated with cell growth, whereas in the respiring system it was attributable to carbohydrate formation. While the carbon continued to be channelled into polysaccharide formation under respiring conditions, it was used for the synthesis of protein and other cellular constituents under growth conditions. Symons and Mc Kinney164 also reported a slow build-up of extracellular polysaccharide when acetate was used as the sole external carbon source. Synthesis continued under both respiring and growth conditions; the difference lay merely in the nature of cellular component synthesised.

Although recognition of intracellular storage was definitely a milestone in the interpretation of substrate utilisation, this concept was not immediately introduced in modelling. All early multi-component activated sludge models (ASM1, ASM2, ASM2d) formulated after 1987 stipulated that substrate utilisation by the microbial community was solely related to microbial growth.170 However, extensive research on the subject continued, relating storage to a sequence of feast and famine periods created by transients in substrate feeding. During the feast period, excess substrate was converted into internal storage periods and in the following famine period with no external substrate available, storage polymers were used up for internal growth.238 When activated sludge is fed with glucose, the resulting polymer is glycogen as usually observed in the early studies and poly-hydroxy alkanoates (PHAs) are generated when the activated sludge is fed with volatile fatty acids.239,240 Simple fatty acids such as acetate mainly generate poly-β-hydroxy putyrate (PHB).241,242 ASM3 was the first model to incorporate the concept of storage.243 In ASM3, storage was conceived as the only substrate removal mechanism, an assumption not applicable in practice. In the light of experimental evidence, it was modified into a new version (ASMGS), which accounted for simultaneous direct growth and storage.244,245 This modelling approach suggested metabolic competition between growth and storage mechanisms as the main factor on the presence and relative magnitude of storage. Studies also demonstrated that storage would occur under continuous feeding conditions.246 Current modelling studies always make provision for evaluating the relative effect of storage on process kinetics and stoichiometry.247–249

**Energetics**

Introduction to the energetics of biochemical reactions has been a major breakthrough for better understanding and establishing the basis of many microbial mechanisms in the activated sludge practice. With the adoption of COD as the organic substrate parameter, it enabled the electron equivalence between substrate utilised, biomass generated and oxygen consumed to be established, as first suggested by McKinney206 in 1962; this provided a different perspective for the concept of yield, which could be defined as follows:

\[
Y_H = \frac{\text{electron potential of synthesised biomass}}{\text{total transferable electrons in organic substrate}}
\]  

This concept defines organic substrate as the electron donor in the energy and biosynthesis reactions and oxygen as the final electron acceptor. The amount of substrate equivalent to the electrons transferred to the final electron acceptor generates energy for biosynthesis; electrons in the remaining fraction of substrate match that in the synthesised biomass. This fraction basically defines the true yield, \(Y_H\), because it is not oxidised to CO₂ but essentially converted to the overall oxidation state of the biomass. The concept of true yield ended the confusion created by a multitude of different values for this parameter, because they merely

---

**Figure 6.** Stoichiometry of carbon metabolism for a total oxidation system.165 NOS, non-oxidisable sludge.
reflected the observed yield values, \( Y_{NH} \), all including the interference of endogenous respiration. The kinetic evaluation offered by Pearson\(^{127} \) and Jenkins and Garrison\(^{128} \) also enabled to define the observed yield or the net yield \( Y_{BH} \) as a coefficient variable as a function of the sludge age of the system:

\[
Y_{NH} = \frac{Y_{BH}}{1 + k_E \theta_X} \quad (31)
\]

The electron equivalence in biochemical reactions is illustrated in Fig. 7, expanding the approach by Porges et al.\(^{168} \) adopted for establishing process stoichiometry.\(^{176} \)

The free energy released by dissimilative (oxidation) reactions in biological systems is conserved as chemical energy in the ATP molecule transferring the stored energy to biosynthesis reactions.\(^{250} \) The chemical energy of ATP is then utilised in generating new cellular material; with this biochemical mechanism, only a portion of the free energy released may become available for biosynthesis. Servisi and Bogan\(^{126} \) found the microbial yield, \( Y_e \) of aerobic systems to be proportional to the free energy of substrate oxidation, \( \Delta G_S \):

\[
Y_e = -k \Delta G_{ox} \quad (32)
\]

