

POLITECNICO DI TORINO

Department of Environment, Land and Infrastructure
Engineering
Master of Science in Environmental and Land Engineering

Master's degree Thesis

**Modeling and Control of Industrial Wastewater
Physico-chemical treatment processes**



Supervisor:
Prof. Maurizio ONOFRIO

Candidate:
Ivo GRAZIOTIN
Reg. N: 240796

Academic year: 2020-2021

Abstract

The purpose of the present thesis is to shed light on a very complex subject such as the modeling and control of industrial wastewater physico-chemical treatments, with particular attention to the coagulation-flocculation process with simultaneous precipitation of heavy metals. After a brief but comprehensive introduction to the theory of coagulation, some concerns about the applicability of DLVO theory in the context of process simulation are raised, thus urging the need to find simpler but equally explanatory formulations that can be practically implemented within numerical models. In this regard, a fairly in-depth review of the progress that flocculation theory has made over the past forty years is given, going as far as modeling flocs as porous and hydraulically permeable fractal aggregates. After recalling the philosophy behind a paper published by Hogg in 1984 about a modelling approach for the collision efficiency factor for particle flocculation using polymers, an extension to this model is proposed, to take into account for more mechanisms of particle destabilization such as sweep coagulation, adsorption and interparticle bridging that possibly act simultaneously. A discretized population balance equation is then proposed to model the coagulation-flocculation-precipitation process using the above formulation in the framework of a linear driving force model.

The topic of process control is subsequently introduced, showing the main reasons and challenges behind industrial wastewater treatment. After a brief overview of the main physical sensors and analyzers related to coagulation and flocculation available on the market, the problem of fouling is remarked as an issue that places considerable limits on the choice of the amount, types, sampling points and sampling frequency of sensors to be installed. The issue of dead time is also mentioned, as a potential source of disturbance for the control system. A brief review of the main studies conducted over

the last 30 years in the field of online monitoring and control of the coagulation process is then given, raising particular attention on the so-called "software sensors". The important role that machine learning and data-driven techniques could play for model and/or control laws discovery is then highlighted, showing that the relatively high cost associated with pilot plant testing to train these systems usually pays off in the long term. Finally, the results of an experimental study on the coagulation-flocculation of washing wastewaters from the automotive painting plant of FCA Mirafiori (Turin, Italy) are illustrated. An equally effective but less expensive alternative to the treatment currently performed in the Mirafiori plant is proposed, also showing the existence of a relationship between turbidity, solids content and chemical oxygen demand (COD) of the influent, as well as between COD and the integral of the UV absorption spectrum below 280 nm. To conclude, some uses of these formulas are proposed for automatic control purposes.

Contents

Abstract	ii
Introduction	vii
1 Theory and practice of Coagulation	0
1.1 Colloids and Suspensions: solid-liquid dispersions and particle stability	0
1.2 Particle destabilization: destabilization mechanisms	32
1.3 Destabilization of hydrophobic colloids by metal coagulants	47
1.4 Destabilization of hydrophilic colloids by metal coagulants	58
1.5 Interactions of destabilizing chemicals with soluble materials	59
1.6 Combination of Additives	60
2 Flocculation Modeling	62
2.1 Collision mechanisms and the Smoluchowski equation	62
2.2 Collision Frequency: the Long-Range Force Model	70
2.3 Collision Efficiency: the Short-Range Force Model	81
2.4 Modeling of Flocculation with Fractal Dimensions	88
2.5 The Surface Precipitation Model	95
2.6 Hogg's Model (HM) for Particle Flocculation using Polymeric Coagulants	99
2.7 A proposed Extension of HM to model ballasted flocculation with enhanced coagulation	106
3 State of the Art of Online Monitoring and Control of Clariflocculation Processes	119
3.1 Direct Use of Online Parameters for Dosing Control	124

3.2	Indirect Use of Online Parameters for Dosing Control	133
3.3	Conclusions and future developments	139
4	Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)	145
4.1	Site Inspection and First Observations	145
4.2	Jar tests and chemical analyses: materials and methods	147
4.3	Results and Discussion	157
5	Conclusions and Future Developments	172
	Bibliography	176
	List of Figures	178
	List of Tables	187

Introduction

The thesis is organized as follows:

Chapter 1 introduces the fundamentals of the theory of Coagulation, starting from exploring the main factors affecting particles stability in solid-liquid dispersions and the origin of surface charge that promotes such stability. After a brief overview of particle stability according to the Guoy-Chapman model and DLVO theory, the four main particle destabilization mechanisms are introduced, namely: compression of the diffuse layer, adsorption and charge neutralization, enmeshment in a precipitate and adsorption with interparticle bridging. These mechanisms are explained in great detail in order to highlight the main factors affecting their performance and the expected outcomes with respect to typical scenarios encountered in engineering practice.

In Chapter 2 Flocculation modeling is introduced, starting from the first attempt given by Smoluchowski in 1917. After that, several advanced models have been proposed over the last century, starting from the so-called Long-range force model, which identified three main long-range transport mechanisms for interparticle collision, namely collision by fluid shear, by Brownian motion and by differential sedimentation. Thanks to the Short-range force model, it was also possible to take into account short-range phenomena arising from van der Waals, double layer and hydrodynamic interactions due to electrical and viscous forces, which always oppose to interparticle collision. Flocculation modeling with fractal dimensions is then introduced as a new way to improve the previous models by considering flocs as porous fractal aggregates. Surface Precipitation Model and Hogg's model for particle flocculation using polymers are then introduced as two simple and extremely useful theories to build models for the collision efficiency function. An extension to Hogg's model was in fact proposed.

In Chapter 3 the problem of coagulation control is introduced, starting from a review of several studies conducted over the last decades in this field. The importance of machine learning techniques is highlighted as it is probably the most promising approach for both modeling and real-time control of clariflocculation processes.

In Chapter 4 is reported an experimental study conducted on washing wastewaters from the automotive painting plant of FCA Mirafiori (Turin).

Chapter 1

Theory and practice of Coagulation

1.1 Colloids and Suspensions: solid-liquid dispersions and particle stability

1.1.1 Solid-Liquid dispersions: preliminary definitions

Solid matter usually found in waters and wastewaters essentially derives from soil erosion, vegetation decay, mineral dissolution, precipitation processes and domestic and industrial waste discharges. Such material may appear in different forms (suspended and/or dissolved) and natures (inorganic, organic or biological), depending on its origin and on the physical, chemical and biological processes to which it has been subjected during its lifetime. With respect to table 1.1, particles smaller than approximately $10^{-4} \div 10^{-5}$ mm are commonly referred to as *Colloids*; examples of colloidal material are mineral microparticles, small aggregates of precipitated and flocculated matter, silt, bacteria, plankton, viruses, biopolymers and macromolecules. The term *Solutions* is instead used when referring to Solid-Liquid Dispersions of particles smaller than approximately 10^{-6} mm; such material comprises, for example, inorganic ions and complexes, molecules, polymeric species and polyelectrolytes, undissociated solutes and small aggregates. Solid-Liquid Dispersions of particles in the size range of $10^0 \div 10^{-3}$ mm are finally referred to as *Suspensions*, as they can settle out of the mixture (unlike solutions and colloids) if left undisturbed for a prolonged period of time. Both colloids

and suspensions are *heterogeneous mixtures* of particles in a liquid medium, that can be easily distinguished from one another depending on the effect that the dispersed material exerts on a visible light beam crossing the mixture when immersed in a light-transmitting medium: whilst the former (colloids) exhibit the so-called "*Tyndall effect*" (or *Tyndall scattering*), the latter have little to no (significant) effects on visible light, as the size range of suspended particulate matter can be much greater than the wavelengths of visible light. Under the Tyndall effect, the longer wavelengths are more transmitted, while the shorter ones are more diffusely reflected via scattering, which can make a colloidal dispersion appear as either milk-like or bluish/greyish depending on the *grade* of the *particle size distribution* (if the PSD is well sorted or not) and if the PSD is centered on a particular wavelength within the visible spectrum (i.e. a poorly graded PSD with particles that mostly range between 400 and 700 nm). The intensity of scattered light depends not only on the frequencies of the incident ray, but also on the concentration of particles in the mixture, which affects color saturation. Solutions, on the other hand, exhibit the so-called "*Rayleigh scattering*", which is similar to the Tyndall effect in that the diffusion of light is still wavelength-selective (e.g. blue light is scattered more than red light), but particles are here supposed to be much smaller than the wavelengths of the incident beam, leading to a greater wavelength dependence of the (spectral) diffusion coefficient, which in this case is inversely proportional to the fourth power of the wavelengths of the incident beam. One last important difference between Rayleigh and Mie scattering (which refers to the Tyndall effect when the particles diameter and the wavelength of the incident ray are approximately equal), is that in the latter the diffusion coefficient is less isotropic, giving a forward lobe of scattered light wider than the backward lobe (which are roughly equal in Rayleigh scattering). Particles in colloids and solutions are able to retain a dispersed state because of certain inherent characteristics that promote their stability (i.e. the ability to remain as independent entities within the dispersion), which are substantially attributable to *interfacial phenomena* related to the effects of surface charge carried by the particles themselves and the degree of hydration (or solvation) of their surface layers. These interfacial phenomena exist because of the inherent high specific surface area (surface area to mass ratio) that colloids exhibit due to their small size, which eventually reaches a point where the ratio between surface and mass forces is strongly biased towards the former (and

1. Theory and practice of Coagulation

Table 1.1: Classification of Solid-Liquid Dispersions based on particle size. *Source:* from [1]

Particle Size (mm)	Classification	Examples	Total Surface Area (m ² /cm ³)	Time Required to Settle 100 mm if Specific Gravity = 2.65
10	Coarse dispersion (visible to naked eye)	Gravel, coarse sand, mineral substances, precipitated and flocculated particles, silt, macroplankton	6×10^{-4}	0.1 second
1			6×10^{-3}	1 second
10^{-1}			6×10^{-2}	13 seconds
10^{-2}	Fine particulate dispersion (visible under microscope)	Mineral substances, precipitated and flocculated particles, silt, bacteria, plankton and other organisms	0.6	11 minutes
10^{-3}			6	20 hours
10^{-4}			60	80 days
10^{-5}	Colloidal dispersion (submicroscopic)	Mineral substances, hydrolysis and precipitated products, macromolecules, biopolymers, viruses	600	2 years
10^{-6}			6,000	20 years
$<10^{-6}$	Solution	Inorganic simple and complex ions, molecules and polymeric species, polyelectrolytes, organic molecules, undissociated solutes		

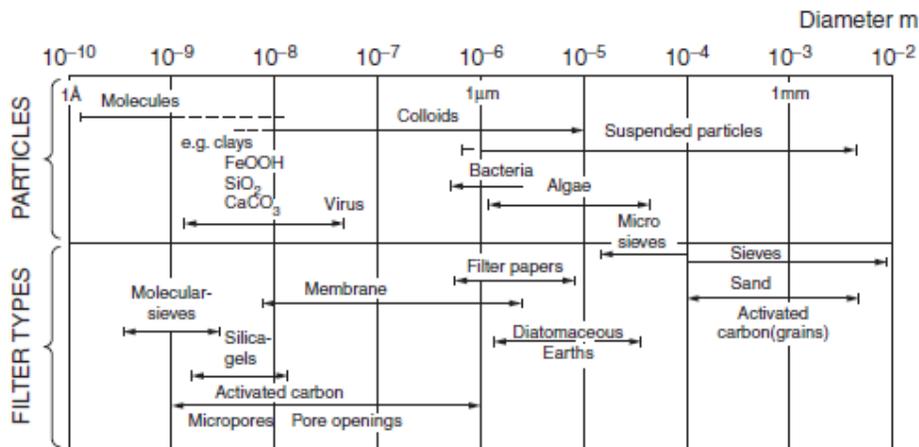


Figure 1.1: Size spectrum of waterborne particles and filter pores. *Source:* from [1]

both gravity and inertia are no longer relevant if compared to electrostatic interactions at the particles surface). As shown in table 1.1, colloids may require timescales of years to settle down through distances lower than 1 m, if considering gravity alone; in reality, particles are subjected to several other phenomena, like (thermal) convection currents and *Brownian motion* (which arise from the molecular and ionic bombardment by the liquid medium), that effectively keep the dispersed particles in permanent suspension. For these reasons, fig. 1.1 shows a few other ways to remove particles from a stable dispersion that don't involve mass forces (either gravity - like settling and flotation processes - or inertia - usually exploited by centrifugation processes -), such as adsorption or physical straining, like granular media filtration and membrane

processes. In granular media filtration, the suspension (particles and water) is forced to pass through a filtering body: particles are captured in the media by physical-chemical processes while water passes through, and when the particles accumulated within the media begin to compromise its performance, the filter is temporarily taken out of service to clean the granular media. Finally, in membrane filtration, particles are removed by *sieving* (i.e. particles larger than the membrane pores are retained, while water passes through) and depending on the design, membranes can also remove soluble materials. Considering the large volumes involved in water and wastewater treatment, most of these technologies may however be economically unfeasible if used alone, although exceptions may exist depending on the treatment purposes and on the quality of the raw influent. Typically, particles in the influent suspension would not be effectively removed in these treatment processes without some kind of pretreatment. One of the cheapest and historically consolidated ways to convert particulate, colloidal and/or dissolved matter into a form whereby separation from a dispersion (or solvent) is practicable is the so-called *physico-chemical treatment*, which basically consists on a *chemical pretreatment* of the dispersion, followed by a *physical separation process* among those already introduced that will intimately depend on the physical properties of the chemically-altered particles. Focusing on the chemical pretreatment alone, two aims may be in general achieved depending on the subsequent separation process:

- Precipitation of dissolved matter into a particulate material that can easily be removed by sedimentation and/or filtration;
- Chemical alteration of the surface properties of the particulate matter aimed to either increase its adsorptivity onto a given filter medium, or its tendency to aggregate into larger particle units (*coagulation* and *flocculation*).

Such conversion of the stable state of a given dispersion or solution to an unstable state is called *destabilization*. Pretreatment processes typically involve a design of the mixing processes as well, to bring about more particle-particle collisions. Ultimately, the success of particle removal depends in fact on the attachment of particles to other surfaces, either other particles in suspension or surfaces provided by the engineered system. O'Melia and Stumm (1967) introduced the concept that the removal of a particle in such processes can be envisioned as two (more or less) independent steps: a long-range transport of the particle to the vicinity of that other surface,

and short-range transport to allow the particle to collide with and attach to that surface. Long-range transport is dominated by forces that are primarily physical in nature, while short-range transport and attachment rely on both physical and chemical forces. Details of the long-range transport step differ among the various types of particle treatment processes (e.g. in granular media filtration and flocculation the long-range transport mechanisms involved slightly differ from one another), while particle attachment is almost indifferent to the removal process, because the short-range forces that allow (or do not allow) particle attachment to occur are essentially the same. It is thus mandatory to understand why particles *do not* form flocs and how to overcome the problem of particle stability for the successful removal of particles from water. As it will be seen later in more detail, the primary factor causing particles to be stable in water is their surface charge: if two similarly charged particles in a suspension approach one another, a repulsive interaction will alter their trajectory and, if the repulsion is strong enough, no collision occurs and the attachment is prevented. Some authors use the terms "coagulant" and "flocculant" interchangeably and even those authors who distinguish between the terms, do not agree on the distinctions. Some use "coagulants" to refer to chemicals that work by the precipitation mechanism and flocculants to refer to those that work by adsorption; others distinguish inorganic additives (coagulants) from organic polymers (flocculants), regardless of the destabilization mechanism(s) involved. There are also people who distinguish chemicals that accomplish charge neutralization (coagulants) from those that work by interparticle bridging (flocculants) and some of these chemicals are referred to by their use (*filter aids, dewatering aids*, etc.) when in fact the same chemicals might be used for other purposes (e.g. to accomplish destabilization for other particle removal processes). Since the literature (and the engineering practice) is often ambiguous on this topic, to avoid these terminology problems the following generally accepted definitions are introduced and used throughout the present work:

- We'll refer to *Coagulation* as the process whereby destabilization of a given dispersion is effected in order to overcome those factors that promote its stability.
- We'll refer to *Flocculation* as the process whereby destabilized particles - or particles formed as a result of destabilization - are induced to come together into larger agglomerates.

- We'll refer to *primary coagulants* as the chemicals added to a given suspension, colloid or solution to promote destabilization (regardless of the mechanism(s) involved); these *destabilizing chemicals* might thus be either inorganic or organic in nature.
- We'll refer to *flocculant (or filter) aids* as the chemicals added to a destabilized dispersion to accelerate the flocculation rate or to strengthen flocs formed during flocculation; often times the term *flocculant* is reserved for high-molecular weight organic molecules that work by *interparticle bridging* because it is the only mechanism that promotes agglomeration of destabilized particles and floc growth even in highly unfavorable hydraulic conditions (e.g. high shear rates).

The former distinction between coagulant and flocculant aids is therefore based not on the nature of the chemicals used (metals or polymers), but rather on the role that they exert once introduced into a stable dispersion. It's important to note that from the exact moment that the primary coagulant is added, flocculation begins immediately as a result of molecular bombardment (Brownian motion); this flocculation stage is often known as *perikinetic flocculation*, and it is the most relevant flocculation mechanism involving particles of colloidal size. However, for appreciable floc formation, some form of induced turbulence in the suspending medium is required: this stage, also known as *ortokinetic flocculation*, can be promoted by either mechanical agitation (using well designed mixing devices) or by passively inducing tortuous flow patterns to the suspending medium, using for example baffles installed in the flocculation tank(s). Once the flocculation process reaches its steady state, the destabilized and flocculated stream is passed to a solid-liquid separation stage where the formed flocs are either separated by gravity (sedimentation/flotation) or filtered from the liquid. The accumulated solids may then be further treated using sludge conditioners (like lime or polyelectrolytes) to enhance their dewatering and facilitate their ultimate disposal.

1.1.2 Colloids and Interfaces: origin and effects of surface charge

The term *stability* refers to the ability of individual particles to remain as separate entities or, in other words, to maintain a dispersed state. The stability of colloidal

material arises from the predominance of surface forces at the solid-liquid interface, that can be attributed to either the presence of a surface charge or the hydration of surface layers of the colloid. Depending on their degree of affinity for the liquid phase, colloids in water and wastewater treatment can be classified into two main categories: hydrophobic and hydrophilic. The amount of water bound by hydrophilic particles may in some cases be three to ten times their own dry mass [1]. On the other hand, examples of hydrophobic colloids are gelatin, starch, gums, proteins and all biocolloids like viruses and bacteria. Both types may co-exist in a particular system and there are also situations in which transitions from one state to the other may occur during the treatment. Although the term hydrophobic infers a definite solid-water boundary, there is in reality a layer of water molecules, perhaps in some cases only a molecule thick, bound strongly enough to the colloid surface such that as the colloid moves in the liquid medium, the plane of shear is still a water-water boundary. Three principal ways in which the surface charge may arise are universally accepted (STUMM):

- Charge could arise from *chemical reactions at the surface*. Several common types of solid particles that are usually encountered in water and wastewater treatment (such as oxides and hydroxides, other inorganics and biogenic organic solids) expose surface sites consisting of acid/base functional groups that are readily ionizable (i.e. that interact with protons and hydroxide ions in solution to form both positively and negatively charged surface species), such as $-OH$, $-COOH$ and OPO_3H_2 . Surface charge depends in this case on the *degree of ionization* (proton transfer) of these functional groups and, consequently, on the pH of the surrounding liquid: just as in solution chemistry, the dominant species shifts from the most protonated form (positively charged) at a lower pH values to the least protonated (negatively charged) at higher pH. Any deprotonation reaction can be described by an equilibrium constant formally known as K_a . It is worth noting that the change from one form to another occurs over a much wider pH range at the surface than for dissolved species, because the change in surface charge accompanying each reaction opposes additional reactions of the same type. For particles that interact in this way the surface charge is zero at some intermediate pH. If no other ions contribute to the surface charge, this value is referred to as the pH_{PZC} , where PZC stands for *point of zero charge*: the net surface charge is negative at all pH values above the pH_{PZC} , while it is positive otherwise. If the

surface charge is the result of interactions with other ions in addition to protons and hydroxide ions, the pH value at which it is zero is called the *isoelectric point* (pH_{IEP}). Surface charges of this nature may finally originate from processes in which solutes bind coordinatively to the solid surface.

- Surface charge could also manifest because of *lattice imperfections* at the solid surface deriving from *isomorphic substitutions*. When a solid forms in a solution, it often contains a small fraction of exogenous species in the crystal lattice that can only occur if the species entering the solid lattice are approximately the same size as the species that they replace, so that the replacement can occur without stressing or deforming the solid structure too severely. If the charges on the two species of interest are not equal, the solid acquires a charge that is expressed at the surface. Isomorphic substitution is common in clays, alumino-silicate solids composed of aluminium oxide (Al_2O_3) and silicon dioxide (SiO_2) layers in a repeating pattern, for which it is common for either ferrous iron (Fe^{2+}) or magnesium (Mg^{2+}) to be substituted for aluminium (Al^{3+}), and for aluminium to be substituted for silica (Si^{4+}); each of these isomorphic substitutions causes the clay particles to acquire a net negative charge. Surface charge of this type is hence pH-independent.
- Surface charge could finally arise from *ion adsorption* due to van der Waals forces or hydrogen bonding, and ions can be either the same or different from those comprising the surface. Let's consider for example a suspension containing $CaCO_3(s)$ particles that it is bound to have only Ca^{2+} and CO_3^{2-} ions in solution. If the adsorption density of Ca^{2+} is greater than that of CO_3^{2-} , the imbalance of adsorbed ions will contribute a net positive charge to the surface. Such an imbalance might be described either in terms of unequal adsorption of the ions or in terms of a *non-stoichiometric* dissolution of the solid. When surface complexes form with ions that are different from those that compose the solid surface, the process is clearly one of adsorption. Perhaps the most important example of adsorption that affects the surface charge of particles is that of large organic molecules, such as those identified as *natural organic matter* (NOM) whose estimation is of crucial importance in drinking water treatment. Such molecules are usually strongly adsorbed onto both inorganic minerals and

manufactured adsorbents (like activated carbon or ion exchange resins) and their functional groups are often charged in the pH range of interest so that when NOM molecules adsorb, they contribute charge to the particle surface. It is now widely accepted that the adsorption of NOM onto particles in natural waters is almost always the principal determinant of their surface charge (O' MELIA ET AL.). NOM mostly arises from animal and plant decay and from the metabolism of plankton and aquatic bacteria. It enters freshwater bodies (e.g. lakes, streams, groundwater) primarily via runoff and infiltration, but can also be produced directly within them. Its structure and chemical composition is complex and can vary substantially depending on its origin, but generally consists of both aliphatic units and aromatic rings that incorporate carboxyl, hydroxyl, amine and carbonyl groups. Invariably, parts of these molecules are unionized and relatively nonpolar and thus hydrophobic, while other parts (those containing strongly polar or ionized functional groups) are hydrophilic. As a result, NOM molecules adsorb onto virtually all of the particles present in any suspension of interest, making their surface properties much more similar than it would be expected from their underlying chemical composition, regardless of the water body under consideration. Because of NOM and their inherent structure and chemical composition, particles encountered in drinking water treatment of both surface waters and groundwaters are always expected to exhibit a negative surface charge in the pH range of interest (6-9).

Both the charge carried by colloids and the dipolarity of the water molecule explain the overall stability of a dispersion, as the number of bonded layers and orientation of water molecules within these layers deeply affect the magnitude of the dielectric constant of water as a function of the distance from the solid-liquid interface (i.e. the *permittivity* profile), as shown in fig. 1.2. Hydrophobic colloids destabilization therefore consists in a reduction of the effective surface charge aimed to reduce both its radius of influence and the extent of zone of ordered water structure. As a consequence, adjacent collidal particles, heretofore remaining apart due to electrostatic repulsion and, to a certain extent, to the ordered water structure around them, are now able to approach each other close enough so that attractive van der Waals forces can hold them together. These latter intermolecular forces are electrostatic in nature as well and arise from four major contributions of which by far the most dominant results from

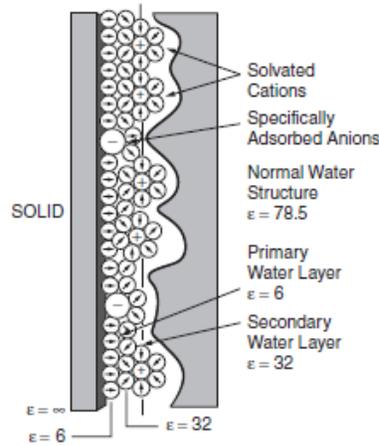


Figure 1.2: Position and orientation of water molecules in the electrical double layer and their role on electrical permittivity. *Source:* from [1]

the interaction of two non-overlapping atoms; the electron cloud of each atom is in fact attracted to the nucleus of the other with an intensity greater than the repulsion due to adjacent electron clouds and nuclei, resulting in a net attractive force of short range. Although reduction of the surface charge does certainly occur and may in some instances be the dominant destabilization mechanism, there are other phenomena of as great an importance, such as the adsorption of polymeric ions and molecules onto colloidal surfaces and the enmeshment of colloids into precipitates. As for the former, although electrostatic considerations play a role during adsorption as well (i.e. charge neutralization), the most relevant contribution is provided by the so-called *bridging* mechanism, which consists in the adsorption of long-chain polymeric species onto the particle surface that later on may bring them together. Finally, the latter mechanism is mostly important for hydrophilic colloid dispersions, as it is likely that the extensive hydration shell, typical of hydrophilic particles, acts like a physical "shielding" that promotes their stability; destabilization is in this case brought by dehydration and coordinative chemical reactions between the applied coagulant and functional groups on the colloid surface.

From the above discussion it is clear that the destabilization of colloids in aqueous media is a complex process that arise from several mechanisms acting simultaneously. Whichever is the predominant, it mostly depends on the following factors [1]:

- the nature of the colloidal dispersion and whether hydrophobic or hydrophilic particles are prevalent, as well as the intensity of surface charge carried by the colloid;
- the type of coagulant added to the colloid dispersion, that is, whether coagulant species are uncharged or charged (and their charge intensity), the adsorptive capacity of the species and the capacity for bridge formation between adjacent colloids.

The surface charge influences the distribution of nearby ions in the liquid: *counter-ions* are attracted towards the surface, while *co-ions* are repelled away from it. Considering at the same time the mixing tendency provided by thermal motion and the mutual ionic repulsion/attraction, we assist to the formation of an *electrical double layer* made up of the charged surface and a neutralizing excess of counter-ions over co-ions distributed in a diffuse manner within the nearby liquid. Electrical double layer (EDL) theory deals with this distribution of ions and hence with the magnitude of the electric potentials occurring in the proximity of the charged surface [1]. EDL can generally be regarded as consisting of two regions: an inner region including water molecules and adsorbed (hydrated) ions, and a diffuse region in which ions are distributed according to the influence of electrical forces and random thermal motion. Historically, the EDL was first considered as a charged surface and a diffuse region of ions around it (Guoy and Chapman model); then, in addition to the latter, a second region of ions held to the surface (i.e. adsorbed) was considered (Stern model). The interfacial region is therefore viewed as one or two layers of molecules that are mostly covalently bound to individual surface sites, surrounded by a layer of aqueous solution wherein the ionic concentrations are affected by electric potential, even though these ions are not bound directly to surface sites.

Adsorption of a molecule (or ion) from solution onto the surface of a solid involves the following sequence of events (Stumm, 1962):

1. Removal of the molecule from solution;
2. Removal of adsorbed solvent from the solid surface;
3. Attachment of the molecule onto the solid surface.

Because of thermal agitation, molecules and ions in solution are in a continuous state of movement, which leads to an haphazard bombardment and consequent random motion of the suspended colloidal particles (Brownian motion) that results in a continuous series of collisions and contact made between molecules and ions in solution and the particle surface. When such contact occurs, three events are possible:

1. molecules and ions could stick to the solid surface via ionic, covalent, hydrogen or dipolar bonding (*chemisorption*); molecules and ions that adsorb in this manner must lose at least the water of hydration on the side nearest the solid, and the surface complexes formed by such reactions are relatively strong and referred to as *inner sphere complexes* (H^+ and OH^- are always presumed to bind in this way). Other ions might also bind to the surface via chemical bonds, but without losing any waters of hydration; such ions are therefore separated from the surface by a water molecule and form weaker complexes (sometimes called *ion pair* or *outer sphere complexes*) with the surface. At distances farther from the surface, ions are assumed to be attracted to or repelled from the surface solely by electrostatic interactions (i.e. not by chemical bonding), therefore adsorption could also arise from London-van der Waals forces of attraction alone (*physisorption*); electrical interactions locally increase the concentration of species that are oppositely charged from the surface (*counter-ions*) and decrease that of like-charged species (*co-ions*), compared to their concentrations in bulk solution. Over a distance of several molecular diameters, the net charge of these counter-ions and co-ions is expected to completely neutralize the particle *net* surface charge (i.e. including the contribution of inner and outer sphere complexes), to progressively approach a net charge of zero in bulk solution (i.e. at infinite distance from particle surfaces).
2. The molecule could be electrostatically repelled from the surface before event (1) can take place.
3. Even if conditions are favourable for attachment, the molecule could once more be transferred in solution because of natural Brownian motion.

A former definition of the *diffuse layer* might then be "*the region in which the concentrations of the nonspecifically adsorbed ions differ from their values in bulk solution*". Affinity of solid surfaces for some solutes may be less than the solid affinity for

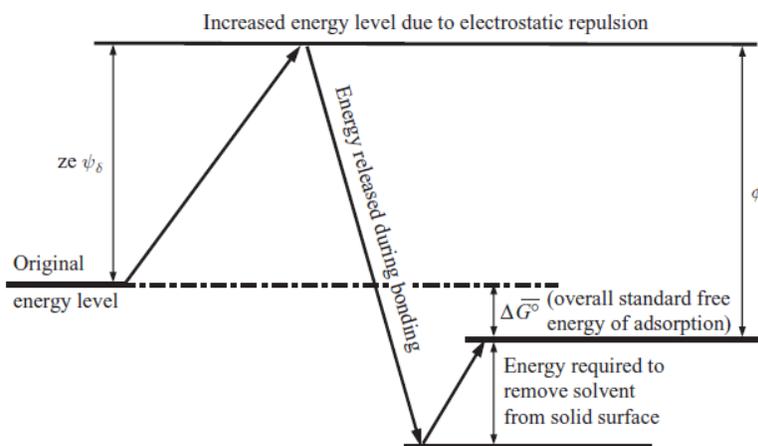


Figure 1.3: Energy changes during adsorption. *Source:* from [1]

the solvent itself. Organic dipoles and large organic ions are preferentially accumulated at solid surfaces mainly because their hydrocarbon parts have little affinity for water (in case of an aqueous solvent). Simple inorganic ions, such as Na^+ , Ca^{2+} and Cl^- , may remain in solution even if they're specifically attracted to the solid surface, because they are readily hydrated in aqueous solutions. Generally, less hydrated ions and many anions tend to seek positions at the solid surface to a larger extent than easily hydrated ions. If adsorption does take place in solution, there is a net release of energy since the energy of attraction between the solid surface and the adsorbate must be greater than the energy of repulsion (if present) due to electrostatic effects: since adsorption involves the removal of a certain amount of solvent from the solid surface, it follows that the energy required to remove the solvent molecules must be less than the energy released by the attachment of the adsorbate. With a very few exceptions, adsorption is generally an exothermic process, because dissolved molecules and ions greatly reduce their entropy when adsorbed/held onto solid surfaces (in order to be a spontaneous process, the change in Gibbs free energy must be negative, therefore heat must be released during interphase mass transfer). Energy released as a consequence of adsorption is called *overall standard (molar) free energy of adsorption*, $\Delta\bar{G}^0$. It is equal to the sum of the energy released as a consequence of whatever forces of attraction prevail (either chemical bonding or van der Waals forces) and either the energy released by the satisfaction of electrostatic forces of attraction or energy to be overcome because of electrostatic forces of repulsion (fig. 1.3). The overall standard molar free energy of adsorption is given by:

$$\Delta\bar{G}^0 = \phi \pm ze\psi_\delta \quad (1.1)$$

where ϕ is the energy released during bonding and $ze\psi_\delta$ is the energy either released or absorbed during electrostatic interactions; if the latter is of opposite sign to ϕ yet is smaller in magnitude, the adsorption takes place quite readily. The standard molar Gibbs (free) energy of reaction, $\phi = \Delta\bar{G}_r^0$, is linked to the *intrinsic binding constant of adsorption* by the following relationship:

$$K_{intr} = \exp\left(-\frac{\phi}{RT}\right) \quad (1.2)$$

which describes the role of chemical affinity between adsorbate and adsorbent during an adsorption reaction. K_{intr} characterizes the species distribution at equilibrium if no other factors (such as electrostatics) affected the binding strength of the adsorbate to the adsorbent. As mentioned above, electrostatic attraction or repulsion alters the net energy change associated with the adsorption reaction and thereby alters the overall binding constant. Experimentally, it is difficult to separate the energy of adsorption into its chemical and electrostatic components, but it is known that many organic ions of similar charge to the surface are readily adsorbed, indicating that energy released in chemisorption may easily be greater than the electrostatic energy of repulsion. Counter-ions not only decrease the effective surface charge, but at higher concentrations may eventually produce a charge reversal. The adsorptivity of a particular system is usually described by empirical or semi-empirical expressions, known as *adsorption isotherms*, which relate the mass (or moles) of solute adsorbed per unit mass (or surface) of adsorbent to the concentration of solute remaining in solution after adsorption (i.e. at equilibrium). Because of their simplicity and generalizability, two commonly used isotherms are the Langmuir and Freundlich isotherms; for systems exhibiting a Langmuirian adsorption behaviour, the relevant parameters fit the equation:

$$\frac{C}{(x/m)} = a + \frac{C}{b}, \quad (1.3)$$

where C is the solute concentration remaining in solution after adsorption (mg/l) and (x/m) is the mass of solute adsorbed per unit mass of adsorbent. Similarly, systems

exhibiting a Freundlichian adsorption behaviour are described by:

$$\log(x/m) = (1/a) \log C + \log b \quad (1.4)$$

Since both isotherms represent two-parameters models, a and b are treated as modeling constants to be fitted with experimental data which are actually defined at a particular temperature and for the particular system (adsorbate-solvent-adsorbent) under consideration. Temperature, solution pH and multiple species adsorption have to be carefully estimated during lab tests because they are three of the most important factors affecting adsorption. Freundlich adsorption behaviour has been reported in the destabilization of dilute clay suspensions by hydrolyzed aluminium sulphate (Black & Chen, 1967), while Langmurian behaviour has often been observed for hydrolyzed iron (Fe III) coagulants and polyelectrolytes (Stumm & O'Melia, 1968; Black *et al.*, 1965; Committee Report, 1971). The difference between the adsorption behaviour of aluminium coagulants compared to ferric coagulants and polyelectrolytes is related to the complexity of the aluminium hydrolysis products formed during coagulation, which strongly depend on the coagulation pH.

1.1.3 The Electrical Double Layer: Guoy-Chapman theory and the Stern Model

While Guoy-Chapman theory assumes point charges in the electrolyte solution, Stern recognized that the finite size of ions will limit the inner boundary of the diffuse part of the double layer, since the centre of an ion can only approach the surface to within its hydrated radius without becoming specifically adsorbed. Stern proposed a model in which the DL is represented by two regions separated by a layer between the so-called *Stern plane* (which is located at about a few hydrated ion radii from the surface) and the so-called *shear plane*, which is the plane that is conventionally assumed to divide those molecules that are rigidly bound to the particle surface from those that are still electrically attracted to the surface, but are free to move to some extent. The adsorbed ions may be dehydrated in the direction of the solid surface, so their centres will lie somewhere between the solid surface and the Stern plane. In its original formulation, ions whose centres lie beyond the Stern plane are considered in Stern's model as part of the diffuse layer, for which Guoy-Chapman theory is applicable. With respect to fig. 1.4,

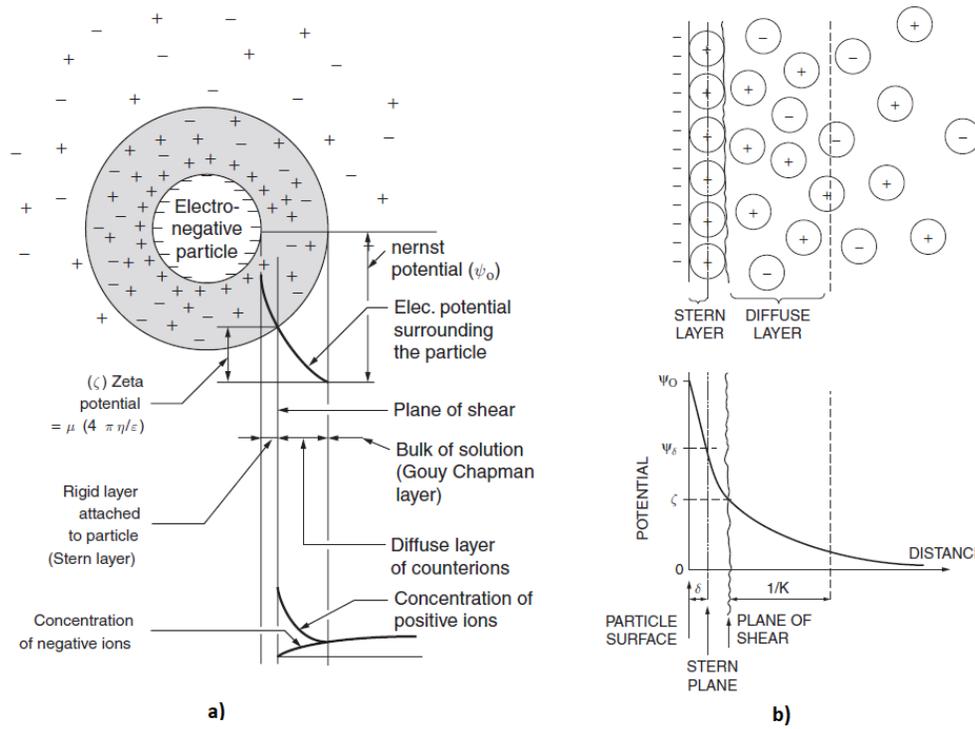


Figure 1.4: Conceptual representation of the Stern's model of the EDL (a, from Priesing,1962) and counter-ion vs. co-ion adsorption in the Stern Model (b). *Source:* from[1]

ψ_0 is the so-called *Nernst potential* (the electric potential at the particle surface), while ψ_δ is the potential at the Stern plane (the so-called *Stern potential*). The potential in the diffuse layer is assumed to decay exponentially from ψ_δ at the Stern plane to zero at infinity (i.e. in bulk solution). When physisorption takes place, counter-ion adsorption generally predominates over co-ion adsorption and charge reversal may take place within the Stern layer when using polyvalent or surface active counter-ions (i.e. ψ_0 and ψ_δ have opposite signs). On the other hand, adsorption of surface active co-ions could eventually create a situation in which $|\psi_\delta| > |\psi_0|$ and $sign(\psi_0) = sign(\psi_\delta)$ (fig. 1.5, b). As mentioned above, even if the electrostatics are unfavorable, the overall reaction could be significantly favorable, if the intrinsic chemical attraction of the metal (or species) for the surface site is strong.

Guoy-Chapman theory introduces the following assumptions:

- The surface is assumed to be flat, of infinite extent and uniformly charged;

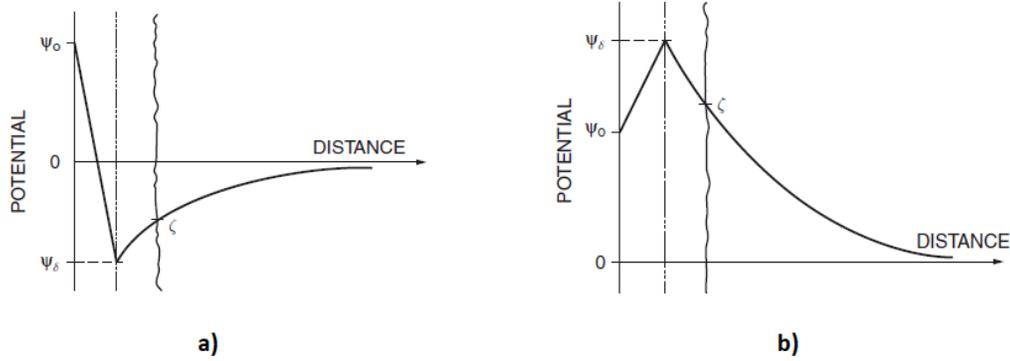


Figure 1.5: Charge reversal within the Stern layer due to polyvalent or surface-active counter-ions (a) and adsorption of surface-active co-ions increasing effective surface charge (b). *Source:* from [1]

- Ions in the diffuse part of the DL are assumed to be point charges distributed according to the *Maxwell-Boltzmann distribution*, which is based on probability theory and leads to an expression describing the number of molecules (ions) in a gas (liquid) at equilibrium that possess a certain amount of energy.

The Maxwell-Boltzmann distribution considers a system made up of n_0 similar "inert" molecules (i.e. that do not attract or repel each other) which possess, in a statistical sense, the following energy distribution:

$$n_i = n_0 \exp\left(-\frac{E_i}{k_B T}\right),$$

where n_i is the number of molecules with energy E_i at absolute temperature T and $k_B T$ is the Boltzmann's constant times the absolute temperature, which is a measure of the average thermal kinetic energy of the system (it is derived from entropy and probability considerations). Finally, the liquid is assumed to influence the double layer only through its permittivity, which is assumed to have the same value throughout the diffuse part of the double layer. Let ψ_δ be the electric potential at the Stern plane and ψ the electric potential at a generic distance x (in the solution) from the Stern plane. Taking the solid surface to be negatively charged and considering at the same time a counter-ion adsorption without charge reversal (which is the usual case), and an excess of counter-ions over co-ions in the diffuse layer, the number of positive and negative

ions are determined by applying the Maxwell-Boltzmann distribution:

$$n_+ = n_0 \exp\left(\frac{ze\psi}{k_B T}\right), \quad n_- = n_0 \exp\left(-\frac{ze\psi}{k_B T}\right) \quad (1.5)$$

where n_+ and n_- are, respectively, the number of positive and negative ions per unit volume at points where the potential is ψ , and n_0 is the concentration of ions at infinity (where $n_0 = n_+ = n_-$). With respect to equation (1.5), z is the *ion valency*, e is the *electron charge*, and hence ze is the ion charge. By definition, ψ is the work done in moving a unit charge (of the same sign as that of the surface) from infinity to the point where the potential ψ is measured: if a charge of ze is moved from infinity (i.e. from bulk solution, where the electric potential is assumed to be zero) to a point, the amount of work done on it is then ψze , which corresponds also to the *potential energy* possessed by the charge at that point. The (net) charge density ρ , at points where the potential is ψ , is given by

$$\rho = ze(n_+ - n_-) = zen_0 \left[\exp\left(\frac{ze\psi}{k_B T}\right) - \exp\left(-\frac{ze\psi}{k_B T}\right) \right] = 2zen_0 \cdot \sinh\left(\frac{ze\psi}{k_B T}\right) \quad (1.6)$$

In the proximity of the particle surface, a *flat plane* approximation for the electrostatic potential may be sufficient, therefore the following Poisson's equation can be easily derived from Coulomb's law:

$$\frac{d^2\psi}{dx^2} = -\frac{\rho}{\varepsilon}, \quad (1.7)$$

where ε is the permittivity (or *dielectric constant*) of the aqueous medium. From (1.5) and (1.7):

$$\frac{d^2\psi}{dx^2} = -2 \left(\frac{zen_0}{\varepsilon}\right) \cdot \sinh\left(\frac{ze\psi}{k_B T}\right) \quad (1.8)$$

The boundary conditions for this equation are:

- $\psi = \psi_\delta$ when $x = 0$;
- $\psi = 0$ and $\frac{d\psi}{dx} = 0$ when $x \rightarrow \infty$,

which lead to the following exact solution:

$$\psi = 2 \frac{k_B T}{ze} \cdot \ln \left[\frac{1 + \gamma \exp(-Kx)}{1 - \gamma \exp(-Kx)} \right], \quad (1.9)$$

where

$$\gamma = \frac{\exp\left(\frac{ze\psi_\delta}{2k_B T}\right) - 1}{\exp\left(\frac{ze\psi_\delta}{2k_B T}\right) + 1}, \quad K = \sqrt{\frac{2z^2 e^2 n_0}{\epsilon k_B T}}, \quad (1.10)$$

and for which ion concentrations at a distance x from the surface (envisioned as a flat plane) are given by:

$$c_i(x) = c_{i,bulk} \exp(-z_i \tilde{\psi}(x)), \quad (1.11)$$

where $\tilde{\psi}(x) = e\psi/k_B T$ is a dimensionless potential, often referred to as the *reduced potential*. It can be shown that for low potentials, eq. (1.9) can be simplified using a first order approximation of both the exponentials and the natural logarithm appearing in (1.10) and (1.9) respectively,

$$\exp\left(\frac{ze\psi_\delta}{2k_B T}\right) \simeq 1 + \frac{ze\psi_\delta}{2k_B T} \iff \frac{ze\psi_\delta}{2k_B T} \ll 1 \implies \gamma \simeq \frac{ze\psi_\delta}{4k_B T}$$

Equation (1.9) therefore simplifies to:

$$\begin{aligned} \psi &= \frac{2k_B T}{ze} \cdot [\ln(1 + \gamma \exp(-Kx)) - \ln(1 - \gamma \exp(-Kx))] \simeq \frac{4k_B T}{ze} \cdot \gamma \exp(-Kx) \\ \psi &\simeq \frac{4k_B T}{ze} \cdot \frac{ze\psi_\delta}{4k_B T} \exp(-Kx) = \psi_\delta \exp(-Kx) \end{aligned} \quad (1.12)$$

Equation (1.12), also known as *Debye-Huckel model* clearly shows that at low potentials, the potential decays exponentially with distance from the Stern plane. Close to the charged surface, where the potential is likely to be relatively high, the above approximation cannot be used and the potential increases at a rate greater than exponential [1]. The Debye-Huckel model is a reasonable approximation when the surface potential is lower than $k_B T/ze$, which for $z = 1$ and $T = 298K$ is equal to 25.7 mV. The Stern potential ψ_δ can be further related to the *net space charge density* σ in the diffuse part of the double layer as follows:

$$\sigma = \int_0^\infty \rho dx = \int_0^\infty \left(2ze n_0 \cdot \sinh \frac{ze\psi}{k_B T} \right) dx$$

Using a first order approximation (valid for low potentials) of the integrand ($\sinh x \simeq x$

for $x \rightarrow 0$):

$$\sigma = \frac{2z^2 e^2 n_0}{k_B T} \int_0^\infty \psi dx \quad (1.13)$$

and introducing equation (1.12) and the second expression of (1.10) in (1.13), it follows that:

$$\sigma = K\varepsilon\psi_\delta \iff \psi_\delta = \frac{\sigma}{K\varepsilon} \quad (1.14)$$

The Stern potential ψ_δ (at low potentials) is therefore linearly dependent on the net space charge density (σ) and inversely proportional on the permittivity of the diffuse layer (ε), and on a quantity K^{-1} whose units are that of a length. The physical meaning of $K\varepsilon = \varepsilon/K^{-1}$ is that of a *capacitance* of a unit area parallel plate condenser: the diffuse layer may therefore be interpreted as a parallel plate condenser with permittivity ε and plates separated by a distance of $1/K$ (sometimes inappropriately referred to as the *thickness* of the double layer), commonly known as *Debye-Huckel length*. The potential ranges in fact from ψ_δ at the Stern plane to zero at infinity, therefore the diffuse layer should theoretically possess an infinite thickness. The Debye length K^{-1} is also defined as the distance needed for ψ to decay to ψ_δ/e , and can be considered as the characteristic length of the diffuse layer. In the Debye-Huckel model, the concentration of a species with charge number z varies with distance (x) away from a planar surface according to the following equation:

$$c_i(x) \simeq c_{i,bulk} \left[1 - z_i \frac{e\psi(x)}{k_B T} \right], \quad (1.15)$$

which is again valid only for low values of ψ_δ , but it holds for both co- and counter-ions (the term in brackets is <1 for co-ions and >1 for counter-ions). From (1.14) and (1.10), it follows that the Debye-Huckel length (K^{-1}) depends on the inverse square root of the *ionic strength* (I) of the electrolyte solution (i.e. increasing the ionic strength results in a decrease in the thickness of the double layer):

$$K = \sqrt{\frac{2z^2 e^2}{\varepsilon k_B T} n_0} = \sqrt{\frac{2z^2 e^2}{\varepsilon k_B T} \frac{N_A C_0}{10^3}} = \sqrt{\frac{4e^2}{\varepsilon k_B T} \frac{N_A I}{10^3}} \quad (1.16)$$

where $N_A \simeq 6.03 \cdot 10^{23} \text{ moles}^{-1}$ is the *Avogadro's number*, $k_B \simeq 1.38 \cdot 10^{-23} \text{ JK}^{-1}$ and C_0 is the electrolyte concentration (if considering a single ion species, *moles l*⁻¹). For a

generic (multivalent) electrolyte solution composed of M different ion species, the ionic strength is defined as:

$$I = \frac{1}{2} \sum_{i=1}^M C_i z_i^2$$

Typical values of K^{-1} in natural waters range from $\sim 10 \text{ nm}$ in fresh water with an ionic strength of $10^{-3} \text{ mol/L NaCl}$ to $\sim 0.4 \text{ nm}$ in seawater [2, p. 317]. Another consequence of the Guoy-Chapman model is that multivalent counter-ions are concentrated in the double layer to a much greater extent than monovalent ions, in fact:

$$n_+ = n_{01} \cdot \exp\left(\frac{ze\psi}{k_B T}\right) = n_{02} \cdot \exp\left(\frac{2ze\psi}{k_B T}\right) = n_{03} \cdot \exp\left(\frac{3ze\psi}{k_B T}\right),$$

therefore,

$$\frac{n_{02}}{n_{01}} = \exp\left(-\frac{ze\psi}{k_B T}\right), \quad \frac{n_{03}}{n_{01}} = \exp\left(-2\frac{ze\psi}{k_B T}\right)$$

While the Guoy-Chapman theory considers the diffuse part of the EDL alone, Stern's complete model for the EDL takes additionally into account charged and/or uncharged species eventually adsorbed to the particle surface (i.e. within the Stern layer), and is based on considerations of electrical neutrality of the whole system; Stern potential ψ_δ must in fact change accordingly to the the number and charge of adsorbed species because ion adsorption and electric potential are implicitly linked to one another (the extent of ion adsorption affects the electric potential at the interface, and vice versa). If σ_0 is the charge density at the particle surface, then treating the Stern layer as a condenser as well of thickness δ and permittivity ε' leads to:

$$\sigma_0 = \frac{\varepsilon'}{\delta} \cdot (\psi_0 - \psi_\delta) \quad (1.17)$$

Remembering that both the chemical energy of adsorption ($\phi = \Delta\bar{G}_r^0$) and the net energy change associated with electrostatic interactions ($ze\psi$) are additive,

$$K_{overall} = \exp\left(-\frac{\Delta\bar{G}_r^0 \pm ze\psi}{RT}\right) = \exp\left(-\frac{\Delta\bar{G}_r^0}{RT}\right) \exp\left(\pm\frac{ze\psi}{RT}\right) = K_{intr} K_{el}, \quad (1.18)$$

Stern assumed that a Langmuir type adsorption isotherm could be used to describe the equilibrium between ions adsorbed within the Stern layer and those in the diffuse part

of the DL [1]. Considering an adsorption of counter-ions alone, the charge density of the Stern layer is given by the following expression

$$\sigma_\delta = \frac{\sigma_m}{1 + \left(\frac{N_A}{n_0 V_m}\right) \cdot \exp\left(\frac{ze\psi_\delta + \phi}{k_B T}\right)}, \quad (1.19)$$

where:

σ_δ is the charge density of the Stern layer;

σ_m is the charge density of the Stern layer corresponding to a *complete monolayer of counter-ions*;

V_m is the molar volume of the solvent;

ϕ is the molar energy of adsorption due to chemical forces;

$ze\psi_\delta$ is the molar energy of adsorption due to electrostatic forces.

As already mentioned (fig. 1.5), it should be noted that if the electric potential at the location of adsorption has the same sign as the ion charge, then $K_{el} < 1$, so electrostatic interactions decrease the overall adsorption equilibrium constant. Conversely, if the adsorbate charge and the electric potential have opposite signs, then $K_{el} > 1$, and electrostatic interactions enhance adsorption. Finally, it is worth noting that whereas K_{intr} is a constant for a given adsorbate-adsorbent couple, K_{el} is *not* a constant, because it depends on the electric potential near the surface and therefore is sensitive to details of the entire system, not just on the reacting species alone. If σ is the charge density of the diffuse part of the double layer, then from eq. (1.14) $\sigma = K\varepsilon\psi_\delta$ and hence, for the overall electrical neutrality throughout the DL,

$$\sigma_0 + \sigma_\delta + \sigma = 0 \iff \frac{\varepsilon'}{\delta} \cdot (\psi_0 - \psi_\delta) + \frac{\sigma_m}{1 + \left(\frac{N_A}{n_0 V_m}\right) \cdot \exp\left(\frac{ze\psi_\delta + \phi}{k_B T}\right)} + K\varepsilon\psi_\delta = 0 \quad (1.20)$$

Equation (1.20), which represents the Stern's complete model for the EDL, contains a number of unknown quantities of which the most significant is by far the Stern potential ψ_δ . Although ψ_δ cannot be directly measured, a close approximation may be obtained by determining, from electrokinetic measurements (like the *electrophoretic mobility* of

the suspended particles when subjected to an electric field), the so-called *zeta potential* (ζ in fig. 1.4 b), which is the potential at the shear plane. A few devices capable of directly or indirectly measure the zeta potential already exist, although they may be of limited use according to the raw quality of the water/wastewater to be treated, as it will be described in more depth in chapter 3.

For the sake of clarity, one last and commonly used model for the EDL that is worth mentioning because of its popularity and completeness (yet retaining the underlying simplicity of the Stern Model) is the so-called *Triple Layer Model* (TLM), which mathematically introduces a further distinction between the inner and outer sphere complexes and molecules and ions that lie between the Stern and the diffuse layer, with respect to the role they play on the decaying of the electric potential between the particle surface and the shear plane. Since the literature can be confusing on this topic when extending the definitions used so far to a triple layer model, a new notation is introduced for the characteristic planes and their corresponding potentials. With respect to fig. 1.6, the *surface plane* (sometimes called the *naught plane* or *o-plane*) is supposed to run through the inner sphere complexes, while the so-called *beta plane* is defined to pass through the outer sphere complexes, leading to a richer representation of the Stern layer, for which the Stern Model assumes one plane only (the Stern plane). The Stern layer is thus re-envisioned as a *compact layer* consisting of a surface and a beta layer that contains the hydration waters of both of them. Nernst and Stern potentials, previously called ψ_0 and ψ_δ , will henceforth be called ψ_o and ψ_β , respectively. The charges shown on species in the *o-plane* represent the charge from that species plus one (a +1 charge is in fact provided by each Fe-atom adsorption site): while some molecules (ions) are presumed to bind directly to the surface oxide ions (like Ni^{2+} in fig. 1.6), others can replace surface oxide ions to bind directly to the underlying Fe ions (it might occur with anionic species, like $HAsO_4^{2-}$ in fig. 1.6). Finally, the *d-plane* appearing in fig. 1.6 has the same meaning of the shear plane in Stern's model (and hence $\psi_d = \zeta$).

While the charge-potential relationship in the diffuse layer is very well agreed, that is not the case for the compact layer, for which the most common modeling approach is again to assume a linear decaying of the potential, but with different slopes across the β plane:

$$C_1 = \frac{\psi_o - \psi_\beta}{\sigma_o}, \quad C_2 = \frac{\psi_\beta - \psi_d}{-\sigma_d} = \frac{\psi_d - \psi_\beta}{\sigma_d}, \quad (1.21)$$

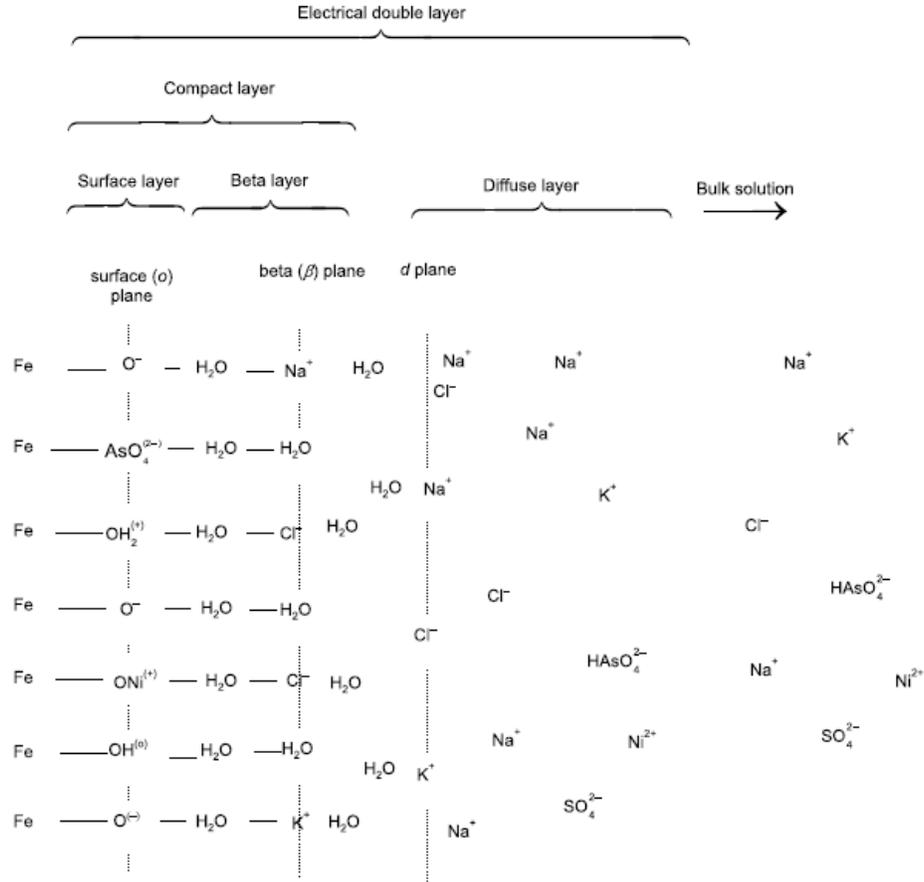


Figure 1.6: Schematic of the interfacial structure near an iron oxide surface, as envisioned in the Triple Layer Model. *Source:* from [2, p. 316]

where C_1 and C_2 are the capacitances of the inner (o -to- β) and outer (β -to- d) layers respectively. Because of the introduced idealizations (e.g. the alignment of adsorbates along planes, the assumption of charge uniformly distributed within them, etc.) and the impossibility to effectively measure them, C_1 and C_2 are usually treated as fitting parameters to improve the match of model calculations to experimental data. It should be finally noted that, as for σ_δ in the Stern Model, σ_d is in reality an *equivalent charge density*, because it does not refer to charges that physically reside *in* the d -plane, but rather to the net charge throughout the diffuse layer, treated *as though* it were all located in the d -plane.

To summarize, the key difference between the Stern and the Triple Layer Model lies in the starting point of the exponential decay of the electric potential, which in

the latter case is assumed to start from the shear plane rather than from the Stern plane. Regardless of the model used, the overall charge around an adsorbent particle *at equilibrium* (including the fixed charge, that of specifically adsorbed ions and the charge in the diffuse layer) must always be zero. Thus, if an ion adsorbs, other processes must combine to expel some like-charged species from the interfacial region and/or bring oppositely charged species closer, so that the overall electroneutrality of the particle and surrounding fluid is maintained. In other words, sorption of an ion is always accompanied by either adsorption of an oppositely charged ion or desorption of a like-charged ion, it never involves exchanges exclusively with water molecules. If the ions participating in the adsorption and desorption reactions involved in an overall sorption reaction bind at different distances from the surface (e.g. if the adsorbing ion formed an inner sphere complex and the desorbing ion an outer sphere complex), the electrical contributions to the adsorption and desorption reactions would generally not cancel (like in *ion exchange* reactions), and thus the surface potential would have a net effect on the reaction. This scenario is typically thought to be the case when strongly binding metal ions, which are usually assumed to bind in the surface or the beta plane, adsorb [2, p. 319]. The importance of electrical interactions in the adsorption of ionic adsorbates onto charged surfaces is hence remarked, as they can increase (counter-ions) or decrease (co-ions) the overall equilibrium constant of adsorption. As already mentioned, the magnitude of these effects (i.e. the magnitude of K_{intr} in $K_{overall}$) depends on the system as a whole, not just on the identity of adsorbates and particles, because the locations where ions bind (or which they vacate) play a role as well in the overall sorption reaction.

1.1.4 Particle stability in terms of the Double Layer

Derjaguin and Landau first (1941), and Verwey and Overbeek seven years later independently developed a quantitative theory based on the Guoy-Chapman model in which the stability of hydrophobic colloids is treated in terms of the energy changes which take place when particles approach one another. The theory, also known as *DLVO theory*, involves estimations of the energy of attraction due to London-van der Waals forces and the energy of repulsion due to the overlapping of electrical double layers in terms of the distance between particles. London-van der Waals forces are thus assumed to be the only phenomenon able to counteract the repulsion that comes when

particles get quite close to one another. At sufficiently short separation distances, these attractive interactions are stronger than the electrical repulsion described above and hence, if one particle approaches another particle (or surface) with sufficient momentum, it can overcome the repulsive interaction and get close enough to the other surface for the attractive forces to dominate.

In the diffuse layer, ions with charge of the same sign as the particle are less concentrated than in the bulk solution, whereas counter-ions are more concentrated. When two particles approach one another, the electrical interactions between them become important at separation distances that are much larger than the thickness of their compact layers. As a result, the electrical interactions are controlled largely by the electric potential profile through their respective diffuse layers, and hence distinctions among the different models for the electrical profiles through the compact layers previously introduced are not of paramount importance. That is the reason why it is common to consider adsorbed species to be part of the solid particles and to define the interface between particles and solution to be at the beginning of the diffuse layer (i.e. at the d -plane) when studying particle behavior in water (and that explains why ψ_d is sometimes referred to as the surface potential). When the Debye-Huckel approximation for low potentials is used, the potential decays away from a spherical surface as follows:

$$\psi(x) = \psi_d \frac{a_p}{r} \exp[-K(r - a_p)] = \psi_d \frac{a_p}{a_p + x} \exp[-Kx], \quad (1.22)$$

where a_p is the radius of the spherical particle (including any adsorbed layers, i.e. up to the d -plane) and r is the distance measured from the center of the particle, so that x is the distance measured from the d -plane, as already assumed. Comparing eq. (1.22) to eq. (1.12), we can see from their pre-exponential terms that the potential decays more rapidly away from a sphere than a plane. As the particles approach one another, the electrostatic repulsion increases, and this repulsive interaction can be quantified as an *increase* in the electrical potential energy of the system. Similarly, as explained subsequently, the attractive van der Waals interactions between the particles can be quantified as a *decrease* in the electrical potential energy of the system. Since energy must be conserved, increases in potential energy are always counterbalanced by an offsetting change in the particles' kinetic energy (i.e. particles slow down as they

approach one another).

The potential energy associated with the interaction of the EDL of the two particles, V_R , is measured relative to a baseline condition in which those interactions do not take place ($V_R = 0$), a situation that theoretically would require an infinite separation distance. However, in practice, it can be defined as a separation that is large enough that, at every point along a line between the two particles, the electric potential of at least one of them is negligible. The derivation of V_R as a function of the separation distance (s) is complex and can be computed as the work (energy) required to bring the particles from the baseline condition ($V_R = 0$) to that separation distance s ; however, for two identical spheres immersed in a symmetric electrolyte solution an expression is given by Lyklema (1978), as corrected by Stumm and Morgan (1996):

$$V_{R,ss} = 64\pi \frac{n_{bulk} k_B T}{K^2} \frac{a_p^2}{s + 2a_p} \left[\tanh \left(\frac{z\tilde{\psi}_d}{4} \right) \right]^2 \exp(-Ks), \quad (1.23)$$

where the subscript "ss" stands for sphere-sphere interaction, $n_{bulk} = c_{bulk} N_A \cdot (1000 \text{ } \text{lm}^{-3})$ is the number concentration of anions (or cations) in the bulk solution (e.g. *no./m*³) and $k_B = R/N_A$. According to this equation, the repulsive energy diminishes with separation distance s as a combination of an inverse function (the second fraction on the right hand side, due to the particle curvature) and an exponential decay, identical to the flat plate case. Hence, the repulsive interaction is a function of both the surface potential (ψ_d) and the solution conditions (K). Since K is proportional to $\sqrt{n_{bulk}}$, the term $n_{bulk} K^{-2}$ is independent of ionic strength, therefore ionic strength primarily affects the exponential term. Conversely, the charge on the electrolyte (z) affects every term in the equation since K is proportional to z , therefore the first fraction decreases too with increasing z value. In other words, *given two solutions with the same ionic strength, but one made with monovalent electrolytes and the other with polyvalent electrolytes, the repulsion between particles is smaller at any distance in the latter case.*

The electron cloud around any molecule is in constant motion, deviating slightly from its most stable configuration first in one direction, then in another. A force opposing these disturbances and attempting to drive the electron cloud to its lowest energy state always exists, but it is contrasted by the momentum of electrons that constantly overshoots this lowest energy distribution, keeping the molecule shifting from

one slightly unstable state to another (and the higher the molecular weight, the greater the momentum of the electron cloud). As a result, even in highly symmetric, nonpolar molecules, the electron distribution is always slightly asymmetric, making them actually slightly polar at any instant and nonpolar only on a time-averaged basis. When two such molecules approach each other quite closely, the electrons in the outer orbit of each molecule affect the motion of electrons in the other. If the instantaneous polarities of the two molecules are the opposite of one another when they first interact, the interaction is favorable (attractive). As a result, the electron cloud of each molecule keeps bouncing endlessly from one side to the other, in such a way that a stable arrangement in which the electron clouds of the two molecules oscillate synchronously is established, making them constantly experience an attractive force toward one another. This attractive force remains in effect until the molecules' kinetic energy causes them to separate. When two *particles* approach each other very closely, the electron clouds of many of the molecules at the surface of each particle will be properly oriented to engage in attractive interactions of the sort described above. If the initial orientation of two molecules (one for each approaching particle) is such that they are not attracted to one another, the electron cloud of each will continue to oscillate independently and randomly. Eventually, if the particles remain close to one another for many oscillations of the electron clouds, an attractive interaction between each pair of molecules on the opposing particles is almost certain to occur. The interactions described above are operative only over very short distances and are known as *London-van der Waals forces* or *dispersion forces*. These attractive forces depend on the *polarizability* of the particles (i.e. how easily the molecules on the particles surface can be polarized by the movement of electrons). For two equi-sized, spherical particles, the energy associated with van der Waals attraction is expressed as follows:

$$V_{A,ss} = -\frac{A_H}{6} \left(\frac{2}{\bar{s}^2 - 4} + \frac{2}{\bar{s}^2} + \ln \frac{\bar{s}^2 - 4}{\bar{s}^2} \right), \quad (1.24)$$

where \bar{s} is the dimensionless separation distance between the particle centers, defined as:

$$\bar{s} = \frac{s + 2a_p}{a_p}, \quad (s = 0 \implies \bar{s} = 2) \quad (1.25)$$

and A_H is the so-called *Hamaker constant*, which is related to the polarizability of the

particles and has a range of values generally considered to be from $3 \cdot 10^{-21}$ to 10^{-19} J , with an average value of 10^{-20} J most commonly chosen.

According to DLVO theory, the total energy of interaction between two charged particles is hence given by the sum of these two contributions:

$$V_{TOT} = V_R + V_A \quad (1.26)$$

The shape of the resulting potential energy - distance curve ($V_R(s)$, fig. 1.7) can be deduced by comparing the analytical expressions of the two forces. While the repulsive energy decays exponentially with the separation distance (s), with a range of the order of the thickness of the electrical double layer (K^{-1}), the attractive energy decreases as an inverse power of the separation distance. Consequently, van der Waals forces will prevail at small separation distances, while at intermediate distances, double layer repulsion may predominate, depending on the relative magnitudes of V_A and V_R . Although not shown in fig. 1.7, at very small separation distances there will be repulsion due to the overlapping of electron clouds known as *Born repulsion*, which prevent particles from interpenetrating each other. The inclusion of this repulsion in the analysis leads to a minimum in the energy curve at some extremely small separation distance which is known as *primary minimum* of the energy curve, and it identifies the most stable distance of separation of the two particles. Figure 1.7 shows two different solution conditions for which the repulsive potential of two colliding particles is less affected in case 1 than in case 2, perhaps because the ionic strength is higher in the latter than in the former (e.g. higher concentrations of ionic species and/or the use of polyvalent cations instead of monovalent cations). Since V_A does not depend on the solution conditions (although adsorption might affect the Hamaker constant), the overall potential of interaction V_{TOT} is substantially affected by the magnitude of V_R . With respect to fig. 1.7, two outcomes are possible:

- case 2: $V_2 = V_A + V_{R2} < 0 \forall s \implies$ the overall potential energy of interaction is always negative and hence particles will always be able to coalesce, regardless of their kinetic energy (*fully destabilized suspension*);
- case 1: $V_1 = V_A + V_{R1} \implies$ the system exhibits a potential energy maximum (A) to be overcome in order for particles to stick together (i.e. to form a floc). If

this potential maximum is large compared to the thermal kinetic energy of the particles ($k_B T$), the system will be stable. If the kinetic energy of the particles is large enough to surmount this *potential energy barrier*, then the particles will coalesce.

The height of this energy barrier depends upon the magnitude of ψ_d and upon the range of the repulsive forces ($\propto K^{-1}$). If we were to think in terms of forces, at separation distances greater than that associated with the energy barrier the slope of the V_{TOT} vs. s plot would be negative, meaning the force is positive and acting to keep the particles apart (force is the negative derivative of energy with respect to distance). At smaller separation distances, the force is negative, meaning the particles would be drawn together. There could be situations in which the net interaction exhibits a *secondary minimum* at separation distances greater than that of the primary minimum. The position and depth of this secondary minimum is much more dependent on V_R than the primary minimum, such that the binding strength of two particles held together at this separation distance is much more sensitive to both the solution conditions and the kinetic energy of the particles, which in some cases may lead to a *restabilization* of the suspension. Binding at the secondary minimum is hence a *reversible* process, whereas binding at the primary minimum could be considered as an irreversible process, with an energy well in the secondary minimum that is relatively small in comparison to that in the primary minimum. As a last important remark, it can be shown that *under identical conditions of surface potential and electrolyte concentration, the effect of particle size on the energy barrier and secondary minimum can be dramatic, suggesting that larger particles are mainly captured and held together by the secondary minimum because the energy barrier that must be overcome to reach the primary minimum is generally higher than that of smaller particles.*[2, p. 529]. While the secondary minimum might still be important in filtration (as amply demonstrated by O'Melia and coworkers in several papers), it is generally mandatory to reach the primary minimum in order to create long-lasting flocs and avoid their destruction (or detachment from filter grains) by hydraulic (shear) forces.