McCarty\(^{251} \) has also developed a method for evaluating \( Y_e \) based on equivalence between substrate and biomass with the following fundamental energy expression:

\[
k_E Y_e \Delta G_S + \Delta G_S = 0 \quad (33)
\]

where \( \Delta G_S \) is the free energy required for the synthesis of an electron equivalent of biomass; \( \Delta G_{ox} \) is the free energy released by the oxidation of one electron equivalent of substrate, and \( k_E \) is the free energy transfer efficiency. McCarty also reported that \( k_E \) values would vary in the range of 0.4–0.7 for both heterotrophic and autotrophic microorganisms and suggested \( k_E = 0.6 \) as a consistent average value. Later, a yield value of 0.64 e\(^{-}\)e\(^{-}\).biomass/e\(^{-}\).e\(^{-}\).substrate was theoretically calculated using the same average \( k_E \) value for domestic sewage which was approximated with the empirical formula of \( C_{10}H_{13}O_3N \); \( Y_H \) of 0.64 mg cell COD mg\(^{-1} \) COD basically reflects the default value suggested in current activated sludge models.\(^{176} \) Similarly, a \( Y_H \) of 0.46 mg VSS mg\(^{-1} \) COD reported by Eckhoff and Jenkins\(^{153} \) on the basis of experiments on sewage, closely approximates this level using the routine conversion coefficient, \( f_x \), of 1.42 mg COD mg\(^{-1} \) VSS.

The concept of energetics has continued to serve for exploring the biochemical fundamentals of different processes taking place in activated sludge. A lower yield value \( Y_{ID} \) of 0.50 was suggested for anoxic growth on the basis energetic of using nitrate as electron acceptor in the absence of dissolved oxygen.\(^{252} \) Phosphorus removal by activated sludge could only be understood and properly implemented as a result of extensive studies on the biochemistry and metabolic modelling of enhanced phosphorus removal mechanisms.\(^{253,254} \) The energy requirement for microbial storage was assessed through the analysis of the biochemical mechanisms involved under different substrate removal conditions.\(^{255} \)

### Concept of bioreactor and mass balance

Wastewater treatment by activated sludge involves the same principles as the continuous culture of microorganisms. Therefore, understanding and controlling process performance and efficiency of activated sludge systems require, aside from the analysis of reaction kinetics, a proper definition of reactor hydraulics that would reflect mixing properties of biological reactors for the adopted process configuration. Reactor kinetics associated with system performance may then be evaluated, combining reaction kinetics and reactor hydraulics in terms of basic mass balance for substrate and biomass.

The concept of reactor hydraulics was first introduced to the field of sewage treatment in 1936 by Kehr\(^{256} \) who basically studied the retention time of the mixed liquor in continuous flow tanks. He indicated that particle residence time in continuous-flow ‘short mixing tanks’ agreed reasonably well with the theoretical curve derived with the assumption that the contents of the tanks were at all times completely mixed. A mass balance for complete mixing was derived for a mono-molecular reaction with discussion of possible application for activated sludge systems. Thomas and McKee\(^{257} \) interpreted ‘short circuiting’ in the aeration tanks in terms of rapid longitudinal mixing. They offered theoretical analysis for the residence time distribution of particles under complete mixing as well as for a series of two to 10 completely mixed tanks. It was concluded that the desired degree of mixing could be achieved with baffles to control the activated sludge process in relation to problems created by toxic wastes, bulking sludge, etc., especially in plants operated as tapered aeration or step feeding systems.

Garrett’s work\(^{258} \) may be considered as the pioneering effort in imposing a physical ‘hydraulic control’ mechanism over microbial mechanisms in activated sludge; however, the proposed approach did not find much support as it failed to rely on the basic principles of reactor kinetics. First, the suggested flow scheme was quite obsolete and costly to implement because it involved a second settling tank for excess sludge, as shown in Fig. 8; the flow rate of the mixed liquor to the excess sludge tank, \( Q_{w} \), presumably provided the hydraulic control for the system. Second, the following relationship was not recognised as the sludge age of the system:

\[
\theta_D = \frac{Q_{R}}{Q_{w}} \quad (34)
\]

where \( \theta_D \) is the the sludge age and \( V_R \) is the the volume of the aeration tank.

Finally, mass balance did not include the rate expression for endogenous decay or any term defining reactor hydraulics:

\[
\frac{dC_B}{dt} = \left( \frac{Q}{V_R} \right) (C_{B0} - C_B) - Y_B \mu_B X_{SS} \quad (35)
\]

where \( Q \) is the influent flow rate; \( C_{B0} \) and \( C_B \) are the influent and effluent BOD; \( Y_B \) is the the yield coefficient; \( \mu_B \) is the specific growth rate, and \( X_{SS} \) is the mixed liquor SS concentration. However, evaluation of experimental results was quite interesting. A plot
of effluent BOD against \( \mu_m \) yielded no correlation with a cluster of around 20 mg BOD/L and, similarly, a \( Y_H \) value of 0.9 mg BOD/mg MLSS was obtained when plotting unit BOD removal rate versus \( \mu_m \); the \( y \) intercept of the plot was 0.1 mg BOD/mg MLSS day\(^{-1}\), presumably indicating the substrate consumed for maintenance energy not accounted for in the corresponding mass balance.

Later, the concept of reactor hydraulics was emphasised by McKinney and O’Brien\(^{46} \) who postulated that increased stabilisation in activated sludge systems could be achieved with a ‘completely mixed aeration tank’ where the incoming wastewater would be mixed within the entire tank volume. McKinney\(^{206} \) also derived the ‘mathematics’ of the complete-mixing activated system. The paper mostly deserves recognition now because it defined for the first time active biomass, \( X_H \), in terms of ‘the oxygen equivalent’, i.e. cell COD, and basic relationship between substrate utilised, biomass generated and oxygen consumed when all are expressed as their COD equivalents. The same relationship is now adopted as the backbone of current activated sludge models:\(^{170,176} \)

\[
\frac{dC_s}{dt} = \frac{dX_H}{dt} + \frac{dS_0}{dt} \tag{36}
\]

where, \( C_s \) is the total biodegradable COD and \( S_0 \) is the dissolved oxygen. It is also one of the most difficult papers to follow mainly because process kinetics and stoichiometry are described using eight different coefficients (\( k_1 \) to \( k_8 \)) without any indication of biochemical reactions involved; for instance, \( k_1 \) defined the yield coefficient \( Y_H \) as:

\[
k_1 = 1 - \frac{Y_H}{X_H} \tag{37}
\]

and substrate removal was expressed in terms of \( k_5 \), as a first-order reaction:

\[
\frac{dC_s}{dt} = k_5 C_s \tag{38}
\]

Mass balance equations were derived using similar empirical terms for three modifications of complete mixing activated sludge, namely (1) aeration only without sludge separation and return; (2) aeration with sludge separation and return with excess sludge discharged in the effluent, i.e. extended aeration; and (3) a conventional process with separate sludge discharge. The paper should also be remembered by its final remark:

It is hoped that the lack of data – about different \( k \) values – will not be used as an excuse by engineers to maintain the empirical approach in preference to the scientific approach to biological waste treatment design.