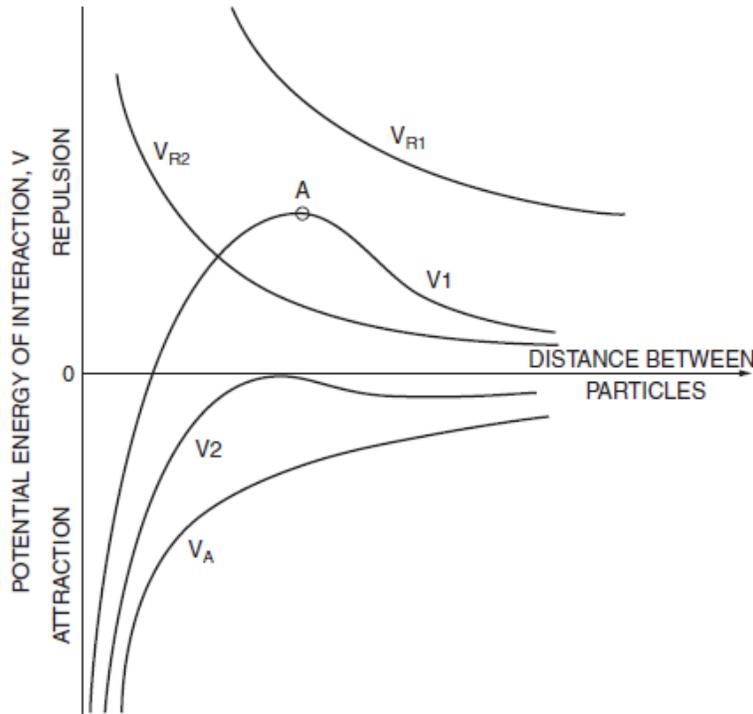


Figure 1.7: Potential energy of interaction between two particles. *Source:* from [1, p. 29]

1.1.5 Experimental Measurements related to Charge and Potential

Measuring the surface charge or potential of suspended particles is of great interest when determining the stability of a solid-liquid dispersion or, in other terms, the degree to which added chemicals change that stability. *Acid/base titrations* of surface charge, *electrophoresis* and measurement of *streaming potential* are the most commonly used approaches to characterize these parameters.

Electrophoresis has been the most common of these measurement methods. When a suspension is placed in an electric field, charged particles tend to move toward the electrode of opposite charge. Each particle rapidly reaches a steady-state velocity that ultimately depends on the applied electrical force, on fluid drag and on the hydrodynamic retardation due to the electrical work done in the solution to move ions around a particle as it moves. The *electrophoretic mobility* is defined as this steady-state velocity (e.g. $\mu\text{m/s}$) divided by the electrical field strength (V/cm); the

sign of the electrophoretic mobility corresponds with the sign of the surface charge. When a particle moves during electrophoresis, the water motion around the particle alters the surrounding diffuse layer, as shown in fig. 1.8. If it did not, the particle would not move because the combination of surface charge, compact layer and diffuse layer must always be electrically neutral. Thus, what electrophoretic mobility measures in reality is a property not only of the particle, but of both the particle and part of the surrounding compact and diffuse layers that move along with it. Research has been focused a lot in finding the location of the shear plane relative to the particle surface, which is often assumed to be at the border between the compact and diffuse layers (regardless of the model used for the compact layer). Its exact location is not known since it depends on the adsorbed ions in the Stern layer and the degree of hydration.

The measured electrophoretic mobility is often converted to the potential at the shear plane (i.e. the *zeta potential* ζ); while various models for the conversion exist (and depend on K), the zeta potential is usually taken as the product of the electrophoretic mobility and a term that depends on the diffuse layer thickness and the dielectric properties and viscosity of the medium (details are provided by Hiemenz and Rajagopalan, 1997 and Hunter, 2001). Zeta potential measurements are also useful in ion adsorption studies and in the characterization of particle surfaces (e.g. the effects of pH, surface active agents and enzymes on the surface of bacteria cells). The usefulness of zeta potential measurements in the control of coagulation-flocculation processes is largely dependent on whether the destabilization mechanism is predominantly due to charge effects.

A third measure of surface charge and/or potential characteristics is the *streaming current*. In streaming current detectors particles are held relatively stationary as fluid flows around them. In modern instrumentation, a piston is oscillated fairly slowly inside a vertical open cylinder of slightly larger diameter; the small annular space between the two cylinders and a small reservoir below are filled with the suspension. Since the piston has a diameter much larger than the annular space, the fluid velocity induced as water escapes from or fills the annular space is far greater than the piston velocity, though that flow is still laminar. Since particles are thought to adhere to the surfaces of both walls, the relatively high fluid velocity near them induces a shear plane which slightly separates particles from their diffuse layers; the net excess of counterions contained in the diffuse

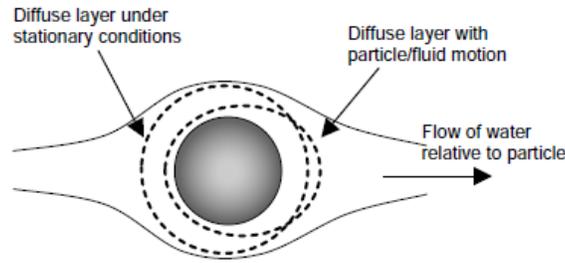


Figure 1.8: Shearing of diffuse layer by flow of water around a charged particle. *Source:* from [2, p. 531]

layer of each particle thus generates an electrical current, which can be measured with electrodes at both ends of the annular space. Streaming current detectors have proven to be valuable on-line instruments to control the dosing of destabilizing chemicals when the mechanism for destabilization involves surface charge neutralization.

1.2 Particle destabilization: destabilization mechanisms

The success of particle treatment processes depends on overcoming (or eliminating) the problem of stability induced by the repulsive energy barrier (i.e. the magnitude of the maximum - or maxima - in the V_{TOT} vs. s plot). Depending on the solution conditions and on the coagulant(s) used, particles can be destabilized in at least four different ways, namely:

1. Compression of the diffuse layer.
2. Adsorption and charge neutralization.
3. Enmeshment in a precipitate.
4. Adsorption and interparticle bridging.

The degree of destabilization achieved is generally a nonlinear function of the concentration of the added chemical, as it will be shown in more detail. The minimum concentration of the added chemical required for effective destabilization is commonly referred to as the *critical coagulation concentration* (CCC).

1.2.1 Compression of the diffuse layer

When the ionic strength of a solution is increased, the availability of counterions to surround a charged particle increases, hence reducing the extent of its diffuse layer. The effect of ionic strength on the composition and thickness of the diffuse layer has already been illustrated, and could be graphically pictured like fig. 1.9 (the higher the ionic strength, the more compressed is the diffuse layer). This reduction in size of the diffuse layer comes with a reduction in the energy barrier that must be overcome to allow particles to collide, effectively promoting the destabilization of the dispersion. At some ionic strength, the net interactive energy becomes attractive at all separation distances (fully destabilized dispersion); since the ionic strength is a function of both the molar concentration and the valence of the dissolved electrolyte, an increase in either of these parameters results in a decrease in the energy barrier. An expression for the (theoretical) optimal concentration of an indifferent (i.e. non adsorbing) electrolyte required to destabilize a stable dispersion (e.g. mostly hydrophobic colloids) could be derived by imposing the condition $V_{TOT}(s) = V_A(s) + V_R(s) = 0$ (net interaction energy of zero) at the location of the energy barrier (where $dV/ds = 0$); to easily find this critical coagulant concentration it is recommended to use an alternative formulation of both $V_A(s)$ and $V_R(s)$ valid for $s \ll a_p$ (Overbeek, 1977 and Verwey and Overbeek, 1948). The diffuse layer model predicts the CCC computed this way to depend on $1/z^6$, which translates to a ratio of 1 : 1/64 : 1/729 (mono-, di-, and trivalent compounds respectively), in close agreement with the empirical *Schultze-Hardy rule* (1890s). In full-scale engineered systems (in particular for drinking water treatment plants), compression of the double layer is not practical as a destabilization mechanism because the treated water would also have a high ionic strength (i.e. a high salt content). Nevertheless, it could be used in situations in which the blending of two waste streams - one with high ionic strength and another with high particle content but low ionic strength (for example) -, could lead to the effective flocculation of the particles without need for additional chemicals. Compression of the double layer has also frequently been employed in research on physical aspects of particle processes since it is one of the best mathematically understood mechanisms among those mentioned above ([2, p. 536]). Compression of the diffuse layer is also important in estuaries, where the salt content increases as the estuary approaches the ocean (the coagulation of particles near the mouth leads to accumulations of sediment near the entrance

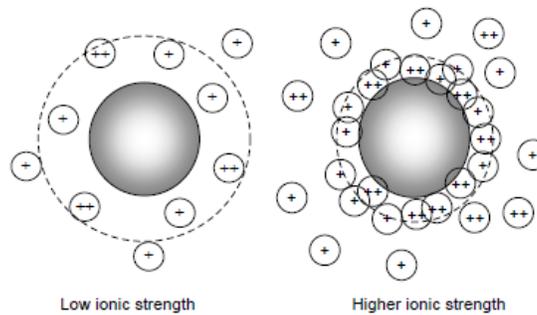


Figure 1.9: Compression of the diffuse layer at high ionic strength (particle is assumed to be negatively charged and the Debye-Huckel length is represented by the dotted circles). *Source:* from [2, p. 535]

to the ocean). As the depth of the secondary minimum in the total energy curve decreases in response to an increase in salt concentration, destabilization might occur to some degree by capture in the secondary minimum at concentrations less than that needed to make the interaction attractive at all distances (i.e. the CCC). Although the Schultze-Hardy rule is inapplicable in most cases (because it does not take into account the extensive adsorption involved during colloid destabilization with metal coagulants), there are some instances where it is experimentally observed, like those where highly charged polynuclear coagulant hydrolysis species are present at low concentrations in solution. At sufficiently low concentrations, adsorption may be negligible and the destabilization mechanism is probably one of double layer depression analogous to that of non-adsorbable ions. At higher coagulant concentrations, adsorption increases and the Schultze-Hardy rule becomes inapplicable (Matijevic, 1973).

1.2.2 Adsorption and charge neutralization

A second destabilization mechanism is the neutralization of particles surface charge by the adsorption of ions (or charged molecules/polymers) of the opposite sign. If just the right amount is adsorbed, the net charge on the particles becomes zero, the particle have no surrounding diffuse layer, electrostatic repulsion is then nonexistent and particles can no longer remain in stable suspension. Adsorption of a charge neutralizing ion (or molecule) can occur because of specific chemical interactions between the destabilizing chemical and the particle surface, or because the molecule is hydrophobic and prefers to attach to a particle surface rather than stay in the bulk solution (higher affinity

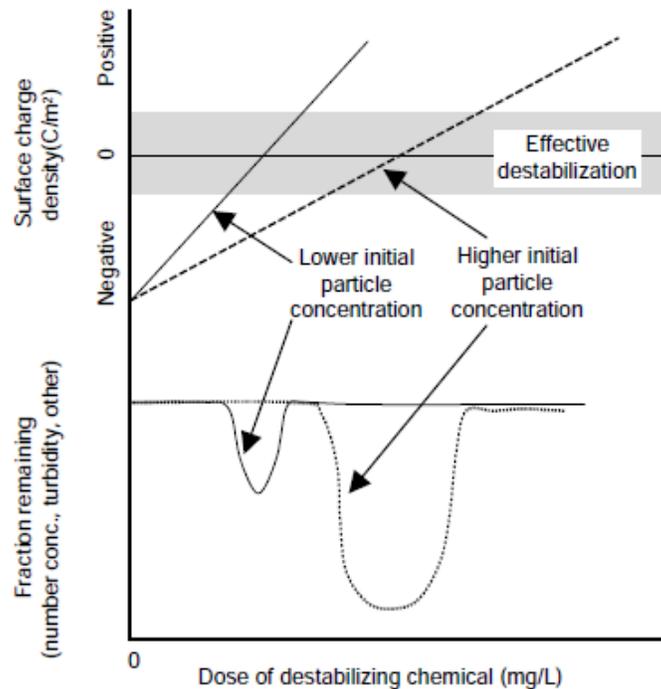


Figure 1.10: Particle destabilization by charge neutralization. *Source:* from [2, p. 536]

for the solid than the liquid). If such adsorption occurs and the added chemical has a charge opposite to that of the particle, the net charge on the particle surface is reduced (i.e. the surface charge is effectively being titrated by the addition of the coagulant). With respect to the top portion of fig. 1.10, the charge provided by the coagulant (positive, assuming the particle surface charge to be negative) progressively changes the net surface charge. At low doses of the chemical, the surface charge remains negative, but its magnitude is reduced; at some dose the original surface charge becomes neutralized, leading to a net surface charge of zero. The driving force for this adsorption is not primarily electrostatic, so further addition of coagulant beyond the dose needed for charge neutralization results in charge reversal, as illustrated in the bottom portion of fig. 1.10. This bottom portion shows the results of a procedure widely known as *jar test*. In such a test, several samples of a given dispersion are subjected to different doses of the destabilizing chemical, rapidly mixed for a short period of time (15-90 s), slowly mixed for a longer period (to allow flocs to grow, 15-45 min) and then allowed to settle for some time period (20-60 min) or filtered with a coarse filter. To evaluate

the effectiveness of particle removal, some measure of residual particle concentration (turbidity, suspended solids or total particle number concentration) is taken. Given the huge variety of coagulants available on the market, jar testing has always been the standard testing procedure for determining the proper chemical dose regardless of the mechanism of destabilization involved and irrespective of which particle removal process is being considered. A series of jars provides several points on one of the continuous curves shown and the most effective dose of the destabilizing chemical is that for which the chosen measure of residual particle content is lowest. It is worth noting that the transition between stability and instability is not abrupt (as shown in fig. 1.10), meaning that destabilization is effective not only when the net particle surface charge is exactly zero, but also in a range of doses around that optimal dose. One explanation for this phenomenon is that the energy barrier has been lowered enough to allow most of the particles in the dispersion to overcome it (remember that the thermal kinetic energy of particles follows a Maxwell-Boltzman distribution, hence there is always a certain amount of particles whose kinetic energy is greater than the average kinetic energy of the whole particle system). Another plausible reason has been offered by Gregory (1973) and Kasper (1971) with their so-called "patch model". According to their model, the destabilizing chemicals (usually relatively small and positively charged organic polymers) adsorb in patches onto the particles surface, forming an uneven charge distribution with some areas still carrying the original surface charge and others (where extensive adsorption has occurred) being of opposite charge. A negative patch on one particle is attracted to the positive patch of another, resulting in effective destabilization even at doses below or above that required for exact charge neutralization.

Another aspect that emerges from fig. 1.10 is that of *restabilization*, which occurs when too much destabilizing chemical is added that charged is reversed (fig. 1.11). Overdosing of the coagulant leads in this case to a dispersion that has the same problem as it did originally: charged - and therefore stable - particles. This phenomenon of overdosing and restabilization can sometimes lead to operational problems; if a dispersion is stable (as indicated by poor removal in the subsequent separation process), operators might mistakenly assume that the current chemical addition represents an under-dosing, prompting them to increase the dose. If, in reality, the current dose is an overdose, the correct decision would be to reduce the amount of coagulant instead. To summarize, when particle destabilization is accomplished by adsorption and charge

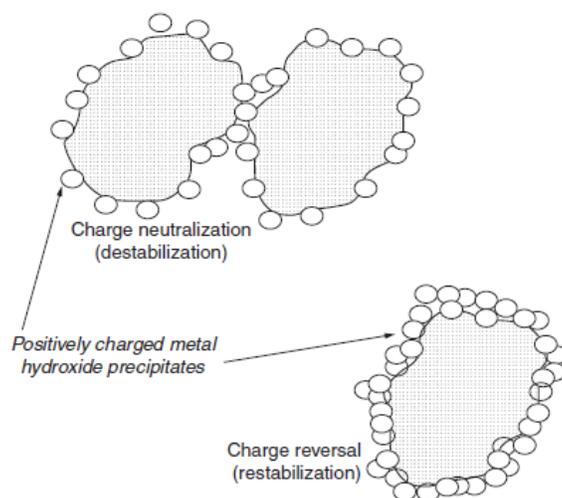


Figure 1.11: Schematic of possible metal hydroxide-particle interaction (adapted from Dentel & Grosset, 1987 and 1988)

neutralization, the chemical dose required is correlated with the particle concentration (i.e. there is a stoichiometric or quasi-stoichiometric relationship between the optimal coagulant dose and particle concentration). This stoichiometry is actually between the dose required and the cumulative surface charge, which is usually assumed to be proportional to the total surface area concentration of the particles in the dispersion. This stoichiometry is also illustrated conceptually in fig. 1.10 by the inclusion of curves (continuous and dashed) for two dispersions with different particle concentrations but otherwise identical characteristics. The same chemical dose in the two dispersions leads to a lesser change in the surface charge of the higher concentration suspension because the dose is spread among a larger number of particles. What may be not as obvious is that this spreading leads also to a wider range of chemical doses that are effective; that is, the same range of low (near zero) surface charge density leads to effective destabilization in both cases, but with a wider range of doses of coagulant for the higher particle concentration suspension (the lower particle concentration suspension presents a narrower range of effective coagulant doses, as represented in the bottom part of fig. 1.10). The stoichiometric effect can also lead to operational difficulties if the quality of water to be treated (i.e. the particle concentration and/or the solution conditions) is highly variable. In such a case, the coagulant dose must be varied to match the variation in water quality, which is not always an easy task especially in industrial wastewater

treatment, where discontinuous (and often unpredictable) discharges of concentrated and highly polluted streams into a main stream of less concentrated water continuously lead to peaks in the mixed flow (more or less damped/smoothed depending on the size of the installed equalization tank(s)). Use of on-line measurement of the streaming current, electrophoretic mobility, or zeta potential is particularly valuable to control the dose of destabilizing chemical in such situations. As a last remark, it should be noted in fig. 1.10 that the effective destabilization zone is not only wider but also deeper for the higher concentration suspension, because as it will be shown later in more detail, higher concentrations of particles are more easily contacted with one another, resulting in a greater extent of flocculation for a fixed time. Both inorganic chemicals (aluminium and iron) and charged organic polymers (like low-molecular weight cationic organic polymers such as polyDADMAC and epiDMA) can be used for adsorption and charge neutralization. Iron and aluminium are effective at pH values where their positive species dominate and they can also form precipitates with soluble NOM at relatively low pH. This specific chemical interaction makes these inorganic coagulants particularly effective for particles whose surface charge is caused primarily by adsorption of NOM; at high doses of the metal salts, precipitation in the form of metal hydroxides occurs.

1.2.3 Enmeshment in a Precipitate (Sweep Flocculation)

A third mechanism of destabilization called "enmeshment in a precipitate" occurs when an inorganic salt is added in such quantity that a precipitate is formed. Most commonly, the inorganic salt is $FeCl_3$ or alum, so that the precipitates that form are ferric hydroxide ($Fe(OH)_3(s)$) or aluminium hydroxide ($Al(OH)_3(s)$). In this destabilization mechanism, the cations that are the precursors of the precipitate first adsorb onto particles that are already in the water, just as occurs in the adsorption and charge neutralization mechanism. However, in this mechanism, the surface catalyzes the formation of nuclei of the new solid, so that precipitation can proceed at a significant rate at lower coagulant doses than are required in the bulk solution (*heterogeneous nucleation*). As a result, the original particles are caught up in the precipitate as it forms, as schematically shown in fig. 1.12. In most cases, the forming precipitates join with each other to form large flocs. One would expect that the hydroxide precipitates would be positively charged at normal pH values (because the PZC of both $Fe(OH)_3(s)$

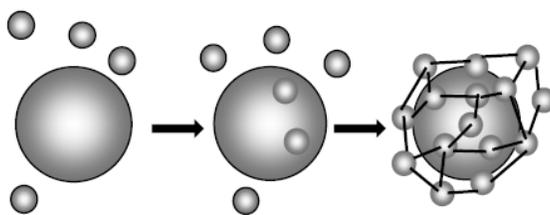


Figure 1.12: Adsorption of metal ions (the little spheres) onto a particle (the large sphere) leading to the formation of a precipitate. The lines in the right part of the figure indicate the lattice formation of the precipitate. *Source:* from [2, p. 538]

and $Al(OH)_3(s)$ is at alkaline pH). However, anions from solution (e.g. sulfate ions or NOM) adsorb onto the forming solids and neutralize the surface charge sufficiently to allow the flocs to continue to grow. This floc growth, combined with the fact that hydroxides incorporate a substantial amount of water into their formation, results in the rapid formation of large and fluffy flocs. If any of the original particles have not been caught up in the initiation of precipitation, they can be contacted quickly by these large flocs, leading to the notion of the particles being swept out of the suspension by the precipitate. As a result, this mechanism of destabilization is commonly referred to as *sweep flocculation* (or, simply, sweep floc), even though most particles are probably destabilized by acting as condensation nuclei that initiate the precipitation reaction. Because particles in the suspension act as nucleation sites, the higher the particle concentration, the higher the precipitation rate is expected to be. This behavior leads to a mild inverse stoichiometry; that is, the more particles are originally present, the less chemical is necessary to accomplish effective destabilization. Nevertheless, sufficient amounts of the metal salts must be added to achieve precipitation, and that often takes a factor of oversaturation of ~ 100 , so this inverse stoichiometry is often minor in comparison to the chemical requirement for precipitation. The truth is that the forming precipitate becomes the dominant species dictating the surface chemistry of all the particles in suspension, hence this mechanism can be used for virtually any dispersion as long as the pH is favorable for precipitation to occur.

To summarize, aluminium and iron can effectively destabilize particles by two different mechanisms: adsorption and charge neutralization at low metal concentrations and sweep flocculation at higher concentrations. Both the solution pH and the nature and concentration of particles in suspension influence these phenomena, so that concepts

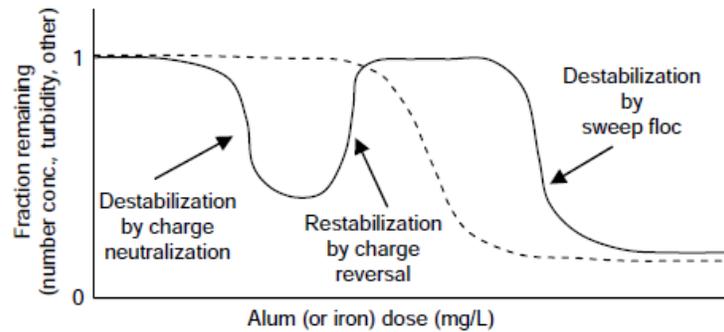


Figure 1.13: Destabilization by two different mechanisms by metal addition. Solid line represents a given suspension at a certain pH, where the two mechanisms are clearly distinct; dotted line represents a more concentrated suspension at the same pH (*Source*: from [2, p. 538])

like "high" or "low" coagulant concentrations are indeed not unique and well defined. The solid line in fig. 1.13 illustrates that the two mechanisms can be clearly differentiable in some cases, and this situation generally applies at relatively low particle concentration and relatively low pH values. If the pH is held constant, restabilization can occur with increasing dose when adsorption (without precipitation) is the dominant mechanism of interaction between the metal ion and the particle surface. Restabilization occurs in fact when the amount of adsorbed coagulant is sufficient to reverse the original particle charge. Destabilization is again accomplished at higher metal ion doses by enmeshment in the precipitate of the metal hydroxide. Conversely, if particle concentration is relatively large, the charge neutralization and sweep flocculation regions can blend into one another, leading to the behavior represented by the dotted line in fig. 1.13. In this case, the destabilization can be accomplished by charge neutralization at doses below those required for sweep flocculation, but *the dose where restabilization would begin is higher than that required for precipitation to be initiated*. Hence, *no restabilization occurs, and it is not possible to delineate where one mechanism of destabilization ends and the other begins* ([2, p. 539]). This behavior is most likely to occur with highly insoluble metal coagulants like iron(III), for which the pH range where charge reversal can be observed ($\text{pH} < 5-5.5$) is rarely encountered in the engineering practice (as can be seen in fig. 1.15).

Amirtharajah and coworkers summarized a large body of literature on the

effectiveness of aluminium and iron salts as effective destabilizing chemicals and presented the results graphically. The original diagrams by Amirtharajah and Mills (1982) for aluminium and by Johnson and Amirtharajah (1983) for iron were subsequently updated, as shown in fig. 1.14 and fig. 1.15, respectively. In each case, the

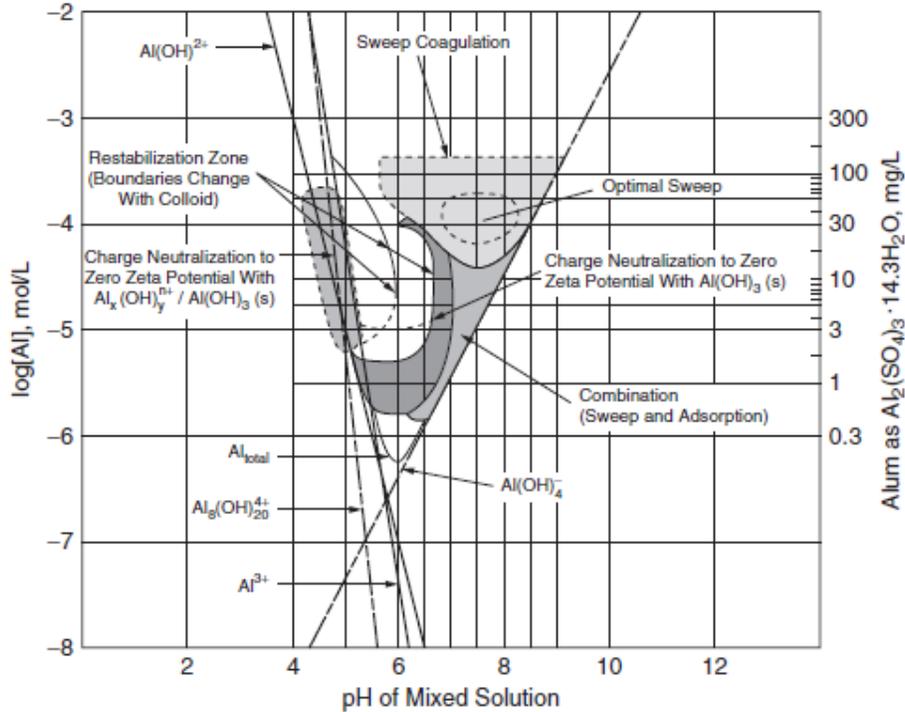


Figure 1.14: Alum destabilization diagram. *Source:* from Amirtharajah and Mills (1982), as redrawn in Edzwald et al. (1998)

authors reviewed many papers in which results of destabilization had been presented, and identified the effective zones of pH and coagulant dose for each mechanism on a log concentration versus pH solubility diagram for the hydroxides. Using for example the alum destabilization diagram reported in fig. 1.14, we can compare the expected behavior of two suspensions, one at pH 6 and the other at pH 7.5 when they are dosed with alum. We assume that a base is added simultaneously with alum, so that the pH is held constant. For the suspension at pH 6 we expect doses below $\sim 0.5 \text{ mg/L}$ to have essentially no effect, doses between 0.5 and 2 mg/L to be effective via the adsorption and charge neutralization mechanism, doses in the range of $2\text{--}30 \text{ mg/L}$ to be ineffective because the suspension undergoes charge reversal, and higher doses to cause

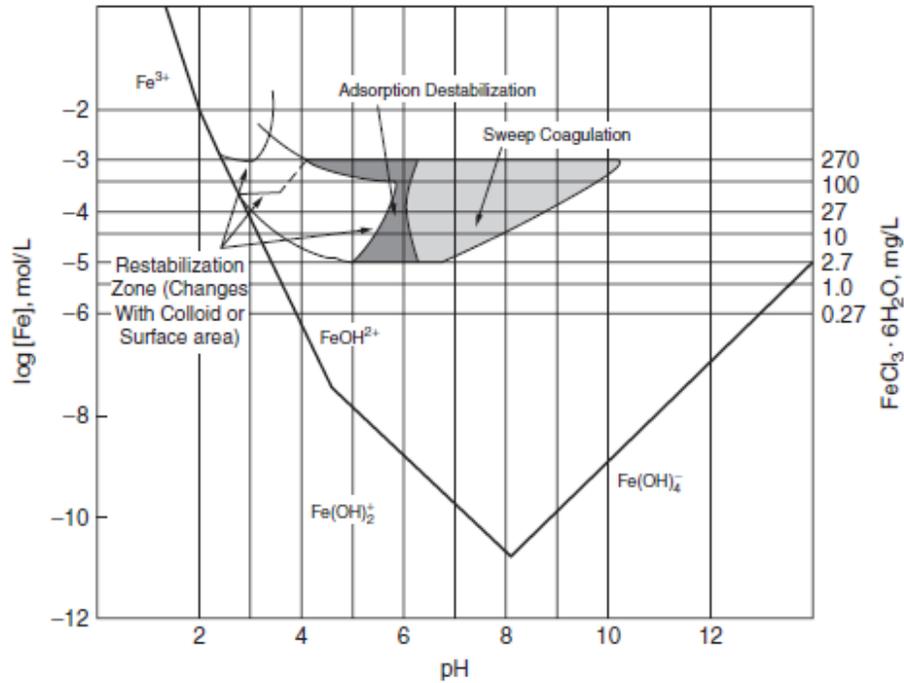


Figure 1.15: Iron destabilization diagram. *Source:* from Johnson and Amirtharajah (1983), as redrawn in Edzwald et al. (1998)

precipitation and destabilization by enmeshment in $Al(OH)_3(s)$ flocs. As indicated on the diagram, the boundaries of these different regions vary with different suspensions (as it will be discussed later in more detail), both because of particle concentration and because of the surface charge density. For the suspension at pH 7.5, destabilization is effective at all doses higher than $\sim 3 \text{ mg/L}$. In the range $3 - 12 \text{ mg/L}$ destabilization is accomplished by a combination of charge neutralization and sweep flocculation (it might be initiated by charge neutralization, but the surface concentration of aluminium builds to the point where precipitation occurs). At higher doses, precipitation occurs rapidly and sweep flocculation is the predominant destabilization mechanism. Because alum is acidic, the pH would tend to decline (more or less aggressively depending on the initial alkalinity) as the alum dose is increased, unless a base is simultaneously added to the system to keep the coagulation pH constant. A suspension starting from pH 7.5 might then end up at pH 6 when higher doses of alum (or ferric chloride) are applied, with a trend of destabilization with dose that is somehow intermediate between those

noted for the two constant pH suspensions.

Iron is less complex than aluminium in its behavior, primarily because it is less soluble. The distinction between charge neutralization and sweep floc is rarely clear at $pH \geq 6$, because doses that are sufficient to cause charge neutralization are a few orders of magnitude higher than that needed to initiate precipitation. Hence, as indicated in fig. 1.15, the mechanism is thought to be primarily charge neutralization at lower pH values and enmeshment in a precipitate at higher pH values. Comparing the two diagrams, it is immediately apparent that the region of insolubility of $Al(OH)_3(s)$ is much narrower than that of $Fe(OH)_3(s)$, which means that alum is effective as a coagulant in a much smaller pH range than iron. In wastewater or drinking water treatment, where the pH is usually expected to be higher than ~ 8.0 , iron is always the choice because of this difference in solubility which also affects the overall reliability of the process under uncertain conditions. Any other precipitate that is formed in solution as part of the overall treatment can also act as a coagulant. Lime for example may be added to raise the pH to cause metal precipitation (e.g. $Cu(OH)_2(s)$, $Zn(OH)_2(s)$, $Fe(OH)_2(s)$, $Mn(OH)_2(s)$) in many industrial wastewaters and either lime or iron can be used to cause phosphate precipitation. All of these precipitates can also help capture suspended particles in the same manner as $Al(OH)_3(s)$ or $Fe(OH)_3(s)$.

1.2.4 Adsorption and Interparticle bridging

The final mechanism of destabilization is adsorption and interparticle bridging. The destabilizing chemicals used in this mechanism are usually high-molecular weight (acrylamide-based) synthetic organic polymers and polyelectrolytes. According to this mechanism, polymer molecules are thought to adsorb to the surfaces of two or more particles by some specific chemical interaction (not simply electrostatics) and form bridges between them (fig. 1.16, a). Large polymers are needed to accomplish destabilization by this mechanism because the polymer bridge has to span at least two diffuse layers. Generally speaking, polymers are quite strong, and once attached to two particles, the floc is unlikely to break apart. As a result, this mechanism of destabilization is particularly useful in situations in which there is substantial hydraulic shear or breakup is particularly undesirable, such as in dewatering applications and granular media filtration.

What might be surprising is that polymers can cause destabilization through this bridging mechanism regardless of the charge on the particle surface; that is, even anionic and nonionic polymers can sometimes be equally effective in destabilizing negatively charged particles, thus confirming the hypothesis that the attraction to the surface is not simply electrostatic in nature. One possible explanation for this phenomenon is that if the polymer used has the same charge as the surface (usually negative), then it is common that an inorganic polyvalent ion of the opposite sign (e.g. Ca^{2+}) has to be added simultaneously to make the polymer effective. Thus, this calcium ion might be acting either as an indifferent electrolyte to compress the diffuse layer (and allow the polymer to overcome the electrostatic repulsion and reach the particle surface), or it might form complexes with surface sites on the particle and/or ligand sites on the anionic polymer, forming a particle-Ca-particle bridge that facilitates the attachment of the anionic polyelectrolyte to a negative surface.

Destabilization by bridging requires the same polymer molecule to attach to at least two different particles. As a result, *the effectiveness depends somewhat independently on both the particle concentration and the polymer dose, and not just on their ratio as in the charge neutralization mechanism* ([2, p. 541]). This experimental evidence arises from the fact that *at lower particle concentrations, the probability that an unattached end finds another particle is low, and so the molecule might bend around and attach to the same particle before finding another*; in such a case, no destabilization is achieved (fig. 1.16, b). In a high concentration suspension, it is far more likely that the unattached end will find another particle before bending around the same particle. This particle concentration effect is another reason that this mechanism of destabilization is often used in sludge dewatering applications, where the particle concentration is very high (for the same reason, it is rarely used in low concentration suspensions). Similarly to the adsorption and charge neutralization mechanism, *overdosing a suspension where destabilization by bridging is the goal can lead to restabilization, but for different reasons*. If a high concentration of polymer is added to the suspension, one end of several polymer molecules can attach to each particle, taking up most of their adsorption sites; the unattached ends are therefore prevented from attaching to another particle not only because most sites are already taken, but because the unattached ends are sterically hindered from making their way to the few untaken sites. Such a situation is schematically depicted in fig. 1.16 (c), where it is difficult to imagine the polymers

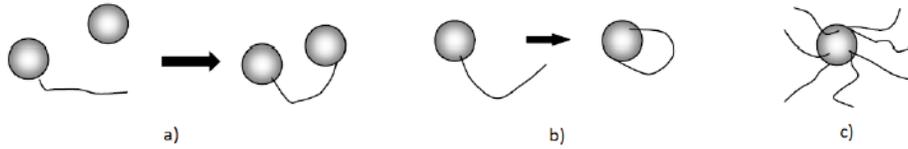


Figure 1.16: Interparticle bridging by a high-molecular weight polymer (a), ineffective destabilization at low particle concentration (b) and effects of polymer overdosing (c) (adapted from [2])

of another similarly overdosed particle successfully finding their way to the available adsorption sites on the surface of the one shown.