In an evaluation of the extended aeration activated sludge, McCarty and Brodersen\(^{69} \) suggested the following mass balance for excess sludge generation:

\[
P_{VSS} = Y_{HB} M_{BOD} - b_n X_{VSS} \tag{39}
\]

where \( P_{VSS} \) is the excess sludge calculated in terms of VSS; \( M_{BOD} \) is the mass of BOD\(_5 \) removal per day and \( b_n \) is the endogenous decay coefficient. The value of the yield coefficient \( Y_{HB} \) was reported as 0.65 mg VSS mg\(^{-1}\) BOD\(_5 \), representing ‘the synthesis of both degradable and non-degradable solids, these fractions being 0.53 and 0.12, respectively’. The value of \( b_n \) was given as 0.18 day\(^{-1}\), referring this time to active biomass:

\[
\frac{dX_H}{dt} = -b_n X_H \tag{40}
\]

The mass balance suggested for excess sludge remains valid today, with the net yield value \( Y_{HB} \) replacing \( Y_B \). Washington and Hettling\(^{207} \) interpreted the results of seven different sets of experiments conducted with sodium acetate or glucose as the organic substrate, in terms of the following mass balance equation for active biomass including a process rate expression for the generation of particulate residual microbial products:

\[
\frac{dX_H}{dt} = \frac{Y_{HB} M_s}{V_R} - b_n X_H + f_n b_n X_H - \frac{X_{HB} Q}{V_R} \tag{41}
\]

where \( M_s \) is the mass of organic substrate removed per day; \( X_{HB} \) is the active biomass lost in the effluent; \( Q \) is the flow rate, and \( V_R \) is the reactor volume. Middlebrooks and Garland\(^{70} \) suggested a similar mass balance as indicated in Eqn (41), for the extended aeration configuration of activated sludge with no sludge wastage other than solids loss in the effluent, and they derived the following expression implying the control of hydraulic residence time, \( \theta_H \) on microbial growth:

\[
\mu_H = k_0 + \frac{1}{bb_H} \tag{42}
\]

where \( k_0 \) is the the overall decay coefficient; \( b \) is the \( X_{VSS}/X_{VSS_e} \) ratio and \( X_{VSS_e} \) is the biomass concentration in the effluent.

The interpretation of all previous suggestions in terms of process parameters that could be related to microbiological mechanisms involved and, at the same time, readily adaptable to process design and operation was offered by Pearson\(^{127} \) in a major work for the activated sludge process, as it is understood today, with the net yield value \( Y_{HB} \) replacing \( Y_B \). The first basic parameter was the specific substrate removal rate, \( q \), derived on the basis of mass balance for organic substrate (BOD\(_5 \) or COD) around the process:

\[
q = \frac{C_{T1} - C_T}{\theta_x X_{VSS}} \tag{43}
\]

where \( C_{T1} \) and \( C_T \) are the influent and effluent total substrate concentrations, respectively.

The second parameter was the sludge age, \( \theta_x \) (sludge retention time, SRT), which defined the culture history and the net growth for the biomass sustained in the aeration tank:

\[
\theta_x = \frac{V_x X_{VSS}}{P_{VSS}} \tag{44}
\]

The specific substrate removal rate, \( q \), could be related to the specific growth rate, \( \mu_H \), with the yield coefficient, \( Y_H \):

\[
\mu_H = Y_H q \tag{45}
\]
After 1920, however, practice was then directed towards fill and draw plants were placed into operation in England and in the fill and draw setup. From 1914 to 1920 several full-scale fill and sludge plant built at Salford was an upscale of the experimental operation in the purification of the tested sewage. The first activated appropriate cycle time which would ensure complete nitrification for biomass, defining the net growth rate, 1/θ on the basis of solids mass balance:

\[
\frac{1}{\theta_X} = \mu_H - k_D = Y_H q - k_D
\]

The study by Jenkins and Garrison, which shortly followed Pearson’s work, should also be considered as an equally significant contribution, firmly establishing the sludge age – or the mean cell residence time – as the major parameter for biomass activity and growth, and COD as the modelling and design parameter for organic substrate. An excellent account of the concept of the sludge age is provided by Lawrence and Middleton. Table 5 outlines information compiled in these studies related to process kinetics and stoichiometry of activated sludge. It should be noted that the pioneering efforts outlined above, contributed a great deal to the concept of reactor kinetics, conveniently expressed in terms of corresponding matrix representation in activated sludge models structure for different purposes.