A relationship between polymer dose and particle concentration was first proposed by La Mer and coworkers (Smellie and La Mer, 1958; Healy and La Mer, 1962) and then refined by Hogg (1984) in such an interesting and elegant way that it deserves to be explored in more detail, as it will be later on in Chapter 6. To give just a hint of the hypothesis underlying Hogg's model, he assumed that the number of possible adsorption sites on a particle is proportional to the particle surface area, and that a polymer attached to a particle at one end is free to rotate so that the other end would tend to find an unoccupied site on another approaching particle. According to this bridging model, effective destabilization involves then two particles at once, and hence *it is not possible to completely separate destabilization from the processes that lead to floc formation* (i.e. coagulation and flocculation must be considered at once when applying this kind of model). The ingenious intuition behind Hogg's theory was to evaluate the probability of attachment of two particles by considering the inverse problem, that is: "*the probability that particles would **not** adhere to one another*". This (inverse) situation would occur only if the sites on the two particles were both either completely filled by polymer, or completely devoid of it. Although it may not seem immediate, this diametrically opposite point of view completely changes the modeling approach first proposed by La Mer, allowing to indirectly take into account the short-range forces responsible for free rotation and reorientation of particles before bridging. If f_{occ} is the fractional coverage of the particles and m_i and m_j are the number of possible adsorption sites on particles i and j , respectively, he found the attachment probability to be:

$$\alpha_{ij} = 1 - f_{occ}^{m_i+m_j} - (1 - f_{occ})^{m_i+m_j} \quad (1.27)$$

For a suspension with particles that are all identical (same size and same number of possible adsorption sites), α_{ij} is close to 1 over a broad range of values of f_{occ} , meaning that destabilization is effective over a wide range of doses, as confirmed by almost every study on particle destabilization by high-molecular weight polymers. For two particles of different sizes, this model suggests that the range of doses (i.e. the fractional surface coverage) that achieves effective destabilization broadens even further, as the size ratio of the two particles (smaller to larger) involved in the interaction decreases. Since the desire in engineering systems is to find the *minimum* effective dose (or surface coverage), overdosing to the point of restabilization is much less likely than in the case of charge neutralization. Unlike with charge reversal (where the surface charge can be measured), *no independent measure is available to determine directly whether an ineffective dose represents an under-dose or an overdose.* Hence, *empirical evidence directly related to the application (flocculation, filtration or dewatering) is the only tool available to determine the optimal (i.e., minimum effective) dose with this mechanism* [2, p. 542].

As a last remark, it is worth mentioning that use of high-molecular weight polymers does not necessarily mean that destabilization is accomplished by the bridging mechanism. Polymer charge often plays a significant role in destabilization, and the destabilization might well be accomplished by charge neutralization even for large polymers. What determines the mechanism is how the polymer is adsorbed to the surface; if it attaches at one end or forms large loops so that part of the molecule sticks out into the water and is available to attach to another particle, then the bridging mechanism is possible. However, it is also possible that the molecule will attach closely to the particle throughout the chain length. Walker and Grant (1996) have made the analogy that this latter phenomenon is like a zipper: after the first part of the molecule attaches to the surface, molecules along the chain keep attaching one by one, so that it closes onto the surface like a zipper. It is then important to know exactly how and when the polymer is added to the water stream in relation to the number concentration and size distribution of particles at the time of mixing, because ultimately which of the two mechanisms described above is most likely to occur will also depend on the length and structure of the polymer.

1.3 Destabilization of hydrophobic colloids by metal coagulants

A mechanism of destabilization of hydrophobic colloids with indifferent electrolytes has already been discussed in the previous sections (i.e. the so-called *compression of the double layer* mechanism). It has been seen that by increasing the ionic strength of the solution, the range of interparticle repulsion is progressively reduced, and the transition from stability to destabilization occurs over a narrow range of electrolyte concentration. Furthermore, if the indifferent electrolyte is increased to excess, there is no effect on destabilization, therefore *the electrolyte concentration required for destabilization is independent of particle concentration*. That is the case with indifferent electrolytes, but in reality no electrolyte is "indifferent" at all, and the re-stabilization of a suspension often occurs if an excess of coagulant is applied (excessive adsorption and charge reversal). It has been seen that when the exploited destabilization mechanism is that of *adsorption and charge neutralization*, a *stoichiometric relationship between particle concentration and optimal coagulant dose required for destabilization is observed*, which means that even for multivalent coagulant species, the magnitude of destabilization can be even greater than predicted by the $1/z^6$ Schultze-Hardy rule. As for the *bridging mechanism*, it is necessary to specify that lots of metal coagulants have a pronounced tendency to polymerize during hydrolysis reactions, and as the extent of hydrolysis increases, progressively higher polynuclear species form. In the same way as seen above for high-molecular weight polymers and polyelectrolytes, when such polymerized species adsorb, they tend to form bridges spanning between adjacent particles, thus promoting destabilization by bridging as well. The last destabilization mechanism that has been identified is that of *precipitate enmeshment* (or *sweep flocculation*). Under appropriate conditions of coagulant concentration and pH, metal coagulants in aqueous solution form metal-hydroxide precipitates. Such species serve to enmesh particulate matter, thus effecting destabilization essentially by a sweep action; apart from perhaps serving as nuclei for precipitation, *the presence of particulate matter is not strictly necessary, although the flocculation rate is considerably higher in more concentrated suspensions* [1, p. 84]. Precipitation mechanisms are of high importance during destabilization of hydrophilic colloids. Here, because of the extent of hydration, electrostatic effects are relatively unimportant [1, p. 84]. Coordination reactions occur between metal ions and

1. Theory and practice of Coagulation

Table 1.2: Characteristics of destabilization mechanisms with metal coagulants (*Source:* from [1, p. 83], adapted from Stumm & O'Melia, 1968)

Parameter	Influence of Indicated Parameter According to Mechanism			
	Physical Double Layer	Adsorption Destabilization	Bridging	Precipitation
Electrostatic interactions	Predominant	Important	Subordinate	Subordinate
Chemical interactions and adsorption	Absent	Important	Predominant	May occur but not essential for removal
Zeta potential for optimum destabilization	Near zero	Not necessarily zero	Usually not zero	Not necessarily zero
Addition of excess coagulant	No detrimental effect	Restabilization usually accompanied by charge reversal; may be blurred by precipitation	Restabilization due to complete surface coverage	No detrimental effect
Fraction of surface coverage (θ) for optimum floc formation	Negligible	$0 < \theta < 1$ (See Chapter 5)	$0 < \theta < 1$ (See Chapter 5)	Unimportant
Relationship between optimum coagulant dosage and particle concentration	Optimum dosage virtually independent of colloid concentration	Stoichiometry possible but does not always occur	Stoichiometry between dosage and particle concentration	Optimum dosage virtually independent of colloid concentration
Physical properties of flocs produced	Dense, high shear strength but poor filterability in cake filtration	Flocs of widely varying shear strength and density	Flocs of 3-dimensional structure; low shear strength, but excellent filterability in cake filtration	Flocs of widely varying shear strength and density

certain functional groups on the particle surface which serve as ligands; destabilization is then visualized as being the result of metal ion - functional group - hydroxide precipitate formation. Furthermore, with the removal of certain anions from solution (for example, phosphorus removal in tertiary wastewater treatment), precipitation reactions occur between metal ions, anions and hydroxides. To summarize the overall picture of the four mechanisms, in table 1.2 are reported various pertinent characteristics, from which is evident that even destabilization with metal coagulants cannot be exclusively attributed to any one particular mechanism; in a particular instance, destabilization may be contributed to by one or several of the proposed mechanisms.

1.3.1 Extent of hydrolysis and adsorption

During the transition from the free aquo metal ion to the insoluble metal hydroxide precipitate, metal coagulants undergo a series of hydrolytic reactions. It is strongly evident that the hydrolysis products of metal coagulants in aqueous solution are adsorbed more readily than free aquo metal ions; the greater the degree of hydrolysis, the more extensive is adsorption. In fact, replacing at least only one coordinate

water molecule of the hydrated metal ion by a hydroxyl ion enhances significantly the adsorptivity of the ion (Matijevic, 1973). The reasons for the enhanced adsorptivity of hydrolyzed metal ions on hydrophobic colloid surfaces are not well understood. Evidently, the presence of hydroxyl ions in the coordination sheath is to a large extent responsible, but just why hydroxyl ions are able to promote counter-ion adsorption remains to be resolved [1, p. 84]. One can visualize that with extensive hydrolysis, successively higher numbers of hydroxide groups per compound can be adsorbed at the interface. One may further postulate that the replacement of water molecules by hydroxyl ions in the coordination sphere of a metal ion imparts a certain degree of hydrophobicity to the complex (Stumm & O'Melia, 1968). Hydrophobic colloidal particles do not adsorb unhydrolyzed aquo metal ions for the same reason that they do not adsorb water molecules (Matijevic, 1973). Furthermore, when a dispersion is destabilized with hydrolyzed metal coagulants, if the suspension is made sufficiently acidic, restabilization occurs as evidenced from electrokinetic measurements by the reversion of particle surface charge back to its original value. Such desorption of metal ion species is due to de-hydrolyzation and a return to the aquo metal ion state, and it is slower than the adsorption of hydrolyzed species (Matijevic, 1973). This indicates a slow replacement of complexed hydroxyl ions by water molecules on acidification [1, p. 84].

1.3.2 Effect of coagulant dosage

As seen in table 1.2, the effect of coagulant dosage on destabilization depends very much on the destabilization mechanism operative. When the mechanism of EDL repression by increasing ionic strength is the predominant mechanism, then a destabilization response to coagulant concentration curve of the form shown in fig. 1.17 (a) will probably be obtained. The suspension passes from stability to destabilization over a narrow range of coagulant concentration. The *critical coagulant concentration* (CCC) is identified as the inflection point on the curve. Since destabilization relies solely on the concentration of counter-ions in solution, the CCC for this mechanism is independent of particle concentration. Furthermore, increasing the coagulant concentration beyond the CCC has little effect, because once the ionic strength is sufficient to compress the diffuse part of the DL enough for spontaneous destabilization to occur, no further advantage in

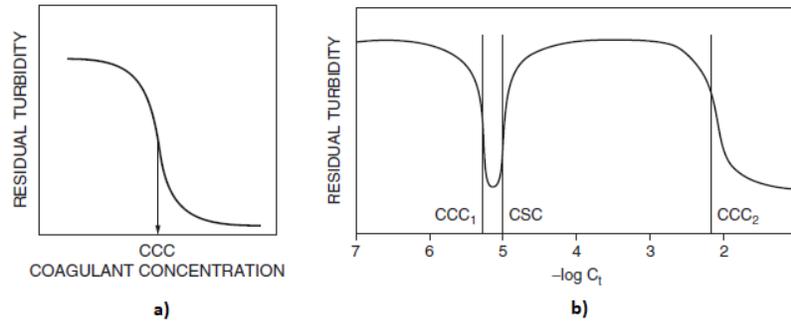


Figure 1.17: *Double layer compression* as the predominant mechanism, with little effect of applied coagulant concentration on the residual turbidity (a) and *adsorption and charge neutralization* as the predominant mechanism (b). (adapted from [1])

increasing ionic strength is attained. Finally, the valency of counter-ions has a profound effect on destabilization: the higher the valency, the lower is the CCC for destabilization. However, from experimental results, a typical situation (e.g. fig. 1.17, b) would be one where more than one destabilization mechanism is involved, depending on the coagulant concentration applied, on particle concentration and last but not least, on pH. The first region, identified as CCC_1 , occurs when sufficient metal hydroxide coagulant species are present to become adsorbed and thereby destabilize the suspension. In this case the destabilizing mechanism would be one of either adsorption of charged metal hydroxide species (giving rise to reduction of charge or charge neutralization), or adsorption of polymerized metal hydrolysis species giving rise to a bridging mechanism. The first mechanism is probably predominant at low colloid surface concentrations, whereas the second may become predominant at relatively high colloid concentrations. A mechanism of destabilization predominantly by diffuse layer compression is discounted because of the subsequent rise in the residual turbidity vs. coagulant concentration plot, thus if a higher coagulant dosage is initially applied, restabilization occurs. The *initial* concentration of coagulant required for restabilization to occur is identified as the CSC (*critical stabilization concentration*). *Restabilization could occur either because of an excess of counter-ion adsorption giving rise to charge reversal, or because of excess adsorption of polymeric species occupying too many adsorption sites, thereby preventing bridge formation between adjacent particles* [1, p. 86]. By further increasing the *initial* applied coagulant concentration sufficiently, a second destabilization region (whose CCC is identified as CCC_2 in fig. 1.17) would occur. Again, there are two

possible mechanisms, the predominance of each *depending on the colloid concentration*. At relatively high colloid concentrations, the predominant mechanism is probably due to ionic strength effects. That is, the presence of charged counter-ions in solution gives rise to double layer repression. Alternatively, perhaps predominantly at lower colloid concentrations, metal hydroxide precipitates are formed and destabilization in this case arises from a sweep action of colloid enmeshment by precipitates [1, p. 87].

From the above discussion, it is evident that besides coagulant concentration, there are at least two other factors influencing the destabilization characteristics and behavior of a given suspension of hydrophobic colloids. These are the concentration of colloids in the suspension and the pH of coagulation. It should be noted that each coagulant concentration in fig. 1.17 refers to an *initial applied coagulant concentration*, thus the two plots are built by interpolation of several experimental results on the same suspension (e.g. jar tests) and not by gradually increasing the coagulant concentration in one experiment.

1.3.3 Effect of colloid concentration

The preceding section focused the attention on the effect of coagulant dosage on the destabilization of hydrophobic colloids. This section extends the discussion to include the effects of colloid concentration. The relationship between coagulant dosage and colloid concentration (expressed in terms of surface area concentration) is cleverly depicted in fig. 1.18 and it is referred to one particular value of the coagulation pH (most likely between 4.0 and 5.0). The shaded area indicates zones where destabilization has occurred. It should be noted that there will be several such diagrams, each at a particular pH value (i.e. the *coagulation domain* of a suspension is not even a surface, but a 3D solid). Four zones are identified as follows:

- **Zone 1** indicates that insufficient coagulant has been applied to the colloidal suspension and that destabilization does not take place.
- **Zone 2** refers to the region where destabilization has taken place.
- **Zone 3** is that region where destabilization (and then restabilization) has taken place, due to excessive coagulant addition.

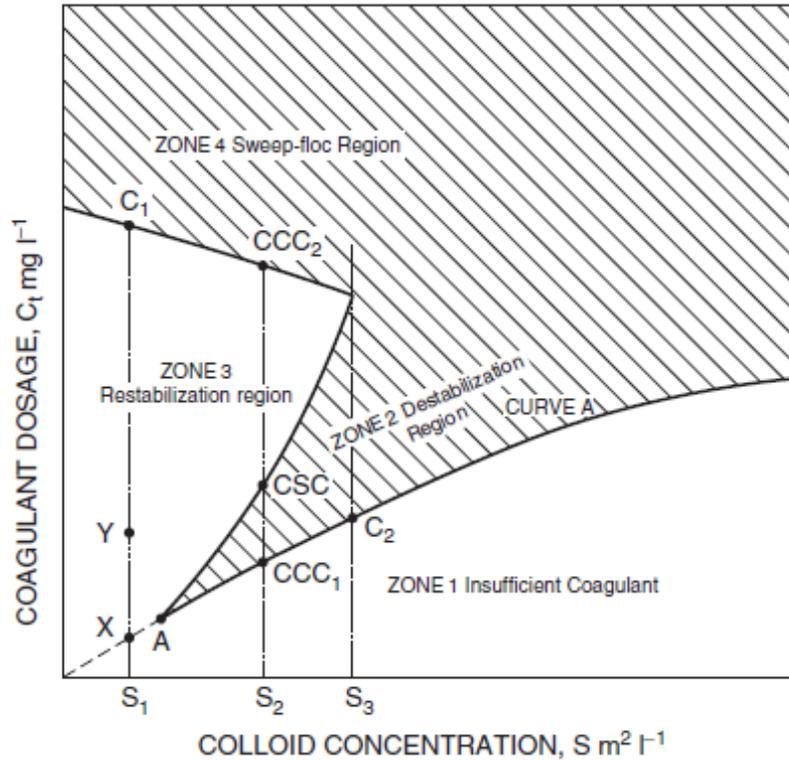


Figure 1.18: Destabilization and restabilization regions at a given pH value as related to colloid concentration (S) and coagulant dosage (C_t) (*Source: from [1, p. 87], adapted from Stumm & O'Melia, 1968*)

- **Zone 4** is the region where the coagulant dosage is high enough for oversaturation, and precipitation of metal hydroxide species occurs.

At very low concentrations of colloids (e.g. S_1 in fig. 1.18) there is very little contact opportunity for colloid-colloid interactions. When a quantity of metal coagulant such as X is introduced, even though the extrapolated stoichiometric coagulant dosage may be applied, uneven adsorption occurs, and hence a certain number of particles will remain stable, some will suffer charge reversal and others will be destabilized. The problem here is therefore more related to an inadequate dispersion, rather than to a chemically unfavored situation (*diffusion limited regime*). Suppose now an increased coagulant dosage Y had been applied. Although colloids are still greatly dispersed, there is an ample supply of coagulant species for each colloid. However, the concentration of coagulant with respect to the colloid concentration is relatively high, so that

excessive adsorption takes place and restabilization occurs as described in the previous sections. *The only way in which it is possible to remove the colloidal solids when particle concentration is too low is therefore to increase the coagulant concentration to a value of at least that given by C_1 , which corresponds to a coagulant dosage beyond its solubility limit that is sufficient to remove particles from water by precipitates enmeshment (sweep floc mechanism).* For low concentrated suspensions this coagulant dosage decreases as colloid concentration increases (as shown by the negative slope of the line passing through C_1 and CCC_2 in fig. 1.18); this behavior derives from the fact that the more particles there are, the more are the contact opportunities in a fixed time between coagulant species and colloids, which also act as nuclei for metal hydroxides precipitation (i.e. the coagulation rate by sweep floc increases). Similarly, as the colloid concentration progressively increases, adsorption of transition hydrolysis products occurs to a certain extent thereby forming nucleic sites for precipitation, and thus less coagulant is progressively required to form precipitates [1, p. 88].

At a certain colloid concentration, such as that corresponding to A in fig. 1.18, contact opportunities are sufficient for hydrolysis species to become adsorbed and promote destabilization by charge neutralization or bridging, before precipitation takes place. It is interesting to note that waters of low solids concentration may be more efficiently destabilized by augmenting the total surface area to a value such as that given by S_2 for example. This step is usually carried out in some cases of water treatment practice where bentonite or activated silica may be added. At colloid concentrations such as S_2 , it is seen that the critical coagulant concentration for destabilization is much less than for S_1 . However, destabilization occurs over a relatively narrow range of coagulant dosage. Beyond a critical dosage (CSC) restabilization of the suspension by charge reversal occurs and significant colloid removal for coagulant dosages greater than CSC would only occur if a quantity of coagulant sufficient for precipitate formation and sweep action is added. As discussed in the preceding section, *the mechanism for destabilization for coagulant dosages between the CCC and CSC is one dependent on adsorption. Because of the predominance of an adsorption type mechanism in this region of colloid concentration, a dependence of the CCC on colloid concentration is somehow expected.* This is seen to be so if considering the linear behavior of curve A in fig. 1.18 in this region. Such stoichiometry probably occurs when suspended solids concentrations are greater than that represented by point A.

At high colloid concentrations (S_3 and higher) it is seen that restabilization does not occur. This may be explained by considering chronologically the action of metal hydrolysis species in solution. The first effect must be compression of the diffuse part of the double layer, rather than extensive adsorption. As stated earlier, adsorption is in fact more favourable with higher hydrolyzed metal species. Hence, before extensive adsorption can occur with its eventual charge reversal, colloids, being in closer proximity than the preceding cases considered, will be destabilized and will have already formed flocs. In other words, *the hydrolysis and adsorption kinetics are now comparable to the collision rate of particles, which results in a destabilization by both adsorption and compression of the DL that is more biased towards the latter than the former the more concentrated is the suspension.* Hence, the region corresponding to colloid concentrations greater than S_3 does not necessarily imply being a sweep floc region (provided a solution pH far from initiating extensive precipitation), but rather a region of destabilization brought about by a double layer repression mechanism. This is supported by the shape of curve A, which is horizontally asymptotic with increasing colloid concentration (remember that the CCC associated with destabilization by compression of the diffuse layer is almost independent of the colloid concentration). It should be noted, however, that if the coagulant dosage is very high, then metal-hydroxide precipitates will obviously form, but there would appear to be little advantage from the point of view of destabilization efficiency. Zone 1 is also known as a region of *reaction limited regime* (as opposed to Zone 3, which was identified as a region of *diffusion limited regime*) because even though the chemical conditions for attachment are now unfavorable, particle concentration (and hence the collision frequency) is now high enough to increase the flocculation rate to such an extent that it becomes practically independent of particle concentration (it will depend solely on the other limiting factor, i.e. the coagulant dosage).

To summarize the above discussion:

- For low colloid concentrations, less than that corresponding to point A, the destabilization mechanism is one governed by precipitate formation.
- For medium concentrations, namely between A and S_3 , the destabilization mechanism is governed by adsorption, and charge reversal may occur. Here, both the Stern and Triple Layer Model for the EDL within the DLVO framework

for particle stability would adequately describe the phenomena (charge reversal is accounted thanks to the additional terms σ_0 and ψ_β in both models).

- Depending on the solution pH, for high colloid concentrations (greater than S_3) a double layer repression mechanism such as that described by the Guoy-Chapman model is probably dominant, and a refined mathematical description of the Compact (or Stern) Layer would probably be unnecessary; still, V_{TOT} should be computed according to DLVO theory as well.
- For medium and high colloid concentrations, destabilization may be brought about by high coagulant dosages in the form of precipitate enmeshment (sweep flocculation), where metal-hydroxide precipitates are induced to form onto the particle surface and grow both by precipitation and floc aggregation.

1.3.4 Effect of pH

The preceding discussions were only focused around the combined influence of colloid concentration and coagulant dosage on the characteristics of destabilization of hydrophobic colloids. The third factor of paramount importance in destabilization with hydrolyzing metal coagulants is the coagulation pH, because ultimately the predominance of a particular hydrolysis species during destabilization is very largely dependent on the pH value. For a particular colloidal suspension, it is logical to consider that there exists a particular hydrolysis species most effective for destabilization, whether by virtue of the charge carried, or the adsorptivity of the species itself [1, p. 90]. At each pH value, curves similar to those represented in fig. 1.19 can be derived. From fig. 1.19 (which refers to a particular case of a silica dispersion using $Fe(III)$ as coagulant) it is seen that for strongly acidic pH values ($pH \leq 2$), at low colloid surface concentrations appreciable destabilization does not occur. At pH=1 restabilization does not occur and the coagulant concentration necessary for destabilization is independent of colloid concentration, probably because of the hydrolysis species existing at this pH value. In strongly acidic environments, hydrated aquo metal ions are in fact the predominant species, therefore the destabilization mechanism is most likely just double layer repression. The absence of destabilization at low colloid concentrations can again be explained in terms of kinetics, meaning that contact opportunities are too low for appreciable perikinetic flocculation to be observed in the time allowed

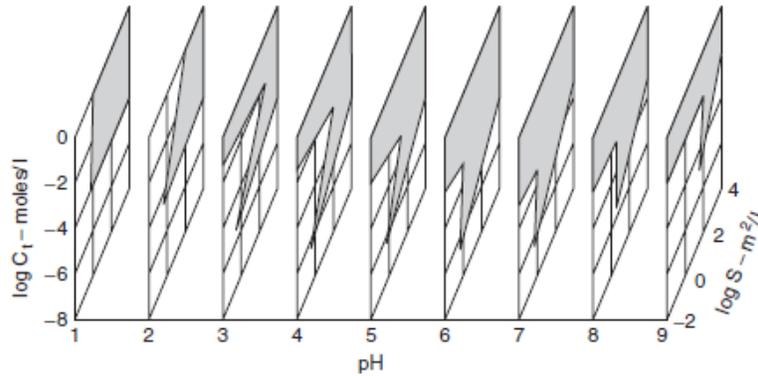


Figure 1.19: Zones of destabilization and restabilization with respect to coagulation pH, colloid concentration (S) and coagulant dosage (C_t) (*Source*: from [1, p. 90], adapted from Stumm & O'Melia, 1968)

for destabilization. At pH=2 restabilization occurs only at low to medium colloid concentrations, because at this pH value hydrolysis products begin to have a pronounced tendency for adsorption, such that overdosing can lead to surface charge reversal and restabilization. At high colloid concentrations, destabilization is due to double layer compression by positively charged transition coagulant species. For pH values from 3 to 5 it is seen that at a given pH the coagulant dosage vs. colloid concentration plot is as discussed in the previous section. From fig. 1.19 it is seen that the slope of the CCC vs colloid surface concentration boundary slightly increases with increasing pH, meaning that there is an increasing dependence of the destabilization mechanism on adsorption due to extensive hydrolysis. To give an example, at a pH of 3 the percentage of hydrolysis of coagulant species at the CCC is approximately 7%. Increasing the coagulant dosage to a value approximately equal to the CSC, the percentage of hydrolysis rise up to approximately 60% [1, p. 91]. Thus, increasing dosage in each case gives rise to increased hydrolysis and consequently increased adsorption: at low pH values (e.g. pH=3) the predominant species are positively charged, hence restabilization is due primarily to charge reversal; at higher pH values (around 5), the predominant species are negatively charged (for example $Fe(OH)_4^-$) and one may postulate a mechanism of bridging with excessive surface site coverage resulting in restabilization. For pH values of 6 and higher, it is seen that restabilization does not occur. This evidence might be explained by a destabilization mechanism of bridging by negative hydrolysis species with a continuous transition to precipitate enmeshment as the

coagulant dosage is incrementally increased, as supported by electrophoretic mobility measurements. The progressive increase in CCC with increasing pH supports the predominance of negatively charged hydrolysis species, since with increasing negativity there are statistically less contact opportunities between coagulant species and colloidal particles. To briefly summarize, it is generally accepted that at low pH values the electrophoretic mobility values are seen to increase from negative to positive values postulating an excess of positively charged hydrolysis species, while at higher pH values the mobilities are seen to increase in negativity for the exact opposite reason.

1.3.5 Effect of Temperature

Coagulation reactions and metal coagulant chemistry are strongly affected by temperature. *The effect of temperature appears to be more pronounced when destabilization reactions mostly rely on the enmeshment mechanism* (Morris & Knocke, 1984), probably because *with decreasing water temperature, the solubility of hydroxide species shifts to higher pH values* (because solubility of solids increases with water temperature), *therefore the "optimum" operating pH also shifts to higher pH*. Surprisingly, *for the adsorption mechanism temperature effects appear to be less severe, especially for high turbidity waters*. In general, *coagulation of low turbidity waters seems to be particularly affected by temperature regardless of the destabilization mechanism. For the sweep floc mechanism, temperature appears to have more influence on the floc characteristics (smaller flocs at low temperatures) rather than on the rate of metal hydroxide precipitation* [1, p. 225]. For the adsorption mechanism, a decrease in temperature negatively affects both the sorption kinetics and the perikinetic flocculation rate, but the equilibrium is shifted towards adsorption. It is worth noting that temperature has also affects the thickness of the double layer, because an increase in temperature also results in a decrease in the Debye-Huckel length. In fig. 1.20 is shown the detrimental effect of overdosing alum at low temperatures contrarily to ferric chloride, which is much less affected by the shift in optimum pH for sweep coagulation at low temperatures (Hanson and Cleasby, 1990). The increase in optimal coagulation pH with decreasing temperature has also been observed with coagulation of humic substances.

One characteristic of hydrolyzing coagulants such as alum and ferric sulfate is that

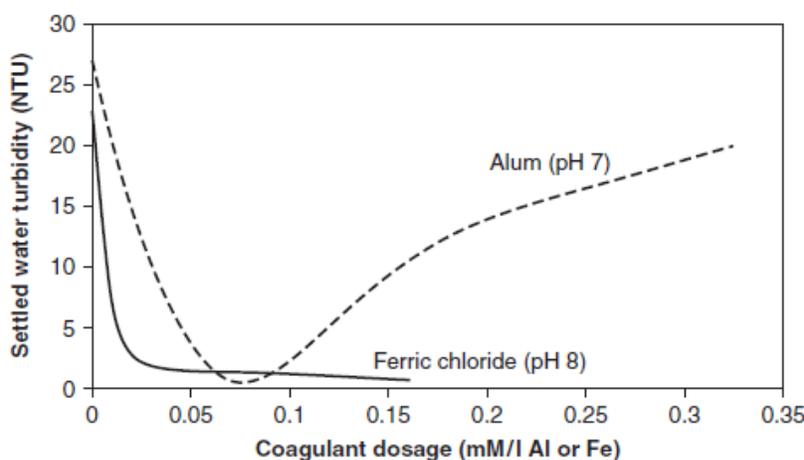


Figure 1.20: Performance of ferric chloride and aluminium sulfate at 1°C (*Source:* from [1, p. 225], adapted from Morris & Knocke, 1984)

there is essentially no control over the coagulant species formed under the prevailing conditions of concentration and pH during treatment. A difference with pre-polymerized coagulants such as PACl and polyferric sulfate is that they can be tailored for particular raw water conditions of temperature and other parameters, and produced with optimal hydrolysis and polymeric species. Such compounds may retard the rate of hydroxide precipitation and allow coagulation reactions to proceed with the best species for a longer time. Therefore, a significant advantage of pre-polymerized coagulants is that the region of optimum coagulation can be expanded over a wider range of raw water conditions, including temperature [1, p. 227].

1.4 Destabilization of hydrophilic colloids by metal coagulants

Because of hydrated layers surrounding them, hydrophilic colloids are not destabilized by such considerations as depression of the EDL. Nevertheless, hydrophilic colloids do carry a surface charge, although the effective range of influence may be confined in some cases to within the hydrated layer alone. This surface charge is often due to the ionization of ionogenic groups (e.g. carboxylic, aliphatic or aromatic hydroxyl, sulphato, phosphato and aminogroups). Stumm and Morgan (1962), by means of

alkalimetric titration curves, demonstrated that metal ions tend to form complexes readily with ligands such as phosphates, pyrophosphates, salicylates and oxalates, which represent the most abundant functional groups of many hydrophilic colloids. As ligands, such substances do not have as great an affinity with metal ions as do OH^- ions, and statistically speaking the selective coordination reaction will depend on the ratio between the concentration of ligand in solution and the concentration of OH^- ions. Ligands will therefore tend to be preferentially coordinated at lower pH values. Furthermore, to form an electroneutral precipitate, less OH^- ions will be required in the coordination sphere of the metal-ion complex, so that the pH for precipitation will be lower in the presence of such ligands than if OH^- were the only ligand present [1, p. 97]. It is generally assumed (Stumm and Morgan, 1962) that *destabilization of hydrophilic colloids carrying functional groups such as those described above is principally due to precipitation of the metal-ligand complexes, requiring both acidic pH values and higher coagulant dosages than their hydrophobic counterparts.*

1.5 Interactions of destabilizing chemicals with soluble materials

It is now commonly accepted that the major source of surface charge of particles in most natural waters is adsorbed NOM (O'Melia, 1989; Buffle et al, 1998). When a metal coagulant is added to water containing both soluble NOM and particles with adsorbed NOM, the interactions between the hydrolysis products of the metal coagulant and NOM dictate the success or failure of destabilization. Natural organic matter is mostly removed either by adsorption and charge neutralization or enmeshment in a precipitate (or as metal-NOM precipitates). *Generally, the ranges of pH and coagulant dose that achieve effective NOM removal overlap with the regions of effective particle removal* (because, the particle surface charge is usually caused by the NOM itself, so the coagulant interacts with the NOM in solution and on the particle in much the same way)[2, p. 542]. However, in the region where both the pH and the coagulant dose are low (and particle removal occurs by charge neutralization), NOM removal occurs by aluminium-NOM precipitation (if using alum for example). At the higher-pH, higher-alum-dose region where sweep flocculation is the main mechanism for particle

removal, NOM removal occurs primarily by adsorption onto the high surface area of the precipitated amorphous $Al(OH)_3(s)$. Current US drinking water regulations require most utilities using surface water supplies to remove some NOM, and an acceptable method is by interaction with aluminium or iron. To accomplish the desired NOM removal, the coagulant is often overdosed beyond that required for effective particle destabilization; this process of achieving both organics and particle removal by the addition of higher doses of destabilizing chemical is known as *enhanced coagulation*, and can be accomplished by any of the destabilizing chemicals leading to precipitation, with iron often being more effective than alum for the reason illustrated in the previous sections. NOM interacts with iron slightly more favorably than aluminium, but can also be removed by other precipitates that are formed during treatment. Besides NOM, metal precipitates are also effective adsorbents for the removal of other species like arsenic, silica and heavy metals (which are adsorbed or coprecipitated), therefore though the primary reason to add these coagulants is typically to destabilize particles, the conditions (mainly the pH), concentrations and even the choice of coagulant can vary according to the desired removal of these other soluble contaminants [2, p. 543].

1.6 Combination of Additives

It is often helpful (or even necessary) to add more than one chemical species to accomplish particle destabilization. If, for example, influent pH varies considerably over time, coagulation with metal coagulants needs to be carefully controlled as pH significantly affects the hydrolysis products and hence the effectiveness of destabilization. *For NOM removal, a relatively low pH is often desirable, and so an acid is sometimes added if the influent pH is not low enough after addition of alum or iron alone* [2, p. 544]. On the other hand, the acidic behavior of these coagulants sometimes must be offset by adding a base. Since most of the surface charge of virtually all particles encountered in environmental engineering applications is a function of pH, the pH can be adjusted as well to accomplish some charge neutralization (if it does not bother too much the effectiveness and controllability of the coagulant). For example, it might be cheaper to add an inorganic acid or base to somehow reduce the surface charge before adding a charge neutralization polymer to complete particle destabilization. Only rarely could the inorganic acid or base be used for the entire neutralization, as the

pH_{PZC} is usually outside the range that would be considered acceptable for the next treatment process, discharge or ultimate use. When the bridging mechanism is exploited and the polymer has the same charge as that of the particles, the same phenomena that tend to prevent interparticle collisions will act to prevent polymer-particle collisions. *To overcome this problem, it is usually suggested to introduce multivalent ions of the opposite charge to further reduce double layer repulsion to accomplish destabilization*[2, p. 544]. For example, if the particle surface charge is negative and an anionic polymer is used, calcium ions could be added prior to the polymer (as, say, $CaCl_2$, or $Ca(OH)_2$ if a pH correction is suggested for other purposes). Finally, bridging polymers are often added not so much for particle destabilization, but rather to improve floc strength and hence to enhance floc growth and/or the effectiveness of subsequent dewatering processes. In such a situation, other chemicals (either inorganic, organic, or both) might be added first to accomplish the destabilization, so that the bridging polymers are added later just to minimize either floc breakup (in flocculation) or particle detachment in deep bed filtration processes.

Chapter 2

Flocculation Modeling

2.1 Collision mechanisms and the Smoluchowski equation

Flocculation basically consists in the collision of particles and their attachment to one another to form a floc. To make collisions possible, particles are brought together thanks to *long-range forces* acting over distances of approximately two or more orders of magnitude times the size of the largest particles in the system ([2, p. 565]). These forces causing long-range motion include *external forces on the fluid* (which generate fluid motion), *gravity* (which causes vertical motion of the particles relative to the fluid) and *thermal activity of molecules and ions in solution*, leading to Brownian motion of the suspended particles. A simplified sketch of the above mentioned mechanisms is depicted in fig. 2.1. When particles get quite close to one another, other forces become significant: these short-range forces include, at the least, the van der Waals attraction and electric double layer repulsion (or, perhaps, attraction) explained in the previous chapter, as well as viscous forces, which are progressively more relevant the shorter the separation distance as the water separating two approaching particles moves out of the way to let particles collide. The interparticle distances over which this short-range forces act are different. The disturbance associated with particle motion across the fluid and generated by viscous stresses due to surface friction generally extends for a distance equal to a few tenths of the particle diameter ([2, p. 566]). When two particles

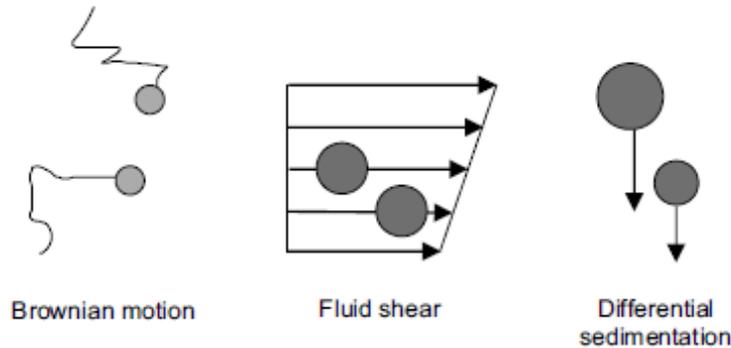


Figure 2.1: Particle paths leading to collisions according to the three most important mechanisms of long-range motion (the length of the arrows reflects the particle velocities). (Source: from [2, p. 565])

approach one another, the water being pushed out of the way by each particle also tends to carry the other particle out of the way, and as the particles get closer, the cross-sectional area through which the water must flow away is restricted (and so the resistive force increases). Thus, *short-range viscous forces always oppose collisions, and they intensify as the particles get closer to one another*. On the other hand, electrical double layer interactions usually extend over a distance that is usually shorter than that of short-range viscous forces. Because of the different distances over which these short-range forces act, even if EDL repulsion is completely eliminated by chemical destabilization of the particles, two particles that are not initially moving in directions that would lead to a collision cannot be made to collide via the short-range interactions only (they first have to feel each other's presence). Rather, for flocculation to occur, two widely separated particles moving independently from one another must first be moving in directions that would lead to a collision (long-range motion), and even then only a fraction of these impending collisions actually occur, because of short-range forces that cause the particles' paths to deviate before collision.

2.1.1 Rate Equation for floc formation

Flocculation can be viewed as a two-step, sequential process, similar to a two-step elementary chemical reaction. In the first step, particles are transported from locations where they have no influence on one another to locations where they do affect each other's motion; in the second one, they have to overcome the short-range barriers to

2. Flocculation Modeling

collision to form a floc. Assuming floc formation to be an irreversible process, the flocculation process might be represented as follows ([2, p. 567]):



where P_i and P_j are two generic particles in the system, $P_i \equiv P_j$ is a pair of particles that have approached each other closely enough that they affect each other's subsequent motion, whereas P_k is a floc formed by the collision of P_i and P_j ; β_{ij} , $\gamma_{f,ij}$ and $\gamma_{b,ij}$ are rate constants for the respective reactions. β_{ij} is commonly known as *collision frequency function* and it characterizes the transport of P_i and P_j particles to a separation distance where the short-range forces become significant; $\gamma_{f,ij}$ and $\gamma_{b,ij}$ are, respectively, the forward and reverse rate constants for the two possible transitions from that state, that is either to form a floc or to separate again. The forces favoring floc formation (i.e. the same ones that brought the particles close to one another plus the van der Waals attraction) are incorporated into $\gamma_{f,ij}$, whereas $\gamma_{b,ij}$ contains those forces opposing floc formation. Assuming that the concentration of particles in close proximity to one another (i.e. $P_i \equiv P_j$ pairs) is approximately at steady state, a balance on the number concentration of those particles in the system (n_{ij}) can be written as follows:

$$\frac{dn_{ij}}{dt} = 0 = \beta_{ij}n_i n_j - (\gamma_{b,ij} + \gamma_{f,ij})n_{ij} \quad (2.2)$$

Looking at each term separately:

$$n_{ij} = \frac{\beta_{ij}n_i n_j}{\gamma_{b,ij} + \gamma_{f,ij}} \quad (2.3)$$

$$r_{k,i+j} = \gamma_{f,ij}n_{ij} = \frac{\gamma_{f,ij}}{\gamma_{f,ij} + \gamma_{b,ij}}\beta_{ij}n_i n_j = \alpha_{ij}\beta_{ij}n_i n_j \quad (2.4)$$

where $r_{k,i+j}$ is the rate of floc formation (P_k particles) via collisions between P_i and P_j particles, n_{ij} is the number concentration of $P_i \equiv P_j$ "proto-flocs" and $\alpha_{ij} = \gamma_{f,ij}/(\gamma_{f,ij} + \gamma_{b,ij})$. It should be noted that $r_{k,i+j}$ is the rate of formation of type-k particles from specific type-i and type-j particles; however, the overall (net) rate of formation of type-k particles should also consider their disappearance as they collide with other particles to

form larger flocs. Conceptually, α_{ij} can be interpreted as follows:

$$\alpha_{ij} = \frac{\gamma_{f,ij}}{\gamma_{f,ij} + \gamma_{b,ij}} = \frac{\gamma_{f,ij}n_{ij}}{\beta_{ij}n_i n_j} = \frac{\text{Rate of formation of } P_k \text{ flocs}}{\text{Rate of close approaches of } P_i \text{ and } P_j \text{ particles}}$$

In other words, α_{ij} can be defined as the percentage of close approaches of P_i and P_j particles that effectively results in floc formation. The analogy between flocculation and chemical reactions is entirely consistent with the physical description of flocculation given above. According to this view, a close approach of two particles is the same thing as the formation of the intermediate species $P_i \equiv P_j$ envisioned in the chemical description. At the same time, if we assume a close approach to be the critical prerequisite for a collision, α_{ij} can be interpreted in the physical description of the process as the fraction of the potential collisions (i.e. close approaches) that actually forms a floc. Based on this understanding, α_{ij} is therefore commonly called the *collision efficiency function*. Although flocculation modeling has progressed over time with an increasing inclusion of phenomena and sophistications impacting floc formation, the predicted rate of floc formation often tends to be overestimated in the models because they consider an idealized view. To account for this difference and match the observed rate of floc formation with model calculations, an empirical correction factor is often introduced; that is:

$$\alpha_{emp} = \frac{\text{Actual } r_{k,form}}{\text{Predicted } r_{k,form}} < 1 \quad (2.5)$$

A common expression for this *collision efficiency factor* (α_{emp}) does not exist because it will ultimately depend on what phenomena are already included and mathematically made explicit in the model (i.e. the mathematical expressions of α_{ij} and β_{ij}). In early efforts to model flocculation, values of β_{ij} for pairs of particles of various sizes were computed considering the three long-range transport mechanisms shown in ??; later modeling efforts have incorporated some of the short-range forces explicitly. As it becomes possible to account for more forces explicitly (and hence to account more accurately for those that are known), fewer forces will need to be lumped into α_{emp} , leading to presumably a better decision making in the design and operation of flocculation systems as a better understanding of the factors controlling flocculation can be gained. In a sense, the goal of flocculation modeling is to make $\alpha_{emp} \rightarrow 1$, so that all the relevant factors affecting successful flocculation would be explicitly taken

into account.

Although the chemical analogy used to model the formation of flocs from smaller particles (eq. (2.4)) is conceptually very useful, its application to flocculating systems is inherently more complex for several reasons. First, flocs of a given size can be formed by a very large number of combinations of smaller particles, even if only pairwise combinations are considered. For instance, the rate of formation of P_k -sized flocs must consider all the possible collisions between two particles individually smaller than P_k , whose sizes add up to the size of the P_k particles. Thus, the rate of formation of P_k -type particles must consider the contributions from many individual reactions occurring simultaneously. Second, even when considering reactions between two particles of given sizes individually, several different reaction mechanisms or pathways can bring them close to one another. Three of such pathways have been introduced earlier (namely fluid shear, sedimentation and Brownian motion) and must evidently be evaluated independently to have a better understanding of the effect they have on flocculation rate (but in reality such independence does not exist and these mechanisms can obviously influence one another). Third, unlike chemical reactions, virtually all products of a flocculation reaction (i.e. flocs) can themselves participate in subsequent reactions to form even larger flocs. Finally, the factors affecting the rate constants for flocculation reactions depend not only on the chemical identities of the reactants and the system temperature, but also on the physical characteristics of the particles and the degree of mixing in the flocculation reactor. A complete analysis of flocculation kinetics would require an evaluation of both α_{ij} and β_{ij} functions for each possible pair of colliding particles in the system, considering all possible ways that they could be brought together, and introduce all these values in the rate expressions describing the formation of flocs of all possible sizes. Ideally, the computation of α_{ij} and β_{ij} values would be based on fundamental principles of hydrodynamics and surface chemistry, and would incorporate all possible phenomena that promote or impede collisions. The history (and ongoing research) of flocculation modeling can be viewed as a progression of attempts to understand and quantify in more and more detail the individual factors controlling these two key functions.

2.1.2 The Smoluchowski equation

In 1917, Smoluchowski published what still remains the fundamental theory of flocculation. He wrote a *population balance equation* (PBE) for flocculating particles in a suspension (i.e. a system of equations describing the temporal evolution of the discretized particle size distribution), which incorporated the following concepts:

- Particles of each size class can be formed by flocculation of smaller particles and lost by the formation of larger particles;
- The probability of three or more particles colliding at exactly the same instant is considered negligible in comparison to that of pairwise collisions;
- Particle volume is conserved during floc formation.

With these assumptions, the Smoluchowski equation for a single collision mechanism of a particular class size (k) can be written as:

$$r_k (M^0 L^{-3} T^{-1}) = \frac{1}{2} \alpha_{emp} \sum_{\substack{\forall i,j: \\ V_{P_i} + V_{P_j} = V_{P_k}}} \alpha_{ij} \beta_{ij} n_i n_j - \alpha_{emp} n_k \sum_{\forall i} \alpha_{ik} \beta_{ik} n_i \quad (2.6)$$

where r_k is the net rate of formation of type- k particles from all possible combinations of type- i and type- j particles, which also takes into account the loss of those k -sized particles flocculating with others; n_i (n_j) is the *number concentration* ($[n_i] = M^0 L^{-3} T^0$) of particles of size i (j), V_p is the volume of particles of size i, j or k ($M^0 L^3 T^0$), β_{ij} is the collision frequency function for the long-range transport mechanism under consideration ($M^0 L^3 T^{-1}$), α_{ij} is the dimensionless function accounting for short-range particle-particle interactions and α_{emp} is the empirical, dimensionless correction factor introduced to match experimental data to the equation. Since i and j can take on all possible values (up to some maximum describing the largest particles expected in the suspension), the same collision is counted twice in the summation when $i \neq j$, therefore a factor of $1/2$ is introduced before the first summation to correct for this double counting. The factor of $1/2$ is also needed when $i = j$ because the collision frequency function would otherwise double count these collisions when $i = j$. As for the last term, a size k particle colliding with *any* size i particle (including size k particles as well) will result in a loss of the size k particle, so that summation is take over all i . Because

2. Flocculation Modeling

flocculation of size i and size j particles can proceed by several mechanisms, each of which acts (almost) independently, the overall rate of flocculation of such particles is given by the direct summation of the individual rates associated with each mechanism:

$$r_k^{tot} = r_k^{Sh} + r_k^{DS} + r_k^{Br} + \dots \quad (2.7)$$

where the superscript on each term indicates whether it refers to flocculation brought about by fluid shear (Sh), differential sedimentation (DS), or Brownian motion (Br). In most of the situations of interest these are the only significant long-range transport mechanisms operating, even though novel treatment systems (like electrocoagulation for example) may also exploit electric or magnetic forces, leading to additional transport modes. Each long-range transport mode is associated with a unique collision frequency function, and a corresponding function accounting for forces relevant to short-range transport. Thus, for example, the rate equation for flocculation by fluid shear would be written as follows:

$$r_k^{Sh} = \frac{1}{2} \alpha_{emp} \sum_{\substack{\forall i,j: \\ V_{P_i} + V_{P_j} = V_{P_k}}} \alpha_{ij}^{Sh} \beta_{ij}^{Sh} n_i n_j - \alpha_{emp} n_k \sum_{\forall i} \alpha_{ik}^{Sh} \beta_{ik}^{Sh} n_i \quad (2.8)$$

It is important to note that α_{ij} depends on the long-range transport mechanism as well despite the van der Waals attraction is independent of it (it depends only on the separation distance between the particles and on the size of each particle); that is because α_{ij} accounts also for the viscous forces encountered as water between two approaching particles escapes to let them collide, therefore hydrodynamic interactions must depend on the long-range transport mechanism. For example, particles approaching one another by differential sedimentation are acted on by gravity and, in absence of other particles, have velocities that are fixed by their physical properties (density, size) and are independent of their location. On the other hand, particles approaching one another by fluid shear move along approximately parallel paths, and their relative velocities depend on their location in the moving fluid but not on their physical properties; in other words, if the approaching velocity is not the same between two different mechanisms, hydrodynamic interactions are not the same. Thus, *the short-range hydrodynamic interactions depend not only on the separation distance between two particles, but also on the fluid motion in their proximity and on*

the external forces acting on the individual particles ([2, p. 570]). Accounting for the three mechanisms of flocculation together leads to the following final formulation of the Smoluchowski equation:

$$r_k^{tot} = \frac{1}{2} \alpha_{emp} \sum_{\substack{\forall i,j: \\ V_{P_i} + V_{P_j} = V_{P_k}}} (\alpha_{ij} \beta_{ij})^{tot} n_i n_j - \alpha_{emp} n_k \sum_{\forall i} (\alpha_{ik} \beta_{ik})^{tot} n_i \quad (2.9)$$

where

$$(\alpha_{ij} \beta_{ij})^{tot} = \alpha_{ij}^{Sh} \beta_{ij}^{Sh} + \alpha_{ij}^{DS} \beta_{ij}^{DS} + \alpha_{ij}^{Br} \beta_{ij}^{Br} \quad (2.10)$$

Speaking of characteristic reaction times, every particle size considered in flocculation has its own rate of reaction (quantified by eq. (2.9)), which can be initially negative for some particle sizes and positive for others. Flocculation reduces the number of small particles and increases the number concentration of large particles in the system, but the distinction between "small" (i.e. sizes with initially decreasing particle number concentration) and "large" (sizes that initially grow in number before reaching a point where they start decreasing) can vary according to the suspension. Keeping in mind that each class of the PSD has its own rate equation describing the change in number concentration, under few hypothesis it can be shown that *the characteristic reaction time for a generic particle size is somewhat similar to that of a second-order reaction; that is, the characteristic reaction time of a generic particle size is proportional to the inverse of the product of a rate constant and the initial number concentration, where the rate constant is not truly a constant, since it changes depending on how the other classes change as flocculation proceeds* ([2, p. 571]). Since flocculation proceeds in a way similar to second order reactions, it might be valuable to design flocculation reactors to achieve a flow pattern that is nearly plug flow, which is the kind of mixing that achieves the greatest amount of reaction for dissolved reactants undergoing second-order reactions. Plug flow is characterized by no mixing among fluid elements in the axial direction and infinite (i.e. perfect) mixing in the transverse directions. A plug flow pattern can be nearly achieved in turbulent pipe flow, but the level of turbulence required is likely to cause flocs to breakup, so this type of reactor is generally considered inappropriate for flocculation. On the other hand, slow mixing promotes collisions between particles, therefore flocculators are generally designed/built as reactors with a residence time distribution equivalent to several CFSTRs in series, a more manageable

yet less efficient way to approximate plug flow behavior. In addition, being a second order reaction, the flocculation rate can be increased if the particle concentration in the reactor is increased. A few ways to do it are, for example, to recycle sludge from a subsequent solid/liquid separation stage (e.g. a settling tank), or to induce some kind of flotating action (i.e. an upflow velocity) to counterbalance the settling velocity of flocs, which therefore would stay longer in the reactor (the net effect to be achieved is to make the residence time of particles greater than the residence time of fluid. Several reactors have been developed to artificially keep particle concentration in the flocculator high enough to increase the flocculation rate and hence the overall effectiveness of the process. Particle concentration could also be increased by direct addition of other particles or by inducing precipitation. The most common method is to promote destabilization via sweep floc mechanism, which has dramatic effects on the PSD; particle volume concentration often increases considerably, but without such precipitation, the flocculation of medium-to-low turbidity waters would be very slow even with complete destabilization. Another method to increase particle concentration is to add some large dense solids such as silica or sand, which increase also the effectiveness of the subsequent sedimentation process thanks to the high density of the resulting flocs (*ballasted flocculation*). Without chemical destabilization, extremely few of the close approaches of two particles would result in the formation of a floc. These effects are taken into account by α_{emp} , α_{ij} or both (depending, to a some degree, by modeling choices); the product of the two α terms can vary over several orders of magnitude depending on the degree of destabilization, which therefore represents the main factor affecting the overall performance of a flocculator [2, p. 571].

2.2 Collision Frequency: the Long-Range Force Model

In the following discussions, the collision frequency functions for collisions caused by fluid shear, Brownian motion and differential sedimentation are presented. Details on the derivation of the three collision frequency functions are provided in [2].

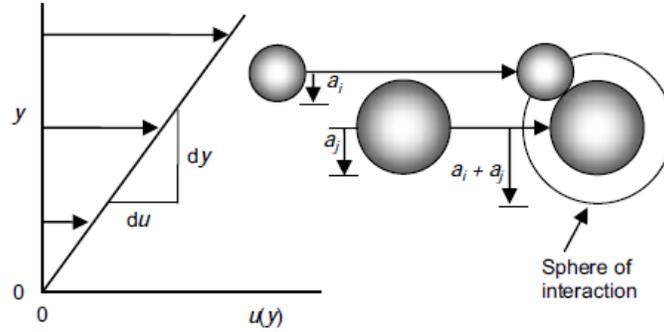


Figure 2.2: Conceptual view of collisions by fluid shear. (*Source:* from [2, p. 572])

2.2.1 Collisions by Fluid Shear

In virtually any reactor - except an ideal plug flow reactor -, different fluid elements travel at different speed; thus, two particles travelling with the fluid but on different streamlines might be brought close to one another, allowing a collision. Flocculation via this mechanism is sometimes referred to as *orthokinetic flocculation*. Let us consider a laminar flow field subjected to shear. Each particle (assumed spherical) is locally assumed to flow on a straight path with a velocity corresponding to that of the streamline passing through its center (fig. 2.2). Hence, particles that are at higher values of y (in fig. 2.2) travel faster than those at lower y values and can catch up to them after some time. If the particles are assumed to continue on rectilinear paths even as they approach one another, a collision will occur if and only if the original vertical distance (i.e. measured in the transverse direction to the streamlines) between their center lines is less than or equal to the sum of their radii. The virtual sphere of radius $(a_i + a_j)$ is called *sphere of interaction* and represents the region around a particle of radius a_i (a_j) containing all the streamlines through which particles of radius a_j (a_i) passing through them will certainly collide with the particle of radius a_i (according to the hypothesis introduced for this mechanism). It can easily be shown (by computing the particle flux through the projected area of the sphere of interaction) that the collision frequency function for fluid shear is equal to

$$\beta_{ij}^{Sh} = \frac{4}{3}G(a_i + a_j)^3 = \frac{1}{6}G(d_i + d_j)^3 \quad (2.11)$$

where G is the *average velocity gradient* of the fluid. In fig. 2.3 are shown calculations

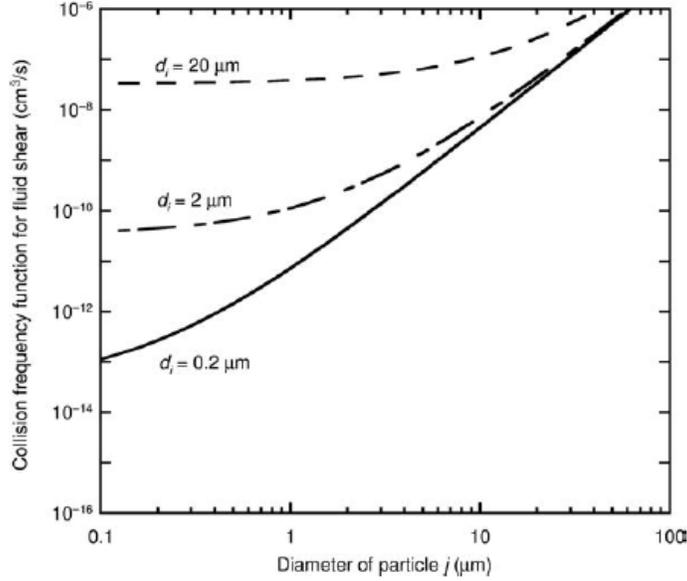


Figure 2.3: Collision frequency function for fluid shear ($G = 25 \text{ s}^{-1}$). (Source: from [2, p. 574])

of the collision frequency function for fluid shear for three different values of particle diameter (particle i in the figure) and a size range of three orders of magnitude for the diameter of the second particle (j , from 0.1 to 100 μm) and a velocity gradient G equal to 25 s^{-1} (common values in operating flocculators range between 15 and 50 s^{-1}). It can be seen (both analytically and graphically) that the collision frequency function for fluid shear is strongly dependent on the size of the larger particles and less influenced by that of the smaller ones. For $d_j < d_i$ (left portion of the plot), the curves become flatter (and hence progressively independent of d_j) as d_i becomes larger than d_j , with asymptotic values that depend on the size ratio between the larger and the smaller particle (d_i/d_j in this case); conversely, at large d_j values ($d_j \gg d_i$, right portion of the graph), the β_{ij}^{Sh} curves become nearly equal (i.e. same asymptotic value), implying that β_{ij}^{Sh} becomes independent of d_i . Introducing the ratio $\phi = d_i/d_j$, the collision frequency function can in fact be rewritten as

$$\beta_{ij}^{Sh} = \frac{1}{6} G d_j^3 (1 + \phi)^3 \quad (2.12)$$

For small ϕ values (i.e. $d_i \ll d_j$, right portion of the graph shown in fig. 2.3), $\beta_{ij}^{Sh} \sim 1/6 G d_j^3$, whereas for higher ϕ values ($d_i \gg d_j \implies \phi \gg 1$, left portion of the

plot), $\beta_{ij}^{Sh} \sim 1/6 G d_j^3 \phi^3$ will depend on both, with ϕ greatly affecting the asymptotic behavior and value of β_{ij}^{Sh} as d_j becomes progressively smaller than d_i . It is interesting to note that the slope on the log-log plot reported in fig. 2.3 approaches a value of 3.0 at high values of d_j , reflecting the exponent in eq. (2.12). The magnitude of G , while not shown in the figure, affects β_{ij}^{Sh} in a linear fashion, hence doubling the velocity gradient to 50 s^{-1} or halving it to 15 s^{-1} would shift the lines up or down, respectively (but the slope would still remain equal to 3.0, being a log-log plot). The effect of G is noticeable, but not dramatic on this logarithmic plot.

2.2.2 Collisions by Differential Sedimentation

Collisions between particles can also occur because they settle at different velocities, as schematically depicted in fig. 2.4, where v represents the settling velocity for the two sizes of particles represented. The particle with the larger settling velocity (i in the figure) will catch up to the smaller one (j in the figure) as they both settle after some time. Similar to the sphere of interaction for collisions by fluid shear, the circle (i.e, oval in the view shown) at the bottom represents the *capture cross-section*, with center located somewhere along the trajectory passing through the center of the larger particle, and radius given by the sum of the radii of the two colliding particles ($a_i + a_j$). If the trajectory of the center of the small particle (j) passes anywhere in that circle, according to this mechanism the particles will certainly collide. Particle j in fig. 2.4 belongs to the so-called *critical trajectory*, which is the trajectory for which the two particles would barely collide and attach to the larger at their equators, as evidenced by the vertical tangent to both spheres. The mathematical derivation of the collision frequency function for this mechanism is similar to that for the fluid shear mechanism; that is, the rate at which larger particles collide with a single, smaller one by differential settling can be written as the product of the concentration of the larger particles, the relative settling velocities and their projected area for potential collisions (i.e. the capture cross-section). In other words [2, p. 575]:

$$\beta_{ij}^{DS} = |v_i - v_j| \frac{\pi}{4} (d_i + d_j)^2 = \frac{\pi g}{72\mu} |(\rho_i - \rho_L)d_i^2 - (\rho_j - \rho_L)d_j^2| (d_i + d_j)^2 \quad (2.13)$$

where g is the gravitational constant, μ is the dynamic viscosity of the liquid, and ρ_p and ρ_L are the densities of the particles and liquid, respectively. The first expression is

general, but the second assumes that both settling velocities follow Stokes' law (which generally applies to particles with Reynolds' number < 0.1). This assumption is usually valid for most of the particles typically encountered in drinking water and wastewater treatment ([2, p. 575]). The collision frequency function for differential sedimentation (β_{ij}^{DS}) is plotted in fig. 2.5 in the same way as fig. 2.3 for β_{ij}^{Sh} (though there is an error related to the fact that the three curves refer to three different values for d_i rather than d_j , which is again plotted in abscissa). It is further assumed that $\rho_i = \rho_j$, for which eq. (2.13) simplifies to

$$\beta_{ij}^{DS} = \frac{\pi g}{72\mu}(\rho_p - \rho_L)(d_i + d_j)^3|d_i - d_j| \quad (2.14)$$

It can be observed (both analytically and graphically) that β_{ij}^{DS} is very small when the two colliding particles are of almost the same size, and is zero when the two sizes are identical (thanks to the term $|d_i - d_j|$ in eq. (2.14)), as shown by the drastic drops schematically depicted in fig. 2.5 for each β_{ij}^{DS} line (which cannot be adequately represented in a log-log plot). If the particles are widely different in size, the collision frequency is dominated by the larger particle for the same mathematical reasons seen for fluid shear, as can be seen by the three curves being almost flat when $d_j \ll d_i$ and being nearly identical when $d_j \gg d_i$. Similar to the collision frequency function for fluid shear, the slope on this log-log plot approaches 4.0 in the limit where d_j/d_i tends to infinity, which is consistent with the order of magnitude of the exponent of the particles diameter in eq. (2.14). Particle density can have a significant effect on the collision frequency as it is the driving force for settling, although such an effect cannot be highly appreciated in a log-log plot (considering particle densities in the range of $1.02 - 2.65 \text{ g/cm}^3$, the lines would just be shifted downward or upward slightly less than one order of magnitude) ([2, p. 576]).

2.2.3 Collisions by Brownian Motion

The third and last long-range transport mechanism here considered is related to a flocculation stage that is often referred to as *perikinetic flocculation*, which refers to collision between particles occurring by Brownian motion. Brownian motion is caused by the random bombardment of ions and molecules of the surrounding fluid on the particle surfaces, and is formally identical to the diffusion of soluble molecules

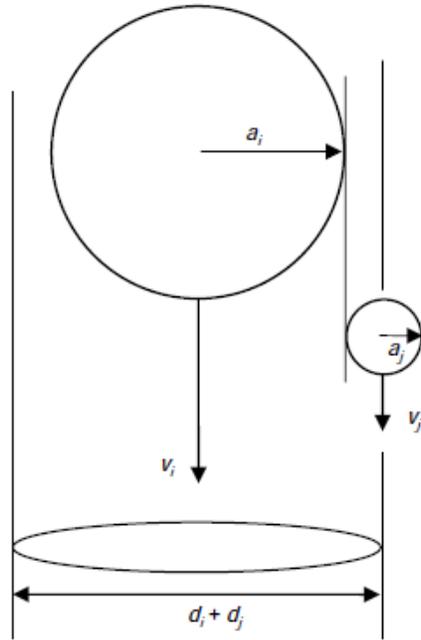


Figure 2.4: Conceptual representation of the collision by differential sedimentation. (Source: from [2, p. 575])

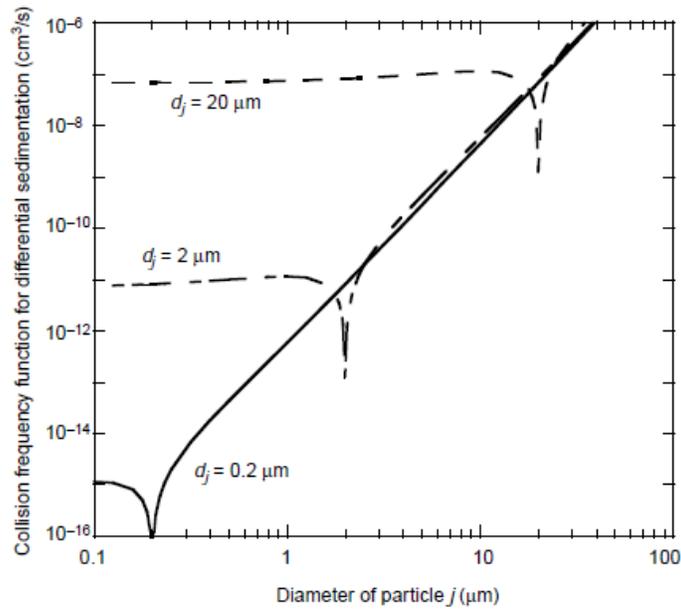


Figure 2.5: Collision frequency function for differential sedimentation ($\rho_i = \rho_j = 1.2 \text{ g/cm}^3$). (Source: from [2, p. 576])

2. Flocculation Modeling

leading to a flux of material down a concentration gradient. Einstein (1905) was the first to propose a description of Brownian motion in terms of a diffusion coefficient combining the concepts of thermal motion of molecules in the liquid (represented by a Maxwell-Boltzmann distribution) and drag on particles (expressed by Stokes' law). Introducing the Stokes-Einstein equation

$$D_p = \frac{k_B T}{3\pi\mu d_p}, \quad (2.15)$$

it is possible to derive the collision frequency function for Brownian motion assuming that the relative motion of two generic particles i and j is the superposition of the motions of each. This assumption implies that the diffusion coefficient describing the relative motion of the two approaching particles can be expressed as the sum of their individual diffusion coefficients, hence:

$$D_{ij} = D_i + D_j = \frac{k_B T}{3\pi\mu} \left(\frac{1}{d_i} + \frac{1}{d_j} \right) \quad (2.16)$$

It can be shown (Smoluchowski, 1917; and Friedlander, 1977) that the collision frequency function for Brownian motion can be expressed as

$$\beta_{ij}^{Br} = \frac{2k_B T}{3\mu} \left(\frac{1}{d_i} + \frac{1}{d_j} \right) (d_i + d_j) = \frac{2k_B T}{3\mu} \frac{(d_i + d_j)^2}{d_i d_j} \quad (2.17)$$

where k_B and T are the Boltzmann's constant and the absolute temperature, respectively. A graphical representation of the collision frequency function for Brownian motion is given in fig. 2.6 for a dispersion at 20°C. Introducing the ratio $\phi = d_i/d_j$, eq. (2.17) can be rewritten as

$$\beta_{ij}^{Br} = \frac{2k_B T}{3\mu} \frac{(d_i + d_j)^2}{d_i d_j} = \frac{2k_B T}{3\mu} \frac{(1 + \phi)^2}{\phi} = \frac{2k_B T}{3\mu} \left(\frac{1}{\phi} + 2 + \phi \right) \quad (2.18)$$

Moreover, β_{ij}^{Br} presents a minimum located at $d_i = d_j$, in fact:

$$\frac{d\beta_{ij}^{Br}}{d\phi} = \frac{2k_B T}{3\mu} \left(-\frac{1}{\phi^2} + 1 \right) = 0 \iff \phi = 1 \iff d_i = d_j, \quad (2.19)$$

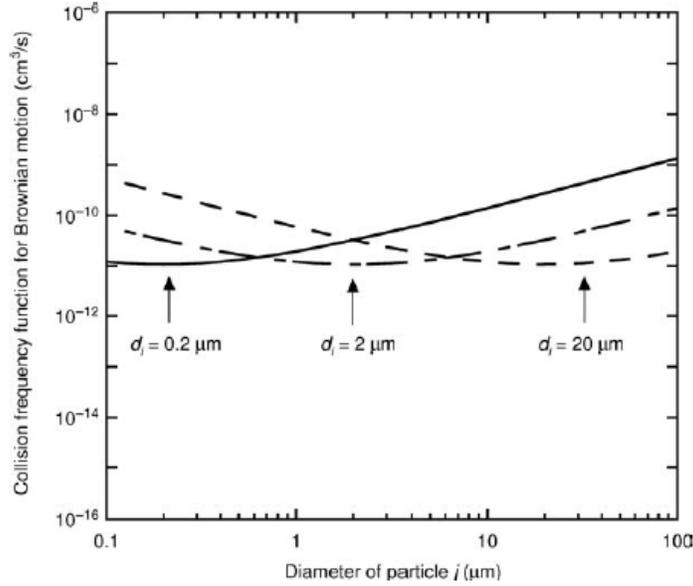


Figure 2.6: Collision frequency function for Brownian motion ($T = 293\text{ K}$). (Source: from [2, p. 578])

At small ϕ values (i.e. $d_i \ll d_j$, right portion of the graph shown in fig. 2.6),

$$\beta_{ij}^{Br} \sim \frac{2k_B T}{3\mu\phi},$$

whereas at high ϕ values (i.e. $d_i \gg d_j$, left portion of the graph),

$$\beta_{ij}^{Br} \sim \frac{2k_B T\phi}{3\mu},$$

Being symmetric around the point of minimum ($d_i = d_j$), β_{ij}^{Br} depends on the sizes of both particles, and the effect of particle size is far less dramatic than in fluid shear and differential sedimentation, as can be seen by the fact that for each choice of d_i shown in fig. 2.6, β_{ij}^{Br} varies by less than two orders of magnitude throughout the range of d_j . It is interesting to note that

$$\beta_{ij}^{Br}(\phi = 1) = \frac{8k_B T}{3\mu}, \quad (2.20)$$

meaning that *the value of the minimum is independent of particle size*. Temperature has a double effect on collisions by Brownian motion as it affects both the thermal

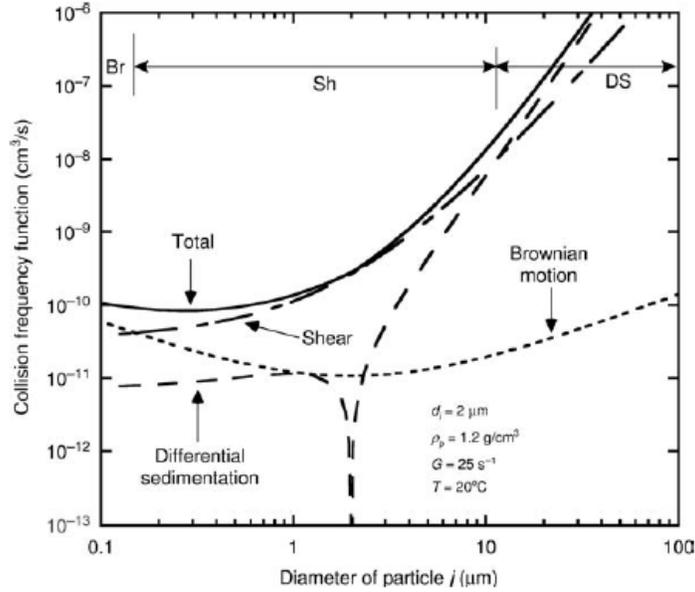


Figure 2.7: Total collision frequency function according to the long-range force model. (*Source:* from [2, p. 579])

kinetic energy of the molecules in the suspension and the dynamic viscosity of the fluid, which decreases as temperature increases.