### Kinetic and stoichiometric coefficients for the activated sludge process

<table>
<thead>
<tr>
<th>Substrate</th>
<th>BOD₅ (mg L⁻¹)</th>
<th>COD (mg L⁻¹)</th>
<th>BOD₅ (mg L⁻¹)</th>
<th>COD (mg L⁻¹)</th>
<th>BOD₅ (mg S mg⁻¹ VSS)</th>
<th>COD (mg S mg⁻¹ VSS)</th>
<th>k_d (day⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic sewage</td>
<td>0.500</td>
<td>0.055</td>
<td>0.480</td>
<td>0.045</td>
<td>0.420</td>
<td>0.087</td>
<td></td>
</tr>
<tr>
<td>Skim milk</td>
<td>0.645</td>
<td>0.370</td>
<td>0.770</td>
<td>–</td>
<td>0.440</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>0.770</td>
<td>–</td>
<td>0.450</td>
<td>0.050</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aromatics and aliphatics</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic sewage</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synthetic sewage</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharmaceutical waste</td>
<td>–</td>
<td>–</td>
<td>0.645</td>
<td>0.370</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical waste</td>
<td>–</td>
<td>–</td>
<td>0.770</td>
<td>–</td>
<td></td>
<td></td>
<td>0.200</td>
</tr>
<tr>
<td>Domestic sewage</td>
<td>–</td>
<td>–</td>
<td>0.450</td>
<td>0.050</td>
<td></td>
<td></td>
<td>0.050</td>
</tr>
</tbody>
</table>

BOD₅ is the 5-day value of the test for biochemical oxygen demand. COD, chemical oxygen demand; VSS, volatile suspended solids.

### Emerging technologies

**Sequencing batch reactor**

The fill and draw system constituting the basic principle of the sequencing batch reactor (SBR) as it stands today was the basic operating scheme of the experiments that initiated the discovery of the activated sludge process. Bartow and Mohlman adopted the same cyclic operation and feeding and investigated the appropriate cycle time which would ensure complete nitrification in the purification of the tested sewage. The first activated sludge plant built at Salford was an upscaled version of the experimental fill and draw setup. From 1914 to 1920 several full-scale fill and draw systems were placed into operation in England and in the US. After 1920, however, practice was then directed towards the operation of the process under continuous-flow conditions. In 1930 Imhoff and Moore argued that ‘continuous flow has in every instance proved itself superior to intermittent flow and in the history of sewage treatment, decades have been lost in the use of intermittent processes.

After the choice of continuous-flow processes, the interest in intermittent operation was only revived in the 1960s. SBR was then developed as a novel activated sludge system, replicating the old fill and draw concept with a single tank that served both as a biological reactor and settler in a temporal sequence. Extensive research highlighted its simple and flexible operation as competitive advantages against conventional activated sludge scheme and encouraged its full-scale application both for sewage and industrial wastewaters.

The operation features of SBR also reflect during each cycle transient responses of all observed parameters such as COD, nitrogen forms, storage products, oxygen uptake rate, etc. Therefore, it is also extensively used as a perfect experimental tool for fundamental and applied research.

### Enhanced biological phosphorus removal

Storage of biopolymers was one of the early observations related to substrate removal by activated sludge; glycogen was the identified storage product. The concept of storage was revived in the 1960s in relation to the hypothesis of biologically mediated excess phosphorus removal. In experiments conducted in 1965, Levin and Shapiro demonstrated that the removal mechanism depended upon biological activity. This concept was extensively pursued in the following studies revealing the biochemical basis of what is now known as the enhanced biological phosphorus removal (EBPR), with a clear indication that substrate storage played a key role in the mechanism. The process requires an anaerobic/aerobic sequence; (2) short-chain fatty acids (VFAs) and particularly acetate generated during the anaerobic phase are readily converted to poly-β-hydroxybutyrate (PHB), a storage product synthesised from acetyl-CoA consuming energy; (3) specific poly-P microorganisms prevail due to their ability for storing PHB; the energy required to store PHB in the anaerobic zone is provided through cleavage of poly-P releasing phosphorus into the bulk solution; and (4) the stored PHB is utilised as a source of internal organic carbon and energy for the microbial growth and excess phosphorus uptake that occurs in the aerobic zone. Different configurations of the EBPR process are now extensively used with or without simultaneous nitrogen removal in cases where nutrient control is required for treated sewage discharges.

Substrate storage is now the basic ingredient of all studies devoted to substrate removal in activated sludge systems operated under dynamic conditions due to structure of the flow scheme, natural fluctuations in the feed flow, substrate
composition, temperature, etc.\textsuperscript{278} Furthermore, polyhydroxyalkanoates (PHAs) also synthesised by activated sludge as storage biopolymers have plastic properties and are now valued as good substitutes of petroleum derived plastics. Consequently, studies are also directed towards bioplastics production from different wastewaters using appropriate flow schemes for activated sludge operation.\textsuperscript{279,280}

The oxic-settling anaerobic process

As previously outlined, subdividing the aeration capacity of the activated sludge process in two different volumes by providing additional aeration to the return sludge was used in most early plants. The basic motivation behind this type of a flow scheme was the primary role attributed to adsorption in substrate removal. It was expected that (1) a major fraction of both suspended solids and dissolved organic matter would be removed by adsorption in the first contact tank; (2) settling properties would be improved by suspended solids entrapped on the activated sludge; and (3) the second re-aeration tank would restore the sorptive capacity of settled sludge. The contact stabilisation was developed on this basis, presumably to provide a balance between the adsorption potential and the ‘reactivation’ of activated sludge. Later, a mechanistic evaluation revealed the ability of the process to virtually differentiate and exercise separate control on the growth and decay of the activated sludge. This flexibility of the contact stabilisation flow scheme offered the advantage of operating at a lower net growth rate compared with a conventional system under the same overall COD loading rate.\textsuperscript{281,282} From a practical standpoint, this meant that the sludge production could be significantly reduced by appropriate manipulation of the process. The popularity of contact stabilisation did not last long, as the practice was soon switched to biological nutrient removal, which required different flow configurations and longer retention times.

Recently, the increasing expense of sludge disposal diverted research efforts to minimise excess biological sludge production. Many alternative technologies were developed for this purpose, involving different processes such as ozonation, chlorination, heating, controlling growth by metabolic uncouplers, etc.\textsuperscript{283,284} The oxic-settling anaerobic process (OSA), also commercially termed the ‘cannibal process’ is particularly interesting because it is inspired by the concept of re-aeration, i.e. a second tank for the sludge return, advocated in the early years of activated sludge plants. In this novel process, re-aeration is substituted with an anaerobic tank; the operation involves the interchange of biomass between the aerobic and the anaerobic phases on daily basis and the recycled biomass is held in the anaerobic ‘cannibal’ reactor for a specified solids retention time.\textsuperscript{285,286} This process, while promising for effective solids control, still requires further investigation on the dynamics of the microbial community sustained under two different environments.

Super-fast membrane bioreactor

Gravity settling of biomass to obtain a clarified effluent has always been a major problem that haunted conventional activated sludge systems. Under unfavourable conditions still difficult to identify and control, a complex filamentous microbial community may prevail and create severe bulking and foaming problems.\textsuperscript{287,288} This is the reason why the system design is primarily targeted to ensure effective flocculent settling, mostly over-sizing the aeration tanks in terms of substrate removal. The concept of high rate activated sludge drastically reduced the aeration time down to around 2.0 h. The idea first flourished in 1941, from the incidental observation of the substantial BOD\textsubscript{5} reductions that took place during the start-up of the Ward Island treatment plant in New York and survived until today.