2.2.4 Consequences and design implications of the long-range force model

Assuming that the three mechanisms described above are independent of one another, it is common to assume that the respective collision frequency functions are additive. Excluding for the moment the effects of short-range forces, one can write:

$$\beta_{ij}^{Tot} = \beta_{ij}^{Sh} + \beta_{ij}^{DS} + \beta_{ij}^{Br} \quad (2.21)$$

A graphical representation of β_{ij}^{Tot} vs. d_j highlighting the relative importance of each mechanism is given by fig. 2.7, which provides a visual comparison of the magnitudes of the three terms appearing in eq. (2.21) under the environmental and operational conditions reported in the graph (which are typically encountered in water treatment plants). It can be seen that close approaches of particles of nearly the same size ($d_j \simeq$

$d_i = 2 \mu m$ in this case) are mostly dominated by Brownian motion and fluid shear, with Brownian motion being the predominant mechanism for particles j smaller than $2 \mu m$ and differential sedimentation being the predominant mechanism when the second particle (j) is quite large. For all sizes in between, approaches by fluid shear are the most likely to occur. The ranges where each mechanism dominates are a function of the choice of the nonvarying particle size (d_i) and all the environmental and operational parameters affecting each mechanism individually (e.g. temperature, viscosity, particle density, velocity gradient). Generally speaking, one can say that the formation of proto-flocs between two small particles is dominated by Brownian motion, while collisions involving a very large particle and a much smaller one are dominated by differential sedimentation, and all others by fluid shear. Looking at all possible combinations of particle sizes, fig. 2.8 shows the regions of dominance for each of the three mechanisms for collisions between any two particle sizes. Only half of the region is represented as the graph would be symmetric around the center line representing equi-sized particles ($d_i = d_j$). It can be seen that when only long-range forces are considered, most collisions are predicted to occur by fluid shear for any reasonable conditions, more or less depending on the specific conditions chosen. For example, increasing the velocity gradient G would obviously increase the region of dominance by shear. Similarly, a decrease in temperature would also increase fluid viscosity and hence reduce both the regions dominated by Brownian motion and differential sedimentation. *In light of this analysis, it is not surprising that design and operation of flocculators has been based almost exclusively on fluid shear as the most important cause of collisions* ([2, p. 580]). Because β_{ij}^{Sh} is linearly dependent on G , the highest mixing intensity that does not result in floc breakup has been thought best by many. Camp (1955), assuming all of the particles in the suspension having the same size, suggested that the dimensionless product $G\tau$ (where τ represents the average hydraulic detention time) could be used as the primary design factor for flocculation basins. In a review of several existing facilities, Camp found that the product $G\tau$ ranged from $\sim 23,000$ to $\sim 210,000$ and suggested this range to be used as a guideline in designing flocculators (and even today, most flocculators are designed to meet this guideline) ([2, p. 580]).

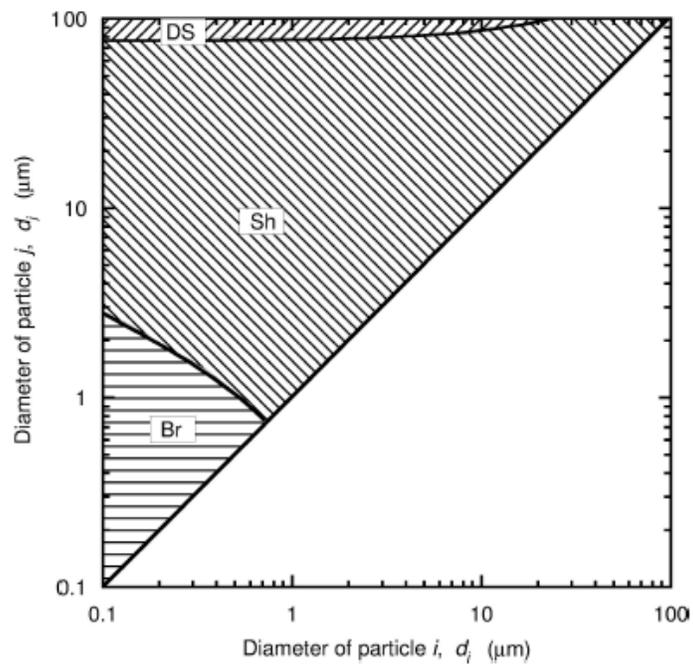


Figure 2.8: Dominant regions for each collision mechanism according to the long-range force model, with borders between regions varying with environmental and operational conditions. (Source: from [2, p. 580])

2.3 Collision Efficiency: the Short-Range Force Model

The collision frequency functions developed in the preceding section describe the movement of particles along paths that are unaffected by the existence of other particles. Such a view is realistic as long as the particles are far from one another, but it becomes unrealistic at short interparticle distances since it does not involve any short-range surface forces (i.e. van der Waals attraction, double layer attraction/repulsion and hydrodynamic effects) [2, p. 581]. These short-range phenomena are represented as force fields, as they depend on both the separation distance and the angle at which the two particles approach one another. Accounting for these short-range forces results in a modification of the collision frequencies developed above, called by some authors as the *Short-Range force model*. These modifications involve explicit numerical simulations done by many investigators under certain hydraulic hypothesis and boundary conditions (like low Reynolds number, no-slip condition for fluid motion around the surface of each particles, etc.), that overall result in a reduction of the predicted frequencies of each interparticle collision mechanism previously introduced, sometimes referred to as *hydrodynamic retardation*. The framework for these calculations is a procedure known as *trajectory analysis*, which consists in evaluating the location of the smaller particle in a coordinate system that moves with the larger one (Lagrangian view), when the smaller particle is subjected to each long-range transport mechanism individually. In other words, the main objective of trajectory analysis in the context of flocculation is to compute the α_{ij} values for particles approaching one another by each of the three long-range transport mechanisms.

This concept is schematically depicted in fig. 2.9, which shows a comparison between the trajectory of a small particle (j in the figure) undergoing differential sedimentation alone according to the long-range model (in which interactions are ignored, a), and the trajectory of the same particle according to the short-range model (which accounts for the hydrodynamic effects and interparticle forces as the two particles approach one another, b). Both particles are settling downward but, to the large particle (i in the figure) with its faster settling velocity, it appears that the small particle is approaching from the bottom. Open trajectories (shown on the left side of the large particle for each model)

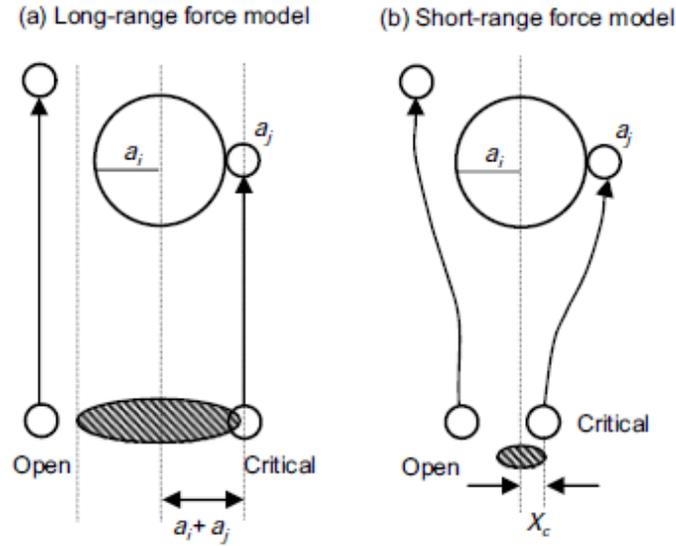


Figure 2.9: Trajectories of particles undergoing differential sedimentation according to the long-range and short-range models. (*Source:* from [2, p. 582])

are those that do not lead to a collision (the particles simply pass by one another). The trajectory with the maximum horizontal separation distance (at an arbitrary large initial vertical separation) that leads to a collision is called *critical trajectory*, and is depicted on the right side for each model. In the long-range force model trajectories are rectilinear, therefore the critical horizontal separation distance between the centers of the two particles (when far removed from one another along the vertical direction) is basically the sum of their radii: any size- j particle whose center passes through the area defined by the circle with radius $a_i + a_j$ in fig. 2.9 (a) will collide with the size- i particle. Conversely, in the short-range force model, trajectories are curved, and the minimum horizontal separation distance is much smaller, as indicated by the radius X_c in fig. 2.9 (b): any size- j particle that passes through the area defined by the circle of radius X_c will certainly collide with the size- i particle. The two models thus differ in the size of the *capture cross-section*, which is therefore overestimated in the long-range force model. In trajectory analysis, the value of X_c is found by repeated trials with different values of the initial horizontal separation distance to find the largest distance that results in a collision. The meaning of α_{ij} is therefore the ratio of the areas of the capture cross-sections that lead to collisions in the curvilinear and rectilinear cases; considering for example differential sedimentation alone, the corresponding collision

Table 2.1: Collision efficiency for fluid shear. (*Source*: from [2, p. 598])

$\log H_A$	a	b	c	d
1	-1.128	2.498	-2.042	0.671
0	-1.228	2.498	-2.042	0.671
-1	-1.482	3.189	-3.468	1.581
-2	-1.704	3.116	-2.881	1.121
-3	-2.523	5.550	-6.098	2.553
-4	-3.723	10.039	-12.569	5.557
-5	-5.775	18.267	-24.344	10.992
-6	-7.037	20.829	-25.589	10.755
-7	-8.773	25.663	-30.703	12.555
-8	-9.733	30.663	-35.703	14.555

$$\text{Sh } \alpha_{ij} = \frac{8}{(1 + \lambda)^3} 10^{a + b\lambda + c\lambda^2 + d\lambda^3}$$

$$H_A = \frac{A}{18\pi\mu d_1^3 G}$$

λ = size ratio ($0 < \lambda \leq 1$); d_1 = diameter of larger particle; A = Hamaker constant; μ = absolute viscosity; G = velocity gradient.

efficiency function is computed as

$$\alpha_{ij}^{DS} = \frac{X_c^2}{(a_i + a_j)^2} \quad (2.22)$$

The approach used to calculate the relative particle motion by fluid shear and the corresponding value of α_{ij}^{Sh} is basically the same, although the details differ because the forces causing the motion are different [2, p. 583]; Adler (BIBLIOGRAFIA) is responsible for the corrections by fluid shear reported here. For Brownian motion, being statistically based, the approach is even more different (the analysis still accounts for the effects of short-range phenomena, but the values are not derived from predicted trajectories of individual particles)[2, p. 583]. The collision efficiency functions (α_{ij}) accounting for hydrodynamic effects and van der Waals attraction (but not double layer interactions) for the three transport modes are shown in fig. 2.10. Details of the development of these figures are given in (BIBLIOGRAFIA), whereas in table 2.1, table 2.2 and table 2.3 are reported calculation equations of α_{ij} for any combination of particle sizes for fluid shear, differential sedimentation and Brownian motion, respectively.

The choice to omit the double layer interactions is related to the fact that on one hand, if destabilization is carried out by charge neutralization and is complete, the double layer repulsion vanishes; on the other hand, double

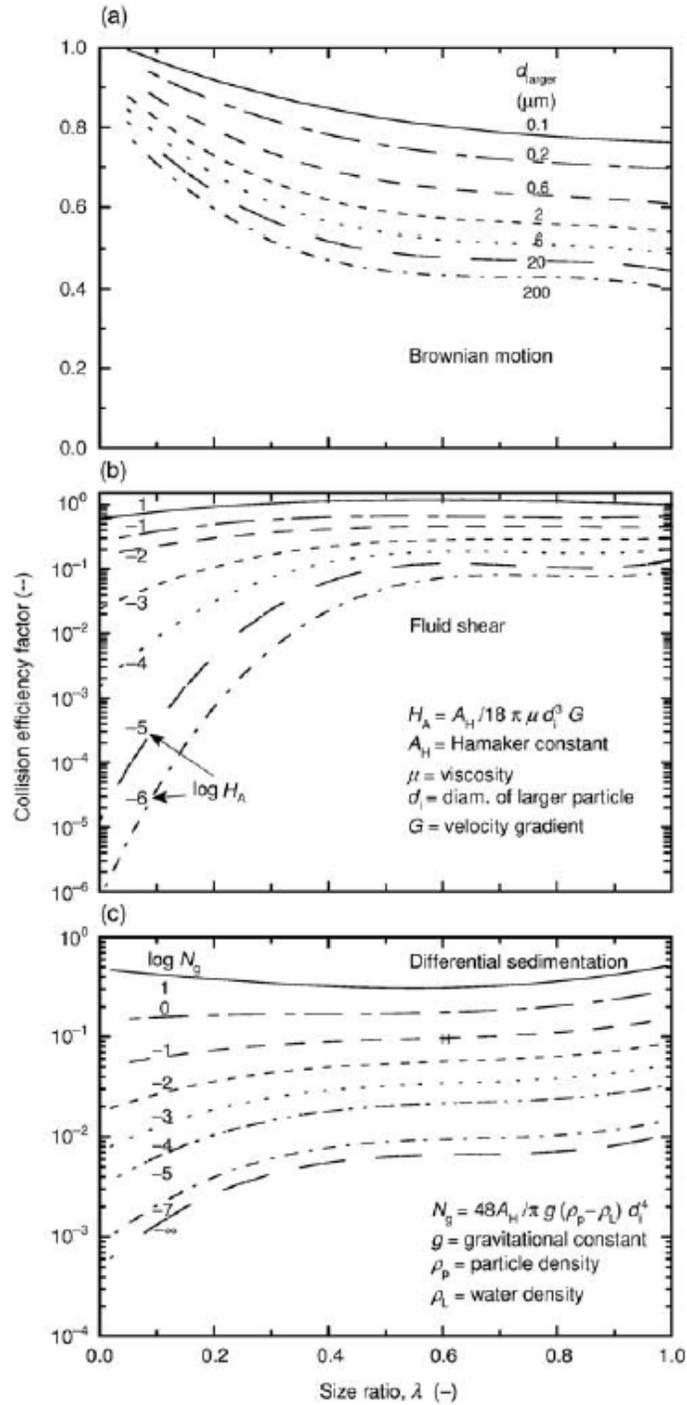


Figure 2.10: Collision efficiency functions in the short-range force model. (Source: from [2, p. 583])

2.3. Collision Efficiency: the Short-Range Force Model

Table 2.2: Collision efficiency for differential sedimentation. (*Source:* from [2, p. 598])

$\log N_g$	a	b	c	d
1	-0.3184	-0.548	0.0865	0.501
0	-0.840	0.423	-1.069	0.930
-1	-1.320	1.318	-2.170	1.361
-2	-1.757	2.137	-3.229	1.794
-3	-2.152	2.880	-4.232	2.230
-4	-2.505	3.547	-5.186	2.668
-5	-2.815	4.137	-6.088	3.108
-6	-3.084	4.652	-6.940	3.551
-7	-3.310	5.090	-7.742	3.996
$-\infty$	-3.928	6.423	-9.449	4.614

$$\text{DS } \alpha_{ij} = 10^{a+b\lambda+c\lambda^2+d\lambda^3}$$

$$N_g = \frac{48A}{\pi g(\rho_p - \rho_f)d_1^4}$$

λ = size ratio ($0 < \lambda \leq 1$); d_1 = diameter of larger particle; A = Hamaker constant; ρ_p , ρ_f = density of particle and fluid, respectively;

g = gravitational constant.

Table 2.3: Collision efficiency for Brownian motion. (*Source:* from [2, p. 597])

Diameter of Larger Particle (μm)	a	b	c	d
0.1	1.025	-0.626	0.516	-0.152
0.2	1.007	-0.860	0.870	-0.322
0.6	0.976	-1.155	1.342	-0.554
1.0	0.962	-1.263	1.522	-0.645
2	0.943	-1.383	1.725	-0.748
6	0.916	-1.533	1.991	-0.887
10	0.905	-1.587	2.087	-0.936
20	0.891	-1.658	2.221	-1.009
60	0.871	-1.739	2.371	-1.090
200	0.863	-1.775	2.439	-1.125

$$\text{Br } \alpha_{ij} = a + b\lambda + c\lambda^2 + d\lambda^3.$$

λ = size ratio ($0 < \lambda \leq 1$).

layer interactions depend on solution and suspension conditions, making them less generalizable than van der Waals or hydrodynamic interactions. In other words, this choice implies to account for double layer interactions by incorporating them into α_{emp} for quantitative work [2, p. 584]. For both fluid shear and differential sedimentation, all of the influencing variables besides the size ratio are incorporated into the following dimensionless numbers that are used for the explicit calculation of the short-range collision efficiency functions:

$$H_A = \frac{A}{18\pi\mu G d_l^3} \quad (2.23)$$

$$N_g = \frac{48A}{\pi(\rho_p - \rho_L)g d_l^4} \quad (2.24)$$

where H_A is related to the fluid shear mechanism, whereas N_g to differential sedimentation. These dimensionless parameters arise from nondimensionalization of the relative equations of motion, and they both include the Hamaker constant (A), the size of the larger particle in the collision (d_l), and parameters associated with the particular collision mechanism. What emerges from fig. 2.10 is that the collision efficiency can be quite low (orders of magnitude smaller than one) for collisions by fluid shear and, to a lesser extent, for collisions by differential sedimentation (especially when the two particles are very different in size). As a result, the overall predicted collision *frequencies* ($\alpha_{ij}\beta_{ij}$) for these two mechanisms are considerably less than when the short range forces were ignored (β_{ij} alone)[2, p. 586]. Conversely, the predicted collision efficiency for collisions by Brownian motion is not nearly as affected, as can be seen in fig. 2.10. This differential effect/action of the short-range phenomena on the three long-range transport mechanisms implies (and imposes) a review of both the charts shown in fig. 2.7 and fig. 2.8 for the specific set of (typical) environmental and operational conditions reported in the figures. It can be seen from fig. 2.11 that in the short-range model far fewer collisions are predicted to occur between particles of vastly different sizes; in this model, particles grow gradually through the size range rather than be scavenged by the largest particles in the suspension [2, p. 587]. Not only is the rate of predicted collisions reduced, but also the relative importance of the three mechanisms is altered for specific couples of sizes of the two particles. For the specific set of conditions reported in fig. 2.12, the image shows a comparison between the regions of dominance for the three transport mechanisms according to the short-range force model

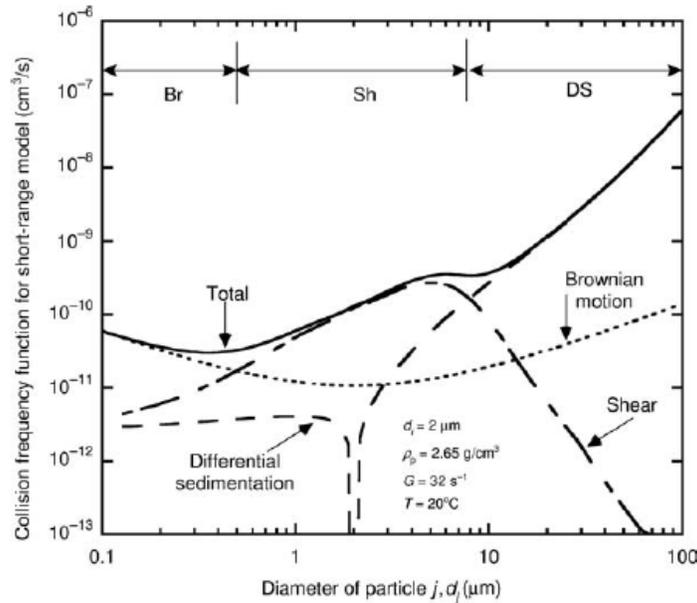


Figure 2.11: Total collision efficiency function according to the short-range force model for the specific set of conditions shown in the picture. (*Source:* from [2, p. 587])

(shaded areas) and the long-range force model, whose boundaries are delineated by the thick lines. While Brownian motion dominates in the left corner (where both particles are small) and differential sedimentation dominates in the upper part of the graph (where one particle is quite large) according to the long-range model, *in the short-range model Brownian motion* (which was the less affected by short-range interactions) *seems to be the primary cause of collisions for a small particle with any other size particle. Fluid shear is only dominant when both particles are greater than $\sim 1 \mu\text{m}$* (more or less depending on the set of parameters chosen) *and nearly the same size, whereas differential sedimentation is dominant when one particle is substantially larger than the other and both are greater than $1 \mu\text{m}$* [2, p. 587].

2.3.1 Consequences and design implications of the short-range force model

The important conclusion is that *fluid shear is predicted to be much less important according to the short-range force model* which, from a design point of view, implies *far less controllability over the collisions that occur*. Most of the collisions between

particles occur by Brownian motion or differential sedimentation, and the parameters that control those collisions are generally not controllable neither by the designer, nor by operators [2, p. 589]. The temperature of water influences its dynamic and kinematic viscosity, and an increase in temperature always results in an increased rate of collisions by all three mechanisms; however, raising water temperature artificially would generally be far too unrealistic and expensive for such an application. The only thing that could be somewhat manipulated to some extent is floc density, which depends also on the type of coagulant used and on the destabilization mechanism(s) exploited. The short-range force model therefore suggests that the mixing intensity is not as important in causing collisions as is predicted by the long-range force model. From a design perspective this implies that *sufficient mixing should be provided to just keep particles in suspension and nothing else*: G values as low as 10 s^{-1} are likely to be able to accomplish that goal in most suspensions [2, p. 589], *more or less depending on the expected floc density* (which implies slightly higher velocity gradients to keep denser flocs in suspension). The proper range for G is then thought to be $10 - 50\text{ s}^{-1}$, where the upper limit stems from a concern to avoid floc breakup [2, p. 589]. Thus, a substantial effort should be made by the designer to ensure a reasonably uniform mixing intensity throughout the flocculator to provide sufficient mixing to keep the flocs in suspension, but not enough to cause their breakup.

2.4 Modeling of Flocculation with Fractal Dimensions

In the last 20-30 years, several advances have been made with respect to particle/floc modeling. The most interesting approach in the context of flocculation modeling is by far the use of fractal dimensions, which allows to represent flocs/aggregates no longer as impermeable and perfectly coalescing spheres (as in Smoluchowski's original formulation), but as porous and self-similar aggregates of discrete primary particles that scale according to a non-Euclidean geometry fig. 2.13. According to the approach formalized by Lee, Bonner *et al.* [3], flocs/aggregates are envisioned as a discrete ensemble of primary (or elementary) particles (for simplicity supposed spherical, impermeable and almost identical) which, following a collision with attachment, grow forming flocs that scale in a non-Euclidian way, preserving both the overall number and volume of primary particles participating in the collision, and the peculiar geometric

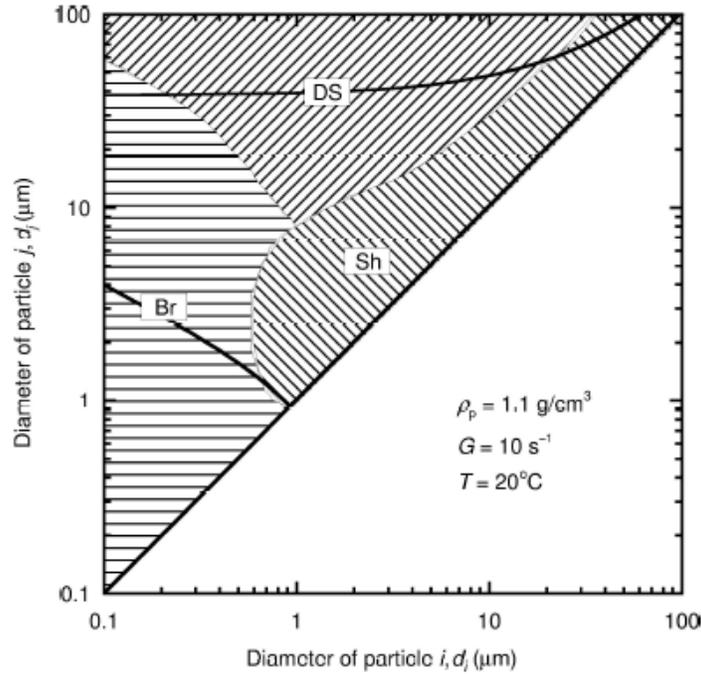


Figure 2.12: Dominant regions for each collision mechanism according to the short-range force model. (Source: from [2, p. 588])

properties of the generating aggregates. Indeed, introducing a scale factor (also known as "Fractal Dimension", D) and a packing factor, it is possible to overcome some important simplifying hypothesis behind Smoluchowski's original formulation, like considering flocs/aggregates as solid impermeable spheres (whose density and hydrodynamic resistance are therefore greater than the real/observed counterparts) or assume a perfectly coalescent behavior that is typical of liquids and emulsions, rather than solids (which usually aggregate extending/branching well beyond the limits dictated by the virtual sphere of equivalent Euclidean volume); in this case, the collision efficiencies predicted by the Euclidean model are lower than expected because of underestimated collision radii. The fundamental hypothesis of conservation of volume of the solids participating in the collisions and to consider negligible, from a probabilistic point of view, collisions that involve more than two aggregates at a time remain still valid. The only hypothesis introduced by the so-called *Coalesced Fractal Sphere* model is to assume both the packing factor and the fractal dimension as characteristic constants of the aggregates. To be precise, the fractal dimension D in

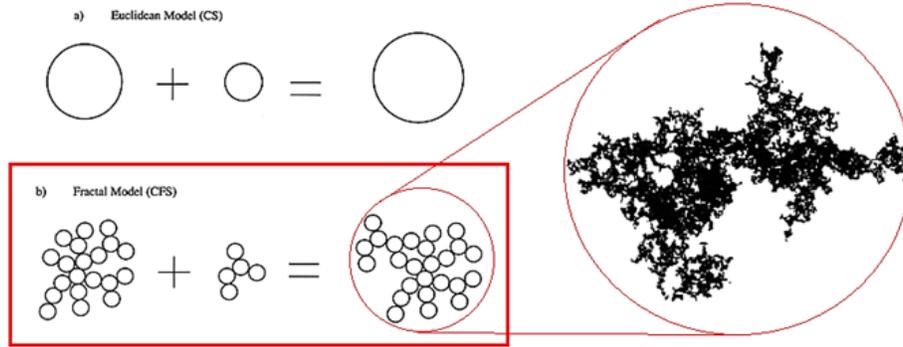


Figure 2.13: Euclidean Model (perfectly coalescing spheres, a) and Fractal Model ("Coalesced Fractal Spheres" model (CFS), b). (*Source:* adapted from [4] and [3])

particular is considered almost as a macroscopic property of aggregates that somewhat reflects the peculiar hydrodynamic and physicochemical conditions in which aggregates form, as amply demonstrated by Logan and other authors in several studies on marine and lake sediments, sewage sludge, etc. According to Logan [4], fractal aggregates possess the following physical properties:

1. Number of primary particles (supposed almost identical and spherical of radius a_p) constituting a fractal aggregate of *collision radius* r (i.e. the radius of the smallest sphere completely containing the aggregate):

$$N_p(r) = \varphi_D \left(\frac{r}{a_p} \right)^D \quad (2.25)$$

where $1 < D \leq 3$ is the fractal dimension, $\varphi_D = \phi^{D/3}$ and ϕ is the packing factor (which is equal to 0.7405 for close cluster packing and 0.637 for random cluster packing). The fractal version of this "Euclidean" packing factor ϕ is φ_D , which also depends on the fractal dimension D . eq. (2.25) represents the fundamental characteristic equation of fractal aggregates, which essentially embodies the fact that if the number of primary particles followed an Euclidean geometry, the fractal dimension D would have been equal to 3. Given the same number of particles contained in the aggregate N_p , a fractal aggregate therefore possess a greater collision radius than its Euclidean counterpart. In order to model also primary particles with the same characteristic equation, the following modification

is proposed:

$$N_p(r) = 1 + \varphi_D \left[\left(\frac{r}{a_p} \right)^D - 1 \right], \quad (2.26)$$

so that for $r = a_p$, N_p is equal to 1. The difference between eq. (2.25) and eq. (2.26) is destined to die out for $N_p \rightarrow \infty$, hence from a practical point of view the two are almost indistinguishable.

2. Volume occupied by solids (i.e. primary particles) in a fractal aggregate:

$$V_s(r) = N_p V_p = \left(\frac{4}{3} \pi \right) a_p^3 \left[1 + \varphi_D \left[\left(\frac{r}{a_p} \right)^D - 1 \right] \right] \quad (2.27)$$

3. Mass of a fractal aggregate (excluding for the moment adsorbed species and precipitates):

$$m_a(r) = \rho_p V_s = \rho_p N_p V_p = \left(\frac{4}{3} \pi \right) a_p^3 \left[1 + \varphi_D \left[\left(\frac{r}{a_p} \right)^D - 1 \right] \right] \quad (2.28)$$

4. Apparent density of fractal aggregates:

$$\rho_{app}(r) = \frac{m_a}{V_a} = \rho_p \left(\frac{r}{a_p} \right)^{-3} \left[1 + \varphi_D \left[\left(\frac{r}{a_p} \right)^D - 1 \right] \right] \quad (2.29)$$

5. Aggregate porosity:

$$\varepsilon(r) = 1 - \frac{V_s}{V_a} = 1 - \left(\frac{r}{a_p} \right)^{-3} \left[1 + \varphi_D \left[\left(\frac{r}{a_p} \right)^D - 1 \right] \right] \quad (2.30)$$

The above expressions describe the relationship between physical and geometric properties of fractal aggregates in the hypothesis that such properties are assumed homogeneous and isotropic over the entire aggregate volume. However, since aggregates grow gradually and continuously during flocculation, it is necessary to pay particular attention to those quantities that depend not only on the magnitude, but also on how these properties are spatially distributed within the volume. A parameter of

fundamental importance is, for example, the *hydraulic permeability* which, as a tensor quantity, depends not only on the spatial coordinates within the aggregate, but also on the direction of the flow. In reality, aggregates are neither homogeneous, nor isotropic (or symmetric); a good compromise could be the approach suggested by Veerapaneni & Wiesner [5], who suggested to model the generic porous fractal aggregate as a sequence of arbitrarily small concentric shells with homogeneous hydraulic permeability (fig. 2.14, right). In this way, aggregates are assumed to be only radially isotropic and homogeneous (spherical symmetry). As for the hydraulic permeability, among the few formulations available in the literature it is suggested the one proposed by Happel [6] ("*sphere-in-cell*" model):

$$k_{Happel} = \frac{2a_p^2 (3 - 4.5\gamma + 4.5\gamma^5 - 3\gamma^6)}{9\gamma^3 (3 + 2\gamma^5)}, \quad (2.31)$$

where $\gamma = (1 - \varepsilon(r))^{1/3} = \gamma(r)$. Since the hydraulic permeability also depends on aggregate size, Veerapaneni and Wiesner suggested to define an average volumetric permeability as

$$k_{Happel}^* = \frac{1}{V_a} \sum_j^{n_{shells}} k_j V_j, \quad \gamma_j = (1 - \varepsilon_j)^{1/3}, \quad (2.32)$$

provided an appropriately small thickness - and hence number - of shells. As already mentioned, unlike porosity and apparent density, hydraulic permeability is a tensor quantity and, as such, in addition to depending on the direction (in this case radial symmetry has been assumed for simplicity), it also depends on how it is spatially distributed within the volume considered. The other properties, if volumetrically mediated, would in fact coincide with the expressions above. The following quantities are finally introduced:

1. Fluid Collection Efficiency (η):

$$\eta = \frac{\text{Flow rate passing through the aggregate}}{\text{Overall approaching flow rate}} \leq 1, \quad (2.33)$$

which represents the ratio between the cross section area (measured at infinite distance) of the streamtube tangential to the aggregate and that of the aggregate

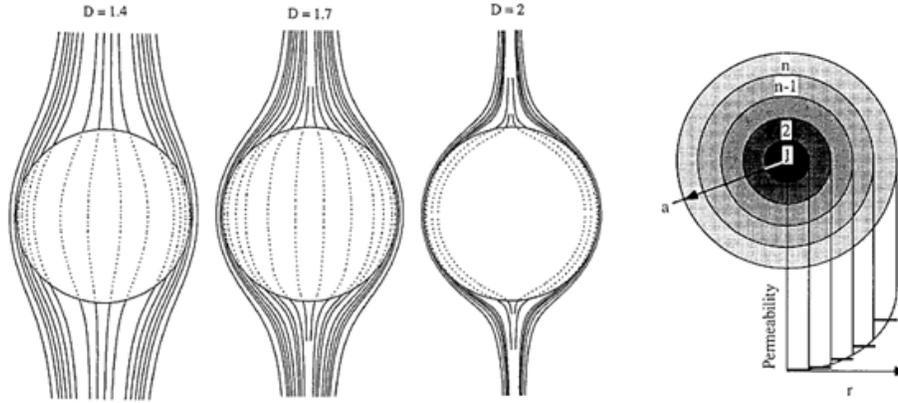


Figure 2.14: Streamlines passing through the same porous aggregate for different values of the fractal dimension D , which deeply affects aggregate permeability (left). Radial variation of permeability within an aggregate for a given fractal dimension D (right). (Source: adapted from [5])

itself (when porosity - and hence permeability - is reduced, the tangential streamtube tapers accordingly).

2. Drag Correction Factor (Ω):

$$\Omega = \frac{\text{Force exerted by fluid on a permeable aggregate of radius } r}{\text{Force exerted by fluid on an impermeable sphere of radius } r} \leq 1, \quad (2.34)$$

which represents the reduction in drag resistance due to aggregate hydraulic permeability.

The mathematical expressions for η and Ω are obtained by solving the flow field inside and outside the aggregate, introducing appropriate boundary conditions and using an adequate filtration model to describe the flow field within the aggregate itself. Since the expressions found in the literature are rigorously valid only under the hypothesis of homogeneous hydraulic permeability, it is possible to use the equivalent average permeability introduced before (k_{Happel}^*):

$$\Omega_{Darcy}^* = \frac{2\xi^{*2}}{2\xi^{*2} + 3}; \quad \eta_{Darcy}^* = \frac{3}{3 + 2\xi^{*2}}; \quad (2.35)$$

2. Flocculation Modeling

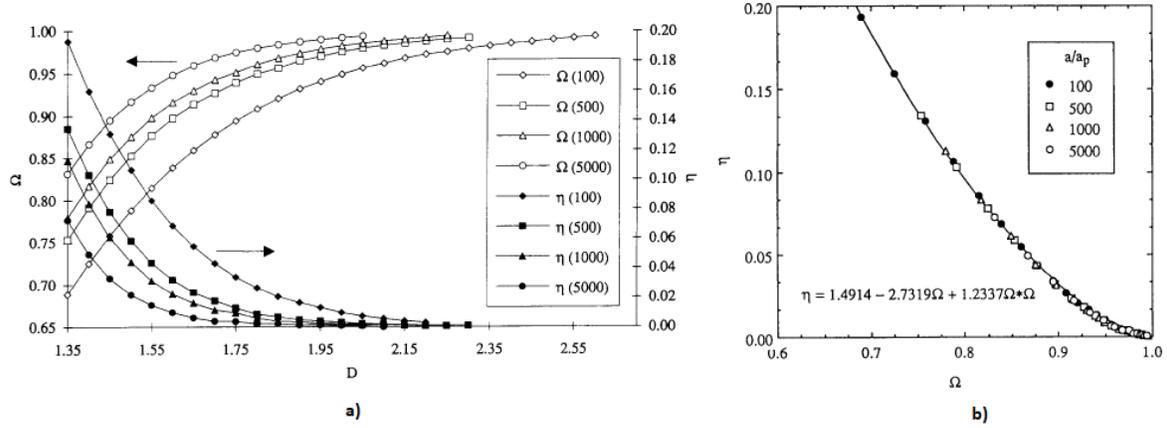


Figure 2.15: Ω and η as functions of fractal dimensions for aggregates of different sizes (expressed in terms of r/a_p ratios) (a) and Ω vs. η at different fractal dimensions for four aggregate sizes (b). (Source: from [5])

$$\Omega_{Brinkmann}^* = \frac{2\xi^{*2} \left(1 - \frac{\tanh \xi^*}{\xi^*}\right)}{2\xi^{*2} + 3 \left(1 - \frac{\tanh \xi^*}{\xi^*}\right)}; \quad \eta_{Brinkmann}^* = 1 - \frac{d}{\xi^*} - \frac{c}{\xi^{*3}}; \quad (2.36)$$

where

$$\xi^* = \frac{r}{\sqrt{k_{Happel}^*}}, \quad (2.37)$$

$$d = \frac{3}{J} \xi^{*3} \left(1 - \frac{\tanh \xi^*}{\xi^*}\right), \quad (2.38)$$

$$c = -\frac{1}{J} \left[\xi^{*5} + 6\xi^{*3} - \frac{\tanh \xi^*}{\xi^*} (3\xi^{*5} + 6\xi^{*3}) \right], \quad (2.39)$$

$$J = 2\xi^{*2} + 3 - 3\frac{\tanh \xi^*}{\xi^*}. \quad (2.40)$$

The quantity ξ^{-2} appearing in 2.37 and in all the equations below is also known as *dimensionless permeability*; eq. (2.35) refers to Darcy's filtration model and is suggested when the fractal dimension D is between 2 and 3 (compact and poorly permeable aggregates), whereas eq. (2.36) refers to Brinkmann's filtration model, which must be preferred when aggregates are significantly porous/permeable (generally when $D < 2$). The above equations can be found in Chellam & Wiesner [7], Neale & Epstein [8] and Adler [9].

To account for "curvilinear effects" in the expressions of the collision frequency functions for permeable fractal aggregates due to short-range phenomena (i.e. double layer interactions, van der Waals attraction and hydrodynamic retardation due to viscous forces), Veerapaneni & Wiesner [5] suggested two possible approaches. One is to make the following corrections to each of the "rectilinear" collision rate kernels for the long-range transport mechanisms previously introduced:

$$\beta_{ij}^{Br} = \frac{2k_B T}{3\mu} \left(\frac{1}{\Omega_i^* r_i} + \frac{1}{\Omega_j^* r_j} \right) (r_i + r_j), \quad (2.41)$$

$$\beta_{ij}^{Sh} = \frac{4}{3} G \left(\sqrt{\eta_i^*} r_i + \sqrt{\eta_j^*} r_j \right)^3, \quad (2.42)$$

$$\beta_{ij}^{DS} = \frac{2\pi g}{9\mu} \left(\sqrt{\eta_i^*} r_i + \sqrt{\eta_j^*} r_j \right)^2 \left| \frac{r_i^2}{\Omega_i^*} (\rho_{eff,i} - \rho_w) - \frac{r_j^2}{\Omega_j^*} (\rho_{eff,j} - \rho_w) \right|, \quad (2.43)$$

$$\beta_{ij}^{tot} = \beta_{ij}^{Br} + \beta_{ij}^{Sh} + \beta_{ij}^{DS}. \quad (2.44)$$

The other basically consists on using the curvilinear collision rate kernels ($\alpha_{ij}\beta_{ij}$) developed in the short-range force model introducing in each expression (tables included) the Hydrodynamic Radius $R_H = \Omega^* r < r$ in each instance of r . The hydrodynamic radius is basically the radius of an equivalent impermeable, spherical aggregate that would feel the same drag force felt by the porous fractal aggregate.

2.5 The Surface Precipitation Model

This section introduces a model that was first proposed by Farley *et al.*[10] in 1985 to model the adsorption of cations onto metal oxyhydroxides. The most commonly used models for adsorption implicitly assume that species do not interact with one another once adsorbed. Furthermore, when considering the adsorption of metal cations onto suspended mineral particles one would expect that, above a concentration in bulk solution approximately equal to the solubility of the corresponding hydroxide, it becomes impossible to discern adsorption from precipitation, since the only thing that we are able to measure is the change in concentration of the metal in bulk solution. The Surface Precipitation (SP) model therefore aims to describe the transition from individual adsorbed metal species to formation of a surface precipitate when the

corresponding adsorption density becomes sufficiently large (a behavior that can be represented like in fig. 2.16). Empirically, it is observed that the amount of adsorbate bound to the particles surface first starts to increase steadily in response to increasing dissolved adsorbate concentration; then, rather than approaching some asymptotic value like in the Langmuir isotherm, it continues to increase without bound, a trend that does not reflect an additional adsorption, but rather the formation of an early precipitate on the solid surface. If the precipitate is a metal hydroxide and solution pH is held constant, the solubility product imposes a limit on the amount of the adsorbate that can be dissolved, hence any addition of dissolved metal leads to precipitation of virtually all the added metal. As mentioned before, if adsorption is quantified by simply analyzing the amount of metal removed from solution, precipitation can be easily mistaken for adsorption, causing the apparent adsorption density to increase indefinitely [11]. According to Benjamin [11, p. 835]:

Few experimental studies have investigated conditions that span the regions where both adsorption and surface precipitation occur. In those studies, although the endpoints of the process (adsorption at low dissolved adsorbate concentrations, and precipitation at high concentrations) are reasonably well defined, the transition between the two states has not been clearly identified. [...] The most widely cited attempt to model the transition between adsorption and surface precipitation was proposed by Farley *et al.* [10] and is based on the formation of an ideal solid solution or mixed solid at the surface. The solid solution is postulated form at very low adsorption densities and to become gradually more enriched with adsorbate as the apparent adsorption density increases.

In the model, «a fixed portion of the adsorbent is assumed to participate to the mixed solid solution. The activity of the adsorbate in the idealized mixed solid phase is presumed to be proportional to its mole fraction in that phase»[11, p. 837]. According to Benjamin, the following expression for the apparent adsorption density attributable to surface precipitation (q_{sp}) can be written:

$$q_{sp} = \frac{\phi(c_{eq}/c_{eq}^{sat})}{1 - c_{eq}/c_{eq}^{sat}} = \frac{\phi K' c_{eq}}{1 - K' c_{eq}}, \quad (2.45)$$

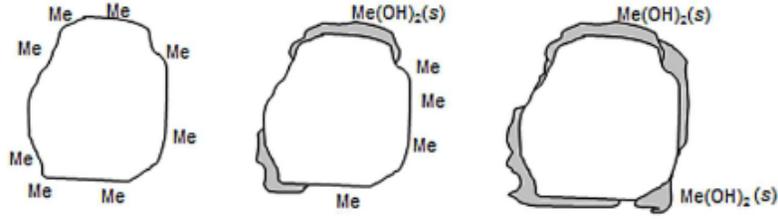


Figure 2.16: Schematic of a metal adsorbate Me that is first bound to discrete sites on the adsorbent, but then forms a surface precipitate covering much of the solid up to a point where the surface of the final product behaves almost like that of a particle of pure metal precipitate. (Source: from [11, p. 836])

where q_{sp} represents the mass of precipitated adsorbate per unit mass of adsorbent in the system, ϕ is the mass of adsorbent that contributes to the mixed solid phase, normalized to the adsorbent surface area, $K' = 1/c_{eq}^{sat}$ and c_{eq}^{sat} is the adsorbate concentration that would be in equilibrium with the pure precipitated solid; if the precipitating solid is a metal hydroxide, then

$$c_{eq}^{sat} = \frac{K_{sp,Me(OH)_2(s)}}{[OH^-]^2}, \quad c_{eq}^{sat} = \frac{K_{sp,Me(OH)_3(s)}}{[OH^-]^3}$$

for divalent and trivalent cations, respectively. According to Benjamin [11, p. 837], eq. (2.45) was derived by combining the solubility product expression with a mass balance, resulting in a simple expression (much simpler than the one derived by Farley *et al.* in [10]) that somewhat recalls the Langmuir isotherm, except for the sign in the denominator (plus for the Langmuir isotherm, minus in the expression for SP). eq. (2.45) includes a constant describing the amount of adsorbent that participates in the reaction forming the solid solution (ϕ , analogous to q_{max} in the Langmuir isotherm) and one related to the thermodynamic driving force for binding (analogous to K_{ads}). The total apparent adsorption density (q_{app}) can be computed as $q_{sp} + q_L$, where q_L is the true adsorption density on the original solid, which is presumed to follow a Langmuir isotherm. Looking at q_{sp} alone in eq. (2.45), it can be noted that at low c_{eq} , q_{sp} is a linear function of c_{eq} , but as c_{eq} approaches $1/K'$, q_{sp} gets very large. The equilibrium dissolved adsorbate concentration c_{eq} cannot in fact exceed c_{eq}^{sat} , since the latter is the adsorbate concentration in equilibrium with the pure solid precipitate.

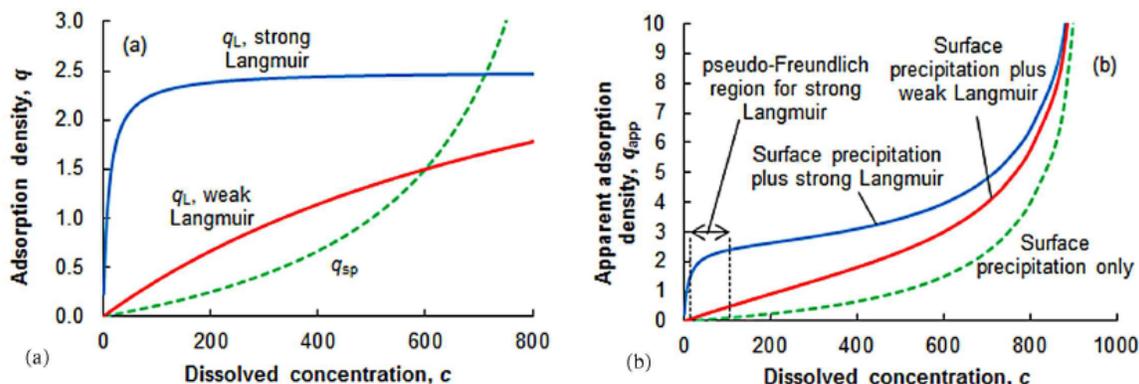


Figure 2.17: Apparent adsorption isotherms for systems in which both Langmuir adsorption and surface precipitation occur simultaneously. On the left, isotherms for adsorption and surface precipitation separately (a); on the right, apparent isotherms for simultaneous adsorption and surface precipitation (b). (*Source*: from [11, p. 839])

As c_{eq} approaches c_{eq}^{sat} , the composition of the mixed surface solid (supposing an ideal solid solution between adsorbent and precipitated adsorbate) approaches that of a pure precipitate of the adsorbate. «Since c_{eq} cannot exceed c_{eq}^{sat} , almost all the adsorbate that is added to solution under these conditions precipitates, and the apparent adsorption density grows without bound»[11, p. 838]. As a conceptual example, in fig. 2.17 is shown the contribution of q_{sp} and q_L to q_{app} for two hypothetical systems, one in which adsorption is relatively strong and one in which it is weak. In both cases, adsorption accounts for most of the adsorbate binding at low c_{eq} , whereas surface precipitation dominates at high c_{eq} . Assuming the two processes occur in parallel (e.g., on different parts of the solid surfaces), the apparent isotherm that would be obtained from experiments is the sum of the Langmuir and surface precipitation isotherms, which is shown for each systems in fig. 2.17(b). For strong Langmuir-binding systems, the empirical isotherm appears very similar to the Langmuir isotherm at low to moderate c_{eq} and like the isotherm for surface precipitation at higher c_{eq} . Thus, in this case (quoting Benjamin): «the isotherm provides a fairly clear indication of the transition from an adsorption-dominated system to one dominated by surface precipitation. However, in the system with weaker adsorption, the curve is qualitatively similar to one for surface precipitation alone». Quoting Benjamin again [11, pp. 839–840]:

Farley et al. suggested that much of the metal ion adsorption data in the

literature might be characterized by relatively strong adsorption and might fall in the region of fig. 2.17 labeled as the pseudo-Freundlich region. The absolute values of c_{eq} and q_{app} in this region depend on the particular metal ion and adsorbent under study, but the shape of the curve would be the same for any system in which adsorption is relatively strong. They also showed that the relationship between q_{app} and c_{eq} in this region can be fit quite well by a Freundlich isotherm. The implication is that a system that obeys a Freundlich isotherm might actually be undergoing a combination of Langmuir-type adsorption and surface precipitation, rather than adsorption onto a surface with a wide range of site affinities.

2.6 Hogg's Model (HM) for Particle Flocculation using Polymeric Coagulants

In 1984, Hogg [12] proposed an interesting reformulation of a simple model (previously introduced by La Mer *et al.* ([13], [14])) of the collision efficiency factor for particle flocculation using polymeric coagulants. Each author agreed with the following modeling assumptions:

- (i) Polymeric coagulants (usually low molecular weight polymers exhibiting a surface charge opposite to that of particles) adsorb in discrete patches on the particles surface (adsorption and charge neutralization).
- (ii) In order to form a "bridge" between two colliding particles, a polymer molecule adsorbed on the surface of one particle must come into contact with a patch of bare surface on the other (adsorption and interparticle bridging).

According to these assumptions, interactions between two adsorbed molecules or two bare surfaces are not expected to cause aggregation. Within this framework, flocculation is therefore regarded as a two-stage process involving:

1. Adsorption of polymer onto the particle surfaces;
2. Interparticle collisions leading to the formation of molecular bridges.

According to La Mer, the collision efficiency factor could be estimated as:

$$\alpha_{ij} = \theta(1 - \theta), \quad (2.46)$$

where θ represents the so-called *surface (or fractional) coverage*, i.e. the fraction of solid surface occupied by adsorbed polymer. Thus, α_{ij} should represent the joint probability that, in the region of contact, the surface of one particle is covered by polymer (θ) while the surface of the other is bare ($1 - \theta$). According to this formula, the maximum collision efficiency is reached when $\theta = 1/2$, for which $\alpha_{ij} = 1/4$. Several criticisms have been raised about this model; first, lots of studies over the past decades have shown how it is possible to reach optimal flocculation even with surface coverages significantly lower than $1/2$ ([15],[16]). Second, from a practical point of view, it is well agreed that flocculation is somewhat "optimal" over a wide range of polymer dosages, whereas according to eq. (2.46) only one optimum is expected (at $\theta = 1/2$). In other words, according to La Mer, for a given suspension there is always one, well defined optimal polymer dosage (which ultimately depends on particle concentration and the degree of affinity between the polymer and the solid surfaces). Third, another study [17] comparing observed flocculation rates with calculated collision frequencies showed that the experimentally fitted collision efficiency factors were closer to 1 than to $1/4$. Hogg [12] criticized eq. (2.46) also from a theoretical point of view, understanding the implicit underlying mathematical assumptions behind such a simple formulation. First of all, «the quantity $\theta(1 - \theta)$ represents the joint probability that the first particle has an appropriately oriented patch of adsorbed polymer facing one of bare surface on the second particle». Clearly the reverse situation is equally likely and would also lead to adhesion, hence α_{ij} should consider both possibilities (i.e. eq. (2.46) should be symmetric, and hence multiplied by a factor of 2). Second, quoting Hogg's words [12]

the probabilities represented by θ and $(1 - \theta)$ refer to specific locations (i.e. contact regions) on the surfaces of the colliding particles. This implies that the orientation of the particles as they approach one another is fixed and that bridging can occur only if they are oriented such that a patch of adsorbed polymer on one particle faces a patch of bare surface on the other. More realistically, it might be expected that particle-particle, particle-polymer and polymer-polymer interactions would lead to reorientation of the

particles and tend to promote those orientations favorable to polymer bridging.

In other words, eq. (2.46) does not take into account those short-range forces that arise when two sufficiently destabilized particles find themselves in close proximity which, after some time in close contact with each other, cause a mutual reorientation that ultimately favors those "bridging-friendly" configurations. Postulating this hypothesis, Hogg [12] reformulated the model introducing the following assumptions:

- (i) Each adsorbed polymer molecule forms discrete patches on the solid surfaces.
- (ii) Adsorbed molecules are distributed at random over the solid surface of each particle to give an overall fractional coverage θ .
- (iii) A generic particle of size d_i has a fixed number n_i of surface sites at which adsorbed polymer molecules can be located. (The word "site" refers to the entire area occupied by a single molecule, not the area occupied by an adsorbed polymer segment).
- (iv) Particles in suspension are in random orientation and are free to rotate in response to hydrodynamic forces or interaction with other particles.

According to this revisited model, interactions between patches of adsorbed polymer or between patches of bare surface are still expected to be repulsive (i.e. they do not cause attachment). The assumption of particles remaining in close proximity to each other long enough to allow their reorientation implicitly implies that for a polymer bridge to form it would simply be necessary for one particle to have *at least* one patch of adsorbed polymer and *at the same time* for the other to have *at least* one patch of bare surface. «It follows that the collision efficiency factor is given by the joint probability of finding adsorbed polymer on one particle and bare surface on the other». The ingenious suggestion given by Metcalfe and Healy as a private communication to Hogg in a review of the original version of [12] was to express the collision efficiency factor as the probability of the complementary event, i.e. that adhesion does *not* occur. Within the above framework, a collision between two particles would not lead to adhesion if:

- both particles were completely covered by adsorbed polymer or
- both particles had completely bare surfaces.

Thus, for two particles of respective sizes d_i and d_j , the probability of nonadhesion is:

$$1 - \alpha_{ij} = (P_1)_{ij} + (P_0)_{ij}, \quad (2.47)$$

where $(P_1)_{ij}$ and $(P_0)_{ij}$ are, respectively, the joint probabilities that both particles are completely covered by polymer or both are completely bare. Following hypothesis (i), (ii) and (iii), if the adsorbed molecules are distributed randomly over the n_i sites on the surface of particle i , the probability that some number m of sites will be occupied can be expressed as a binomial distribution [12]:

$$p(m, n_i) = \frac{n_i!}{m!(n_i - m)!} \theta^m (1 - \theta)^{n_i - m}, \quad (2.48)$$

where the fractional coverage θ represents the (*a priori*) probability that any particular site is occupied. Thus, the probability that particle i is completely covered by adsorbed polymer (all n_i sites occupied) is given by $p(m = n_i, n_i)$, hence:

$$p(n_i, n_i) = \theta^{n_i}.$$

Therefore, the joint probability $(P_1)_{ij}$ that both particles are completely covered by adsorbed polymer is simply:

$$(P_1)_{ij} = \theta^{n_i + n_j}$$

Similarly, the probability that particle i has no adsorbed polymer is given by $p(m = 0, n_i)$, hence:

$$p(0, n_i) = (1 - \theta)^{n_i},$$

and the joint probability $(P_0)_{ij}$ is:

$$(P_0)_{ij} = (1 - \theta)^{n_i + n_j}$$

To summarize, the collision efficiency factor according to Hogg [12] is given by the following formula:

$$\alpha_{ij} = 1 - \theta^{n_i + n_j} - (1 - \theta)^{n_i + n_j}, \quad (2.49)$$

which ultimately depends on the values assumed by parameters n_i and n_j , which are the total number of adsorption sites on the surface of particle i and j , respectively.

Quoting Hogg [12],

they can be determined as the ratio of the surface area of the particle to the area occupied by a single adsorbed molecule. If it is assumed that the adsorbed molecule is coiled to a similar extent to a free molecule in solution, the area occupied can be approximated by

$$A_p \sim 4\bar{R}_g^2, \quad (2.50)$$

where \bar{R}_g^2 is the mean square radius of gyration of the polymer molecules, which is typically observed to be proportional to molecular weight for polymers of the same type.

For a d_k -size particle, the number of adsorption sites n_k can be estimated as:

$$n_k = \frac{k_s d_k^2}{4\bar{R}_g^2}, \quad (2.51)$$

where k_s is a particle shape factor (e.g. $k_s = \pi$ for spheres of diameter d_k), i.e. n_k is proportional to the square of particle size and inversely proportional to polymer molecular weight. The only problem with this formula is that in case of very small particles compared to "large" polymers in solution, n_k would be lower than 1 ($n_k = 1$ would imply a single polymer molecule completely enveloping a particle, which seems fairly unrealistic). On the other hand, $n_k < 1$ would imply the complete opposite situation, i.e. the adsorption of particles onto the polymer, which makes little sense in the context of this model. Paraphrasing Hogg [12], «in typical applications of the flocculation process, particles range in size from about 0.1 to 10 μm , while the molecular weights of flocculants range from about $2 \cdot 10^5$ to $2 \cdot 10^7$. The latter correspond to radii of gyration ranging from about 0.01 to 0.2 μm , therefore values of n_k are expected to vary from about 1 to more than 10^5 ». It is interesting to note that when $n_i = n_j = 1$, α_{ij} reduces to:

$$\alpha_{ij} = 1 - \theta^2 - (1 - \theta)^2 = 2\theta(1 - \theta), \quad (2.52)$$

which is the same as eq. (2.46) (the one originally proposed by La Mer). This means that eq. (2.46) can be interpreted as a very special case of eq. (2.49) for which, considering the explanation given above for cases like $n_k = 1$, it implies that La Mer

implicitly considered only collisions between particles completely covered with polymer and particles with completely bare surfaces, hence excluding all the collisions involving particles just partially covered with polymer.

The effects of n_i and n_j on the magnitude of α_{ij} (or E_{ij}) with fractional surface coverage θ are illustrated in fig. 2.18. Given a generic θ value, one can see in both graphs that α_{ij} is greater for larger particles than for smaller ones, reflecting the fact that it is extremely unlikely that a large particle, with many adsorption sites, would be either completely covered or completely devoid of adsorbed polymer. Efficiency therefore increases with increasing particle size and/or decreasing polymer molecular weight; the larger the particles, the more α_{ij} flattens out to 1 over a progressively large range of θ centered in $\theta = 1/2$. For large particles, the value is essentially 1 over most of the range of surface coverage near $\theta = 1/2$. Scheutjens and Fleer ([18],[19]) and Hogg [20] have shown that polymer adsorption isotherms typically exhibit a behavior like the one schematically depicted in fig. 2.19, which consists of three regions: an essentially linear portion at very low polymer concentration (Henry's Law region), a plateau region at intermediate concentrations, and a third region (at high concentrations) where again the adsorption density increases quite steeply. According to Hogg [12],

Henry's Law region corresponds to the adsorption of isolated molecules on relatively bare surfaces (which tend to spread out), whereas in the plateau region the surface is essentially covered by "spread-out" molecules and further adsorption can occur only by replacement of segments of already adsorbed molecules by segments of molecules from solution. Adsorption of this kind is unlikely to contribute significantly to the bridging process. Consequently, the effective surface coverage to be used in the model probably corresponds to the fraction of the plateau adsorption, i.e.

$$\theta \sim \frac{\Gamma}{\Gamma_{plateau}}. \quad (2.53)$$

To summarize, according to Hogg's model the collision efficiency factors are expected to be very high (close to 100%) over a broad range of surface coverages ($0.1 < \theta < 0.9$ typically), especially for particles that are large compared to the adsorbed polymer molecules. Specifically, α_{ij} increases as the size of the colliding particles increases and

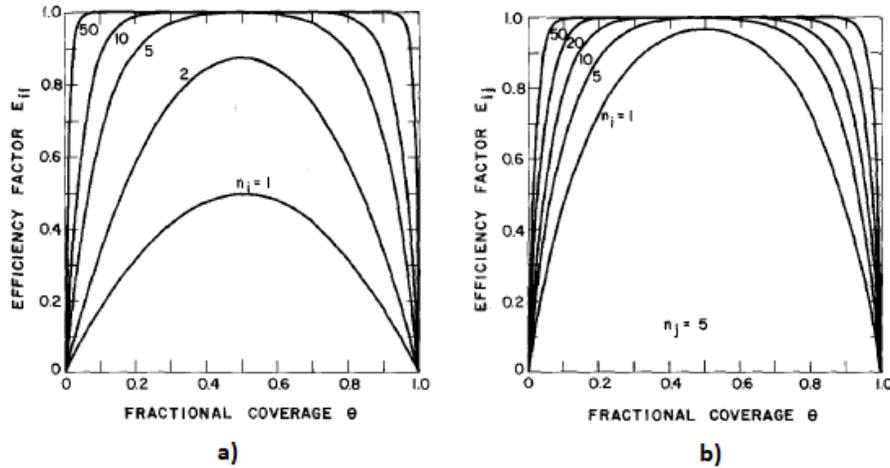


Figure 2.18: Collision efficiency factors for equisized particles (a) and for particles of different sizes (b) as functions of the surface coverage θ . *Source:* from [12]

decreases with increasing polymer molecular weight. It should be noted, however, that the latter effect does not imply that higher molecular weight polymers are less effective flocculants than those of lower molecular weight, because ultimately the average size of the particles in suspension is expected to increase during flocculation. In addition to its effect on the collision efficiency, molecular weight has a major influence on both the adsorption process and the strength of the adhesion between bridged particles (i.e. on the strength of the flocs formed during particle aggregation). While the model presented by La Mer probably underestimates the collision efficiencies by not considering the reorientation of approaching particles, the model proposed by Hogg could perhaps overestimate the collision efficiency, especially at very low and very high surface coverages. Nevertheless, in most situations this model should provide a fairly realistic description of the process. Eventually, an additional parameter (perhaps a symmetric function of θ around $\theta = 1/2$, to compensate for this overestimation at the borders) could be introduced in the expression of n_k given by eq. (2.51), to be fitted with experimental data for model calibration.

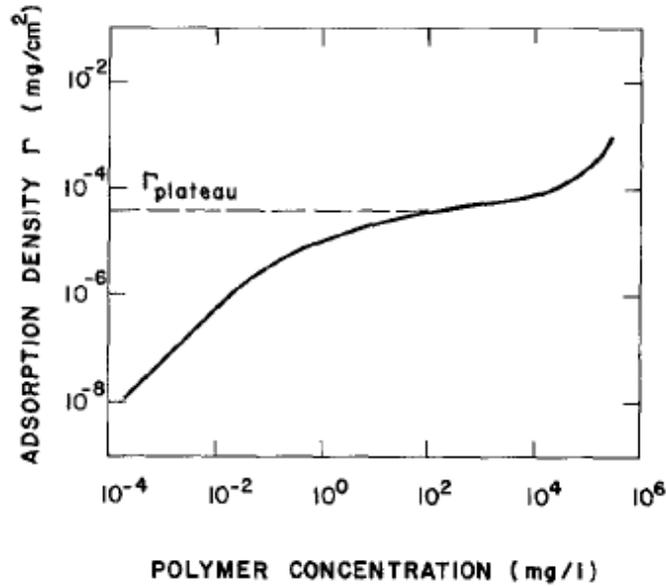


Figure 2.19: Schematic of a typical polymer adsorption isotherm, showing the plateau region at intermediate polymer concentration and surface coverage. *Source:* from [12]

2.7 A proposed Extension of HM to model ballasted flocculation with enhanced coagulation

In this section, an original extension of Hogg's Model is proposed, in order to account for multiple destabilization mechanism at once using a single expression for the collision efficiency function. Rather than an extension, it is actually a complete reformulation that makes use of the probabilistic meaning behind the collision efficiency factor highlighted by Hogg. In particular, the proposed formulation is intended to model the collision efficiency function of a coagulating suspension where particle destabilization is accomplished by both adsorption of a positively charged, low molecular weight polymer and sweep coagulation (through hydroxides precipitation), and flocculation is enhanced by the addition of a ballasting agent prior to coagulation (e.g. microsand or bentonite clay) and of a negatively charged, high molecular weight polymer, to further improve floc strength and settleability.

Let us consider a particle dispersion in which the liquid phase could also be rich in heavy metals and dissolved organic substances, that may or may not adsorb on the

surface of the solid particles. This extension to Hogg's model was in fact originally proposed to simulate the behavior of an existing industrial wastewater effluent (see Chapter 4), whose treatment consists of a primary sedimentation step followed by an enhanced coagulation-flocculation stage with bentonite addition, which serves both as a *ballasting agent* to improve the flocculation rate and flocs settleability (achieving at the same time a slight COD removal through the adsorption of dissolved organic substances) and as a *seeding agent* for hydroxides precipitation. The coagulant used is a liquid mixture of ferric chloride (>50%) and cationic polyamines (<5%), while the flocculant aid is a high molecular weight anionic polyelectrolite. Hydroxides precipitation is induced by adding lime between before the flocculant. The following assumptions - which could eventually be relaxed or slightly modified to account for different wastewater characteristics and/or process configurations - are introduced:

- (i) It is assumed that the residual solids content found downstream the primary sedimentation stage consists mainly of unsettled (colloidal) precipitates (i.e. precipitates that are not grown enough to reach significant settling velocities).
- (ii) After bentonite addition, precipitates are thought to re-dissolve, adsorb and precipitate again onto the bentonite particles (if enough contact time and mixing energy are provided). Bentonite concentration is therefore assumed to be the primary factor affecting particle number concentration (seeding agent).
- (iii) It is assumed that precipitation occurs mainly via heterogeneous nucleation. Hydroxides are thought to provide additional adsorption sites once precipitated on the bentonite particles (uneven coating).
- (iv) There is a *multisite, competitive* Langmuir adsorption behavior between the different adsorbable species, i.e. between the polymeric coagulant - hereinafter referred to as "PA" (assuming it is actually a polyamine) -, the metals of interest - including the eventual metal coagulant, mixed with the polymeric one - and the adsorbable fraction of the soluble COD (hereinafter referred to as COD_s^*).
- (v) It is assumed that heavy metals can adsorb to some extent on the surface of the bentonite particles, with an increase in their removal with increasing bentonite concentration and/or increasing metal concentration (at higher concentrations) explainable through an adsorption isotherm compatible with that described in

2. Flocculation Modeling

the Surface Precipitation (SP) Model.

- (vi) It is assumed that PA and COD_s^* can adsorb either on the surface of the bentonite particles or on that of precipitates, eventually with different affinities; metals are assumed to adsorb only on the former.
- (vii) It is assumed that the anionic polyelectrolyte (hereinafter referred to as "F", which stands for "Flocculant aid") can adsorb only on the *bare surface* exposed by precipitates and PA (regardless of where PA has adsorbed), possibly with different affinities between the two. By "bare surface" is meant a surface not occupied by either PA, COD_s^* or F itself. The choice to not consider also surface sites occupied by adsorbed cations as potential adsorption sites for "F" stems from the fact that the SP model already accounts for a fictitious precipitation of metals at concentrations below their corresponding solubility limit, which ultimately depends on both the solution pH and temperature.
- (viii) It is assumed for simplicity that precipitates have an almost identical adsorptivity for PA and F (which is different for the two species).
- (ix) Finally, all mass transfer kinetics (including surface precipitation) are modeled using a Linear Driving Force model (LDF).

In fig. 2.20 is shown a conceptual diagram of the surface coverage θ of each species mentioned above, grouped by type of adsorption sites exposed on the particle/floc surface (either bentonite or precipitates) on which each species has adsorbed. Sites occupied by the flocculant aid (F) are represented with a solid color fill. As can be seen, F is allowed to adsorb only on the additional surface sites provided by the precipitates (F(SP)) and on those already occupied by polyamine (F(PA)), which in turn can be adsorbed on either bentonite surface sites (PA,b) or precipitates' surface sites (PA,sp). Finally, the term "0" is used to refer to patches that are not covered by flocculant (i.e. the aforementioned "bare surfaces", either of bentonite, precipitates or polyamine).

According to the formulation proposed by Benjamin for the Surface Precipitation model, the pseudo-equilibrium underlying (EQUAZIONE) should realistically be the following:

$$q_{sp,Me}^{eq} (c_{Me}^{sat,eq} - c_{Me}^{eq}) = \varphi SSA_b \Gamma_{b,MAX} \cdot c_{Me}^{eq}, \quad (2.54)$$

2.7. A proposed Extension of HM to model ballasted flocculation with enhanced coagulation

where Γ and q are the adsorption densities expressed in mol/m^2 and mol/g of adsorbent, respectively. SSA_b represents the specific surface area of the adsorbent (bentonite in this case, m^2/g), whereas $\Gamma_{b,MAX}$ is the maximum (molar) adsorption density of the adsorbent (mol/m^2). φ can be interpreted as a degree of affinity/similarity between the hydroxide precipitates and bentonite surface sites (it could be expressed for example as the ratio between the maximum adsorption density per unit mass of the former and the maximum adsorption density per unit mass of the latter). Moles are a better unit of measurement than mass when considering competitive and/or multisite adsorption with adsorbates that vary significantly in size. Looking at $\Gamma_{b,MAX}$ in eq. (2.54), one can write

$$\Gamma_{b,MAX} = \Gamma_{b,0}^{eq} + \Gamma_{b,Me}^{eq}$$

i.e. when considering a single adsorbable-precipitable species alone, the overall sites

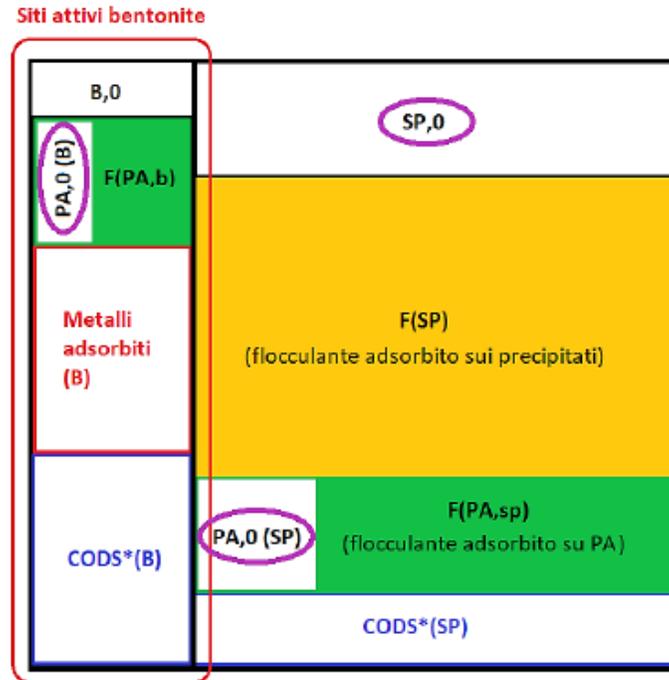


Figure 2.20: Schematic of the surface coverages of a particle/floc with adsorption sites grouped by type of adsorbent exposed on the surface, i.e. bentonite ("b") or precipitates ("SP" - "surface precipitates"). The surface sites responsible for particle adhesion are circled in purple.

available for surface precipitation are then expected to coincide with all those made available by the adsorbent material, which in this case are both the bentonite sites still

2. Flocculation Modeling

vacant ($\Gamma_{b,0}^{eq}$) and those occupied by the adsorbed metal ion ($\Gamma_{b,Me}^{eq}$). Furthermore,

$$\Gamma_{sp,Me}^{eq} = \frac{q_{sp,Me}^{eq}}{SSA_b} \left(\frac{mol_{precipitate}}{m^2_{adsorbent}} \right),$$

so that eq. (2.54) can be rewritten as:

$$\Gamma_{sp,Me}^{eq} (c_{Me}^{sat,eq} - c_{Me}^{eq}) = \varphi (\Gamma_{b,0}^{eq} + \Gamma_{b,Me}^{eq}) c_{Me}^{eq}, \quad (2.55)$$

The left hand side in eq. (2.55) therefore represents the driving force that pushes towards the dissolution of the precipitates that may already be present on the surface (due to local supersaturation at the solid-liquid interface), whereas the right hand side that pushing towards the surface precipitation (wich depends both on the total surface sites available for the heterogeneous nucleation of precipitates and on the concentration of metal in bulk solution). When considering more than one adsorbable-precipitable species, the pseudo-equilibrium of surface precipitation could perhaps be rewritten as follows:

$$\Gamma_{sp}^{eq} \chi_{Me_i}^{sp,eq} (c_{Me_i}^{sat,eq} - \chi_{Me_i}^{l,eq} c_{tot}^{eq}) = \varphi_i (\Gamma_{b,0}^{eq} + \Gamma_{b,Me_i}^{eq}) \chi_{Me_i}^{l,eq} c_{tot}^{eq}, \quad (2.56)$$

where:

- Γ_{sp}^{eq} is the "equilibrium" adsorption density of all the precipitates (moles of precipitates per m^2 of adsorbent);
- $\chi_{Me_i}^{sp,eq}$ is the "equilibrium" mole fraction of precipitate "i" in *ideal solid solution* with the others;
- $(\Gamma_{b,0}^{eq} + \Gamma_{b,Me_i}^{eq})$ is the "equilibrium" adsorption density (mol/m^2) of all sites available/favorable to the surface precipitation of metal "i";
- $\chi_{Me_i}^{l,eq}$ is the equilibrium mole fraction of metal "i" in bulk solution.