The membrane bioreactor (MBR) has been a major breakthrough which totally removed gravity settling; it eliminated major constraints on sludge retention time and biomass concentration, traditionally implemented for sustaining good settling properties.\textsuperscript{289} This potential was initially interpreted as a possibility to sustain substantially higher biomass concentrations without facing operational problems.\textsuperscript{290–292} A number of studies also investigated the idea of operating MBR as a high rate system at low sludge ages below 5.0 days.\textsuperscript{293,294} Results on satisfactory performance instigated the concept of superfast membrane bioreactor, which would be operated below a sludge age of 2.0 days, with the ability for complete removal of soluble substrate, while conserving particulate substrate for energy recovery.\textsuperscript{295} Recent studies confirmed stable operation of the superfast MBR under high loadings with significantly low levels of soluble microbial product (SMP) generation.\textsuperscript{296,297}

CONCLUDING REMARKS

Two statements from the true masters in this field, Sawyer and McKinney, should be remembered for discussing the past and the future of the activated sludge process after the first 50 years. In 1965, Sawyer\textsuperscript{62} addressed the following major question:

Has the activated sludge process been developed to the maximum degree possible? In view of the fact that most recent developments have been based on biological concepts, it does not seem very likely that many new and startling developments will occur.

Indeed, amazing concepts and ideas were developed in this period, thanks to the hard work of many scientists, who not only established the essential basis of organic carbon removal, but also provided invaluable clues for future achievements. It is a pity that most scientists actively working in this field have only a vague notion of this period, very much limited to its discovery by Ardern and Lockett and related empirical practice.

The same year, McKinney\textsuperscript{78} expressed quite a different opinion about the future of the activated sludge process:

The activated sludge process poses a definite challenge for research and development in the future. The process appears to be standing on the threshold of even greater achievements than in the past.

McKinney’s prediction proved to be quite justified, as amazing developments could be attained to better understand and improve the process during the last few decades. It will only be fair to acknowledge that the amazing progress of the activated sludge process today is largely due to the pioneering efforts in the early period. It is strongly believed that future research would and should still benefit from the tremendous scientific information accumulated piece by piece through the early years.

ACKNOWLEDGEMENTS

The dedicated support and contribution of Busra Alli, Cansu Karaca, Dr Hakan Dulkadiroglu, Professor Nevin Yagci and...
Professor Seval Sozen to all phases in the preparation of the review is gratefully acknowledged.

REFERENCES
23 Clark HW and Gage S De M, Experiments Upon the Purification of Sewage and Water at the Lawrence Experiment Station, 44th annual report. State Board of Health, Massachusetts, pp. 275–367 (1912).
24 Clark HW and Gage S De M, Work of the Lawrence Experiment Station for 1913—the Purification of Sewage and Investigation on Allied Subjects, 45th annual report. State Board of Health, Massachusetts, pp. 265–368 (1913).
26 Clark HW and Adams GO, Experiments Upon the Purification of Sewage and Water at the Lawrence Experiment Station During the Year 1914, 46th annual report. State Board of Health, Massachusetts, pp. 295–346 (1914).
35 Bartow E and Mohlman FW, Recent results of experiments on the purification of sewage by aeration in the presence of activated sludge. J Ind Eng Chem 8:151 (1915).
40 Melling SE, The purification of Salford sewage along the lines of the Manchester experiments. J Soc Chem Ind 33:1124–1130 (1914).
52 Ruchhoff CC, McNamie, PD and Butterfield CT, Biochemical oxidation by activated sludge. Sewage Works J 10:661–690 (1938).
The activated sludge process


The activated sludge process


228 Mejia AC and Cisneros BJ, Particle size distribution (PSD) obtained in effluents from an advanced primary treatment process using different coagulants, in *Chemical Water and Wastewater Treatment IV*, ed. by Hanhh HH, Hoffmann H and Odegaard H. Springer-Verlag, Berlin, pp. 258–268 (2000).
233 Hocaoglu SM and Orhon D, Particle size distribution analysis of chemical oxygen demand fractions with different biodegradation characteristics in black water and gray water. *Clean Soil Air Water* 41:1044–1051 (2013).
The activated sludge process