A few observations:

1. All gammas (Γ) are normalized with respect to the surface of the adsorbent, thus it should not be surprising if Γ_{sp}^{eq} can become even greater than $\Gamma_{b,MAX}$. In fact, while the surface of the bentonite particles remains constant, that of the particle-precipitates ensemble grows proportionally to the amount of precipitated

hydroxides, hence the sum of the adsorption densities of all the available surface sites (bentonite + precipitates) cannot be constant, being normalized by the bentonite specific surface area (which is fixed).

2. It should be noted that the surface precipitation model is mathematically valid only for molar concentrations of metals in bulk solution below their corresponding solubility limit (i.e. as long as $c_{Me_i}^{eq} < c_{Me_i}^{sat,eq}$, $\forall i$); it is indeed thermodynamically impossible for concentrations in bulk solution *at equilibrium* to be greater than those at saturation. Since precipitation is also assumed to be limited by mass transfer, in order to keep the difference ($c_{Me_i}^{sat,eq} - c_{Me_i}^{eq}$) as positive as possible during simulations, each of the reacting species introduced in the system (e.g. metal coagulant, lime, etc.) should have a dissolution rate compatible with the surface precipitation rate, in order to avoid abrupt variations in the saturation concentrations at equilibrium due to abrupt variations in pH or molar concentrations in bulk solution.

Given all the above assumptions, the following system of algebraic equations is proposed to model the *equilibrium* conditions:

- Available adsorption sites:

$$\Gamma_{TOT}^{eq} = \Gamma_{b,MAX} + \Gamma_{sp}^{eq} \quad (2.57)$$

- Bare bentonite adsorption sites:

$$\Gamma_{b,0}^{eq} = \Gamma_{b,MAX} - \sum_j \Gamma_{b,Me_j}^{eq} - \Gamma_{b,PA}^{eq} - \Gamma_{b,COD_s^*}^{eq} \quad (2.58)$$

- Bare precipitates adsorption sites:

$$\Gamma_{sp,0}^{eq} = \Gamma_{sp} - \Gamma_{sp,PA}^{eq} - \Gamma_{sp,COD_s^*}^{eq} - \Gamma_{sp,F}^{eq} \quad (2.59)$$

- Adsorbed metals at equilibrium:

$$\Gamma_{b,Me_i}^{eq} = \Gamma_{b,0}^{eq} \lambda_{Me_i} K_{Me_i}^b \chi_{Me_i}^{l,eq} c_{tot}^{eq} \quad (2.60)$$

- Bentonite surface sites occupied by free adsorbed polyamine at equilibrium:

$$\Gamma_{b,PA(0)}^{eq} = \Gamma_{b,PA}^{eq} - \Gamma_{b,F(PA)}^{eq} = \Gamma_{b,0}^{eq} \lambda_{PA} K_{PA}^b \chi_{PA}^{l,eq} c_{tot}^{eq} \quad (2.61)$$

- precipitates' surface sites occupied by free adsorbed polyamine at equilibrium:

$$\Gamma_{sp,PA(0)}^{eq} = \Gamma_{sp,PA}^{eq} - \Gamma_{sp,F(PA)}^{eq} = \Gamma_{sp,0}^{eq} \lambda_{PA} K_{PA}^{sp} \chi_{PA}^{l,eq} c_{tot}^{eq} \quad (2.62)$$

- Adsorbable fraction of the soluble COD adsorbed on bentonite surface sites at equilibrium:

$$\Gamma_{b,COD_s}^{eq} = \Gamma_{b,0}^{eq} \lambda_{COD_s} K_{COD_s}^b \chi_{COD_s}^{l,eq} c_{tot}^{eq} \quad (2.63)$$

- Adsorbable fraction of the soluble COD adsorbed on precipitates' surface sites at equilibrium:

$$\Gamma_{sp,COD_s}^{eq} = \Gamma_{sp,0}^{eq} \lambda_{COD_s} K_{COD_s}^{sp} \chi_{COD_s}^{l,eq} c_{tot}^{eq} \quad (2.64)$$

- Surface Precipitation equilibria:

$$\Gamma_{sp,0}^{eq} \chi_{Me_i}^{sp,eq} (c_{Me_i}^{sat,eq} - \chi_{Me_i}^{l,eq} c_{tot}^{eq}) = \varphi_i (\Gamma_{b,0}^{eq} + \Gamma_{b,Me_i}^{eq}) \chi_{Me_i}^{l,eq} c_{tot}^{eq} \quad (2.65)$$

- Flocculant adsorbed on precipitates' surface sites at equilibrium:

$$\Gamma_{sp,F(sp)}^{eq} = \Gamma_{sp,0}^{eq} \lambda_F K_F^{sp} \chi_F^{l,eq} c_{tot}^{eq} \quad (2.66)$$

- Polyamine adsorbed on bentonite surface sites that is covered by flocculant at equilibrium:

$$\Gamma_{b,F(PA)}^{eq} = \Gamma_{b,PA(0)}^{eq} \lambda_F K_F^{PA} \chi_F^{l,eq} c_{tot}^{eq} \quad (2.67)$$

- Polyamine adsorbed on precipitates' surface sites that is covered by flocculant at equilibrium:

$$\Gamma_{sp,F(PA)}^{eq} = \Gamma_{sp,PA(0)}^{eq} \lambda_F K_F^{PA} \chi_F^{l,eq} c_{tot}^{eq} \quad (2.68)$$

- Mole fractions:

$$\sum_i \chi_{Me_i}^l + \chi_{PA}^l + \chi_{COD_s^*}^l + \chi_F^l = 1 \quad (2.69)$$

$$\sum_j \chi_{Me_j}^{sp} = 1 \quad (2.70)$$

In all of the above equilibria, λ is a coefficient which takes into account the relative size of each species in terms of how many adsorption sites it occupies/takes once adsorbed (\sim *multidentate adsorption*). Each adsorption constant K_i^j will generally depend on the affinity between adsorbate i and adsorbent j and on the pH and temperature of bulk solution. Similarly, metals saturation concentrations ($c_{Me_i}^{sat,eq}$) will depend on the solubility product of the corresponding hydroxide (K_{sp}), which depends on pH and temperature too.

Once solved, the above system of algebraic equations provides a mathematical expression for each equilibrium adsorption density Γ^{eq} appearing from eq. (2.57) to eq. (2.68) in terms of equilibrium molar concentrations in bulk solution. Because of the shape of the surface precipitation isotherm, it is not possible to reverse the above equilibria to find an explicit expression for the equilibrium molar concentrations in bulk solution c^{eq} . Hence, according to the LDF model, the rates of change of molar concentrations in the solid phase should be given in the following form:

$$\frac{d\Gamma_{ij}(t)}{dt} = -k_{ij}^{LDF} [\Gamma_{ij}(t) - \Gamma_{ij}^{eq*}(t)], \quad (2.71)$$

where:

- $\Gamma_{ij}^{eq*}(t)$ represents the adsorption density of species i on particle/floc j that *would be* in equilibrium at time t with the *instantaneous* superincumbent molar concentration of the same species in bulk solution ($c_i(t) = \chi_i^l(t)c_{tot}(t)$) at time t ;
- k_{ij}^{LDF} is the effective LDF mass transfer coefficient, which depends on temperature, local hydrodynamic conditions (i.e. Peclet and Reynolds numbers) and relative size of species i and particle/floc j ;
- $\Gamma_{ij}(t)$ is the current adsorption density of species i on particle/floc j at time t .

2. Flocculation Modeling

Thus, the rate of change of molar concentrations in bulk solution should be given by:

$$\frac{dc_i(t)}{dt} = - \sum_j n_j(t) A_j^{eff}(t) \frac{d\Gamma_{ij}(t)}{dt} = \sum_j n_j(t) A_j^{eff}(t) k_{ij}^{LDF} [\Gamma_{ij}(t) - \Gamma_{ij}^{eq*}(t)], \quad (2.72)$$

where:

- n_j is the number concentration of j -sized aggregates (assuming to write a Discretized Population Balance Equation to model the evolution of the Particle Size Distribution in discrete form);
- A_j^{eff} is the effective surface area of the adsorbent which, in order to account for filtration effects within the porous aggregates, could be written as follows

$$A_j^{eff} \simeq \eta_j^* (m_p N_{p_j} SSA_b) + (1 - \eta_j^*) \left[SSA_b m_p \frac{(r_j \Omega_j^*)^2}{a_p^2} \right], \quad (2.73)$$

where SSA_b is the specific surface area of the bentonite particles (m^2/g), m_p is the mass of a single bentonite particle, N_{p_j} is the number of primary (bentonite) particles contained in a j -sized aggregate (g). The first addend represents the area effectively "touched"/felt by the flow that actually permeates through the aggregate in unit time, which coincides with that of all the primary particles that constitute it. The second addend refers instead to the area effectively felt by the unpermeated flow, which is represented by the surface area of an impermeable aggregate of equivalent radius equal to the Hydrodynamic Radius of the aggregate ($r_j \Omega_j^* < r_j$). The adsorbent area is therefore splitted into these two terms in order to account for the effects that particle aggregation and floc growth have on the evolution of the effective surface area available for adsorption during flocculation. It should be noted that when $r_j = a_p$ (i.e. when flocs are primary particles themselves), $\eta_j^* = 0$ and $\Omega_j^* = 1$ (impermeable particles), the available surface area of the adsorbent coincides with that of all the bentonite particles in suspension ($n_j A_j^{eff} = n_j SSA_b m_p$).

As an example, the temporal variation of aggregate j mass due to the surface

precipitation of metal i is therefore given by:

$$\left(\frac{dM_j}{dt}\right)_{Me_i(OH)_x}^{sp} = A_j^{eff}(t)k_{ij}^{LDF} \cdot MM_{Me_i(OH)_x} [\Gamma_{sp}^{eq}(t)\chi_{Me_i}^{sp,eq}(t) - \Gamma_{sp}(t)\chi_{Me_i}^{sp}(t)] \quad (2.74)$$

Introducing the surface coverage of species m on aggregate k , θ_k^m , as

$$\theta_k^m(t) = \left(\frac{\Gamma_m(t)}{\Gamma_{TOT}(t)}\right)_k = \left(\frac{\Gamma_{b,m}(t) + \Gamma_{sp,m}(t)}{\Gamma_{b,MAX} + \Gamma_{sp}(t)}\right)_k, \quad (2.75)$$

the following "critical scenarios" are proposed to evaluate the probability of non-adhesion according to Hogg's framework:

- (i) The colliding particles/aggregates are both completely covered of polyamine (excess of positive surface charges that leads to a repulsion due to charge reversal);
- (ii) Both aggregates have precipitates' surface sites completely covered (if present) and at the same time there is no free polyamine available (either because it is absent or because it is completely saturated of anionic polyelectrolyte).

In all other cases it is assumed that the two colliding particles/aggregates, once they have reached a sufficiently short separation distance, have enough time to reorient themselves until they find mutual patches favorable to attachment. It is therefore assumed that only precipitates and the polyamine are the primary factors directly affecting particle destabilization (i.e. coagulation). The anionic polyelectrolyte is assumed to have no direct influence on coagulation, as it will be explained in more detail shortly. Given the above assumptions, the collision efficiency function of two colliding particles/aggregates i and j can be expressed as:

$$\alpha_{ij} = 1 - (\theta_i^{PA(0)})_{\nu_i^*} \cdot (\theta_j^{PA(0)})_{\nu_j^*} - (\theta_i^{-sp,0 \& -PA(0)})_{\nu_i^*} \cdot (\theta_j^{-sp,0 \& -PA(0)})_{\nu_j^*}, \quad (2.76)$$

where:

$$\theta_k^{PA(0)} = \left(\frac{\Gamma_{b,PA(0)} + \Gamma_{sp,PA(0)}}{\Gamma_{TOT}}\right)_k, \quad (2.77)$$

$$\theta_k^{-sp,0 \& -PA(0)} = \left[\frac{\Gamma_{b,MAX}}{\Gamma_{TOT}} \left(1 - \frac{\Gamma_{b,PA(0)}}{\Gamma_{b,MAX}}\right) + \frac{\Gamma_{sp}}{\Gamma_{TOT}} \left(1 - \frac{\Gamma_{sp,PA(0)} + \Gamma_{sp,0}}{\Gamma_{sp}}\right)\right]_k \quad (2.78)$$

A few comments/observations:

1. With respect to eq. (2.78), which explores the region of low surface coverages of both polyamine and precipitates
 - Without precipitates ($\Gamma_{sp} = 0 \implies \Gamma_{b,MAX}/\Gamma_{TOT} = 1$), the marginal probability of non-attachment $\theta_i^{-sp,0}$ & $^{-PA(0)}$ tends to 1 for $\Gamma_{b,PA(0)}/\Gamma_{b,MAX}$ tending to zero (and hence $\alpha_{ij} \rightarrow 0$); this situation corresponds to a simultaneous absence of precipitates and an insufficient amount of free adsorbed polyamine (perhaps due to an excess of adsorbed COD_s^* or because of an excess of anionic polyelectrolyte covering the polyamine itself).
 - When the amount of precipitates starts growing, the first ratio in eq. (2.78) ($\Gamma_{b,MAX}/\Gamma_{TOT}$) becomes progressively less important, while the second (Γ_{sp}/Γ_{TOT}) begins to increase; as a result, the higher the adsorption densities of bare precipitates' surface sites and/or of surface sites occupied by free polyamine, the greater is α_{ij} , because interactions between patches of bare precipitates exposed on the aggregates surface are always assumed to favor attachment. Conversely, $\alpha_{ij} \rightarrow 0$ when the terms in the round brackets tend to zero (which could happen in case of an excess of polyelectrolyte and/or COD_s^*).
2. With respect to eq. (2.77), which explores the region of high polyamine coverages, it can be seen that precipitates have again a positive effect on α_{ij} , because the denominator in eq. (2.77) always exceeds the negative contribution given in this case by $\Gamma_{sp,PA(0)}$.
3. According to eq. (2.77) and eq. (2.78), an excess of polyelectrolyte seems to be detrimental only at low concentrations of adsorbed polyamine and precipitates (2.78), as one would actually expect.

It could be useful to look at the following limit cases:

- No precipitates present ($\Gamma_{TOT} \simeq \Gamma_{b,MAX}$)

$$\alpha_{ij} \simeq 1 - \left(\frac{\Gamma_{b,PA(0)}}{\Gamma_{b,MAX}} \right)_i^{\nu_i^*} \cdot \left(\frac{\Gamma_{b,PA(0)}}{\Gamma_{b,MAX}} \right)_j^{\nu_j^*} - \left(1 - \frac{\Gamma_{b,PA(0)}}{\Gamma_{b,MAX}} \right)_i^{\nu_i^*} \cdot \left(1 - \frac{\Gamma_{b,PA(0)}}{\Gamma_{b,MAX}} \right)_j^{\nu_j^*} \quad (2.79)$$

- Excess of precipitates ($\Gamma_{TOT} \simeq \Gamma_{sp}$)

$$\alpha_{ij} \simeq 1 - \left(\frac{\Gamma_{sp,PA(0)}}{\Gamma_{sp}} \right)_i^{\nu_i^*} \cdot \left(\frac{\Gamma_{sp,PA(0)}}{\Gamma_{sp}} \right)_j^{\nu_j^*} - \left(1 - \frac{(\Gamma_{sp,PA(0)} + \Gamma_{sp,0})}{\Gamma_{sp}} \right)_i^{\nu_i^*} \cdot \left(1 - \frac{(\Gamma_{sp,PA(0)} + \Gamma_{sp,0})}{\Gamma_{sp}} \right)_j^{\nu_j^*}, \quad (2.80)$$

and in this case, even without PA, $\alpha_{ij} \simeq 1$ as long as $\Gamma_{sp,0}/\Gamma_{sp} \simeq 1$ (a low collision efficiency would be probably due to an excess of adsorbed COD_s^* and/or polyelectrolyte F).

The proposed extension to Hogg's model, which allows to model the combined use of a low molecular weight polymeric coagulant, a metal coagulant and a high molecular weight polymeric flocculant, actually considers the first two as the main responsible for particle destabilization, whereas the third as an adjuvant of flocculation which is actually useful once particles have already been destabilized. It will be shown in the next section that the flocculant aid could also play an important role within the Discretized Population Balance Equation (DPBE) of the aggregates, since it positively affects floc strength by lowering the magnitude of the fragmentation term in the DPBE, allowing the median of the PSD to evolve towards larger floc sizes over time. Exactly as in Hogg's Model, the terms ν_i^* and ν_j^* in eq. (2.76) refer to the number of all available adsorption sites at the surface of the colliding aggregates i and j . In particular, considering all the particles that make up a k -sized aggregate, one can write:

$$\nu_k = \Gamma_{TOT_k} SSA_b N_{p_k} m_p, \quad (2.81)$$

where m_p is the mass of a bentonite particle and N_{p_k} is the number of primary (bentonite) particles within a k -sized aggregate. The size of adsorbates is implicitly taken into account by the λ coefficients in the adsorption equilibria. Considering that the collision efficiency should refer only to patches exposed on the particle/aggregate surface, ν_k should be re-proportioned in such a way to consider just the primary particles exposed on the aggregates surface, which are the only ones responsible for the attachment between aggregates. To estimate these external surface sites, the following

proportion between primary particles and aggregate can be written:

$$\frac{\nu_k}{N_{p_k} S_p} \simeq \frac{\nu_k^*}{4\pi(r_k \Omega_k^*)^2} \quad (2.82)$$

For which ν_k^* appearing in eq. (2.76) is equal to:

$$\nu_k^* = \frac{\Gamma_{TOT_k} S S A_b m_p}{a_p^2} (r_k \Omega_k^*)^2, \quad (2.83)$$

where a_p is the primary particles radius, r_k is the aggregate collision radius and $r_k \Omega_k^* < r_k$ is the aggregate hydrodynamic radius. Speaking of primary particle radius, it should be noted that because of surface precipitation, the primary particles radius cannot remain constant. Using for example eq. (2.74), the temporal variation of the average volume of primary particles in aggregate j due to precipitates is given by:

$$\left(\frac{dV_p}{dt} \right)_j = \frac{1}{N_{p_j}} \sum_i \frac{1}{\rho_{Me_i(OH)_x}} \left(\frac{dM_j}{dt} \right)_{Me_i(OH)_x}^{sp} \quad (2.84)$$

Thanks to this expression, it is possible to derive the rate of change of primary particles radius due to heterogeneous nucleation of precipitates:

$$\left(\frac{da_p}{dt} \right)_j = \left(\frac{\partial a_p}{\partial V_p} \right)_j \left(\frac{dV_p}{dt} \right)_j = \frac{1}{3} \left(\frac{3V_p}{4\pi} \right)_j^{-2/3} \left(\frac{dV_p}{dt} \right)_j, \quad (2.85)$$

and hence derive, for every aggregate size class, the temporal variation of all the above mentioned properties (collision radius, porosity, hydraulic permeability, collision frequency and collision efficiency functions, etc.) due to the amount of precipitated hydroxides onto the primary particles. In other words,

$$\frac{d\Omega^*}{dt} = \frac{\partial \Omega^*}{\partial a_p} \frac{da_p}{dt}, \quad \frac{d\eta^*}{dt} = \frac{\partial \eta^*}{\partial a_p} \frac{da_p}{dt}, \quad \frac{dk^*}{dt} = \frac{\partial k^*}{\partial a_p} \frac{da_p}{dt}, \quad \frac{d\alpha_{ij}}{dt} = \frac{\partial \alpha_{ij}}{\partial a_p} \frac{da_p}{dt}, \quad \frac{d\beta_{ij}}{dt} = \frac{\partial \beta_{ij}}{\partial a_p} \frac{da_p}{dt}.$$

Chapter 3

State of the Art of Online Monitoring and Control of Clariflocculation Processes

This chapter follows in the footsteps of an extensive review by Ratnaweera and Fettig (2015) [21], who summarized all the major studies conducted over the past 20-30 years on online monitoring and control of clariflocculation processes. The intent is not to go into the detail of each of the reported studies, but rather to highlight the main objectives and approaches proposed in order to have a general picture of what has already been done in terms of research, on which aspects it is perhaps suggested to carry out further investigations and which are the main criticalities of coagulation control in order to better understand where future research should go. What parameters can be measured and how to obtain these measurements are question that were somewhat answered in the previous section. For what purposes and how these measurements can be used are questions that this paragraph is intended to answer. In table 3.1 an overview of the online sensors available on the market and their application for monitoring the coagulation process in practice is provided. The assessment reflects the experiences of the authors, which may deviate to some extent from U.S. practices or even among European countries. If the quality of the incoming water/wastewater were to be constant and the optimal amount of coagulant were well determined, the dosage

3. State of the Art of Online Monitoring and Control of Clariflocculation Processes

Table 3.1: Online sensors applied for monitoring and control of clariflocculation processes according to experiences and applications in Northern Europe/Scandinavia (*Source:* from [21])

Monitoring stage	Parameter	Commonly Measured	Occasionally Measured	Commonly Used for Dosing Control	Occasionally Used for Dosing Control
Raw water (inlet)	Flow	X		X	
	pH		X		X
	Temperature	X			
	Conductivity		X		
	Turbidity		X		X
	Suspended solids		X		X
	UV-absorbance		X		X
	Color		X		X
	TOC		X		
	Ortho-P		X		
	Total-P		X		
	Particle counter		X		
Coagulated water (immediately after dosing of coagulants)	pH	X		X	
	Streaming current		X		X
	Zeta potential				X
Finished water (outlet)	pH	X			X
	Turbidity	X			X
	Suspended solids	X			
	UV-absorbance	X			
	Color	X			
	TOC		X		
	Ortho-P		X		
Total-P		X			

would simply be proportional to the flow rate. Thus, flow-proportional dosing should be (and usually is) applied by treatment plants as part of a feedforward control scheme. This requirement is largely fulfilled both in drinking and wastewater treatment [22, pp. 10–18]. Given, however, that water quality may be more or less variable depending on its origin, and that treatment objectives are closely related to its final use (discharge, recycle or reuse), the increasing availability over the years of progressively sophisticated analyzers and the increasingly pervasive use of programmable logic controllers (PLC) in engineering practice has prompted research to renew attention on the search for optimal strategies for monitoring and controlling coagulation-flocculation processes. The reason why it has not yet been possible to address this problem in the same way that has been done for half a century for electromechanical systems or chemical plants is related to the lack of a universally accepted mathematical description of coagulation

and of appropriate and reliable online physical sensors/analyzers specific for monitoring coagulation. Especially in industrial wastewater treatment, the problems of fouling and interference in sensor measurements are sometimes too great to solve unless substantial modifications are made to the treatment process itself.

With respect to the influence of water quality, Jackson and Tomlinson (1986) have described the following general possibilities for using online parameters directly:

- Feed-forward control based on raw water quality;
- Feed-backward control based on dosed water quality;
- Feed-backward control based on treated water quality.

More advanced configurations are obviously possible (and sometimes suggested) depending on the challenges encountered. A successful feedforward strategy relies entirely on forecasting: given a model of the system we want to control, the controller is simply designed to make the system work where we want to according to a predetermined set point, without actually checking how the system is performing. Ideally, if carefully designed, a feedforward control can completely eliminate the effect of measured disturbances on the process output, sometimes providing better responses than feedback control alone. However, the main issues with feedforward control are that it relies on good model(s) of the system, it can't stabilize unstable dynamics, it generally tends to over-react to noise and disturbances and it is not robust to uncertainties (disturbances that we cannot measure or are out of our control, time-changing parameters, etc.). The best strategy would then be to measure as much as possible about how the system is doing and use those measurements as a feedback signal to update the controller as to whether or not it is actually doing the job we want it to do, and improve its action in case it is not. In this regard, it should be noted that there are indeed systems (as in this case) for which it is not possible to obtain those measurements, either for the lack of specific physical sensors/analyzers or for practical reasons (e.g. suboptimal sampling points, noisy or unreliable measures due to interference or fouling - which may do more harm than good to control performance -, etc.). Nevertheless, the main advantages of feedback control are that it can actually stabilize unstable dynamics, it can attenuate noise and reject disturbances, and it can compensate for uncertainties in the model (even with relatively poor models, it certainly

does a better job than feedforward control alone with respect to noisy measurements, model uncertainties and disturbances, even for complex and high dimensional nonlinear dynamics). The only real (but very important) weakness of feedback control is related to time delays (between measurement and actuation), which usually "kill" robust performance and make the system potentially unstable.

Combining feedforward and feedback control can significantly improve performance over simple feedback control whenever there is a major disturbance that can be measured before it affects the process output. Ultimately, the decision as to whether or not to use feedforward control depends on whether the degree of improvement in the response to the measured disturbance justifies the added costs of implementation and maintenance. This is usually the case with coagulation control, as amply demonstrated by studies conducted on streaming current detectors or zeta potential analyzers, which have shown very poor performance when used alone (they require at least the flow rate of the incoming raw water as an additional measurement, although other parameters may also be recommended). On the other hand, feedforward control is usually used along with feedback control because the latter is able to track setpoint changes and to suppress unmeasured disturbances that are always present in any real process. Speaking of streaming current detectors, it was demonstrated - especially in river water treatment applications - that the instrument set point needs to be tuned regularly (sometimes even daily) because of not-negligible changes or fluctuations in raw water quality (mostly related to NOM). Another commonly used control scheme worth mentioning is the so-called *Cascade Control*, which consists of one or more nested control loops within the architecture of another. Usually cascade control is applied to processes which exhibits a sluggish behavior with respect to some process variables (e.g. temperature, equalized raw water quality, etc.) and a faster one with respect to others (e.g. flow rate, unequalized raw water quality, streaming current, etc.). Essentially, cascade control improves a "slow" control loop's ability to respond to known disturbances by exploiting the dynamics of a faster one, which provides an early warning variable that facilitates disturbance rejection in the outer (slower) loop, helping to keep its performance steady. The faster inner loop therefore reduces the overall variability experienced by the process since is able to respond more quickly to disturbances than the outer loop, and can also help to correct for nonlinearities. The outer loop can hence be tuned more conservatively: since the inner loop makes more rapid adjustments, the

outer loop no longer needs to be nimble. Both feedback control and cascade control usually implement PD and/or PID controllers, for which there is an extensive and rich literature in terms of design and tuning algorithms, provided that a mathematical model of the dynamical system to be controlled is given. PD and PID controllers have always been an industry standard thanks to the mathematical rigor with which it is possible to design and optimize the control of virtually any system according to the desired specifications, knowing at the same time the strengths and weaknesses, as well as the underlying uncertainties and approximations introduced. Although several approaches for the design and optimization of linear controllers applied to nonlinear dynamical systems have been proposed, it is generally accepted that they give poor performance when applied to highly nonlinear systems or to moderately nonlinear systems with *dead time*. Considering that the coagulation-flocculation process can be highly nonlinear and that dead times can reach truly unsustainable values for any type of feedback control - especially when it is based on treated water quality -, the need to develop and implement advanced models and control strategies is quite self-explanatory. One of the best strategies in this regard would be to use a *Model-Based Controller* (which indeed is a feedforward control) combined with a slow feedback controller that actually keeps track of the system response and slowly return that information as an input signal to the controller, in a somewhat supervisory way. According to this architecture, the feedback controller is designed to correct the system behavior slowly, on a wider timescale (in order to avoid magnifying potential instabilities caused by dead times), whereas the model-based controller is designed to provide quick responses to external disturbances, and therefore requires a good model of the system to do so. Unlike a standard feedforward controller, a model-based controller uses both the input measurements and the feedback signal from the slow feedback controller as informations to calculate the most appropriate control action according to some optimality criterion specified by the designer. Intelligent control systems can also make use of all these (stored) informations to further refine (or even recalibrate) the model whenever system performance becomes rather poor.

3.1 Direct Use of Online Parameters for Dosing Control

3.1.1 Feed-forward control based on raw water quality

To be successful, a feed-forward control strategy (fig. 3.1) requires an established and well agreed site-specific functional relationship between the raw water parameters (directly and continuously measured *in situ*) and the optimal coagulant dose required for effective destabilization. For this purpose, laboratory analysis, jar tests data, pilot-plant runs and/or technical-scale operation data must be available to cover ideally all of the possible parameter variations that might occur during the plant's lifetime. Finding such a relationship usually requires data spanning at least over a period of few months, preferably over a whole year ([21]). One of the first applications for this kind of strategy was done at Hazard's Green waterworks by Jackson and Tomlinson (1986), for the direct treatment of abstracted river water with iron sulfate. Several historical data covering about four years were evaluated and implemented using a multiple regression which led to an optimal dosing equation containing nine empirical constants:

$$FeSO_4(l)(mg/l) = c_0 + c_1T + c_2 \log(T) + c_3K + c_4 \log(K) + c_5\theta + c_6 \log(\theta) + c_7Q + c_8 \log(Q), \quad (3.1)$$

where T , K , θ and Q are the turbidity, conductivity, temperature and flow rate of the raw water, respectively. This relationship was successfully implemented at Hazard's Green waterworks however, at two other sites, this approach failed, meaning that such a relationship (or a similar correlation) could not be found. A similar approach was applied in Piteå waterworks in Sweden, where a simple empirical dosing equation was derived by Hernebring (1987) for the treatment of the raw water with aluminium sulfate (AlS). It was found that the required dose of AlS was primarily dependent on the NOM concentration measured via UV absorbance according to the following linear relationship:

$$AlS \text{ dose } (mg/l) = 1.23 \cdot A, \quad (3.2)$$

where A (m^{-1}) represents the UV absorption at 254 nm . This simple relationship was successfully implemented by the waterworks for dosing control purposes as long as pH was kept stable and that the NOM properties were not changing significantly.

For similar surface waters from Norway another empirical correlation was suggested by Eikebrokk and Fettig (1990), who took into account also pH, initial NOM (NOM_i) and residual NOM (NOM_r) concentrations when building their relationship:

$$NOM_r = (c_1 + c_2 \cdot pH + c_3 \cdot pH^2 + c_4 \cdot pH^3) \cdot (Al\ dose, mg/l)^{c_5} \cdot NOM_i^{c_6} \quad (3.3)$$

It was found that the six parameters of this equation (from c_1 to c_6) were quite different between the five types of water studied, suggesting that the nature of NOM plays a very important role with respect to coagulant demand, as already stated in the previous chapters. This equation was not applied in a dosing control scheme because of these uncertainties (it was also found that the resulting stoichiometric ratios $Al : TOC$ varied between 0.2 and 1 mg/mg). This experience pointed out the need for a further characterization of the properties of NOM, as evidenced by the increasing interest around ultraviolet absorbance techniques in water and wastewater treatment that followed. The specific ultraviolet absorbance (SUVA) concept was suggested and further developed by Edzwald (1993); SUVA values are calculated by dividing UV absorbance at 254 nm by the dissolved organic carbon (DOC) concentrations of the NOM. It was seen that if raw water has an average SUVA of 2 – 2.5 and greater, the coagulant dosage is controlled by NOM rather than turbidity. In a case study with five different waterworks in the United States, this concept was successfully applied by Edzwald and Kaminski (2009). For two types of raw water, the following simple equation was derived between the Al dose required and the UV absorbance:

$$Al\ Dose\ (mg/l) = c_1 + c_2 \cdot A, \quad (3.4)$$

where A is the UV absorption at 254 nm (given as cm^{-1}). However, the empirical parameters c_1 and c_2 were different for the two raw water types (with stoichiometric ratios $Al : TOC$ between 0.6 and 0.65 mg/mg for both sites), indicating a site-specific influence of the NOM characteristics on coagulant demand. This was in good agreement with results obtained by Archer and Singer (2006), who evaluated data from several hundred waterworks in the United States using alum as primary coagulant (they derived a mean and median stoichiometric ratios $Al : TOC$ of 0.6 and 0.7 mg/mg , respectively).

Based on an evaluation of jar test data for 21 raw water sources coagulated with

ferric salts and other 39 raw waters coagulated with alum, Edwards (1997) developed an empirical relationship for NOM coagulation where the final DOC was assumed consisting of two contributes: a non-adsorbable DOC (DOC_{na} , not removable by coagulation) and an adsorbable DOC remaining in water after coagulation. The non-adsorbable DOC was related to the influent DOC_i and its corresponding $SUVA_i$ according to the following relationship:

$$DOC_{na} = DOC_i \cdot (k_1 SUVA_i + k_2) \quad (3.5)$$

the final adsorbable DOC fraction ($DOC_{a,f}$) was assumed dependent on pH , DOC_i , DOC_{na} and coagulant dosage. Altogether, six parameters were determined for ferric salts and alum, respectively, by regression analysis. The main purpose of the approach was to estimate NOM removal in terms of DOC and to design more detailed and site-specific jar testing.

The above mentioned studies show that feed-forward dosing control can be based on online measurements of UV absorbance (to be correlated with DOC or TOC) in raw waters where NOM is the primary factor affecting coagulant demand. Several suppliers offer spectrophotometers capable to either measure UV absorbance at a single wavelength (typically 254 nm for drinking water applications) or scan a range of the UV-VIS spectra, depending on the peculiar characteristics of the water to be treated and the need for a better characterization of the organics involved. Some instruments are also equipped with turbidity compensation techniques. A more comprehensive empirical description was developed by Van Leeuwen *et al.* (2003 and 2009), where several data from jar tests and pilot plant studies regarding different Australian surface waters were used to derive correlations between raw water parameters (UV absorbance, color and DOC), removal efficiencies and dosage requirements for two different coagulants (for which pH changes due to hydrolysis reactions could also be predicted). The software mEnCo[®] (SA Water Centre for Water Management and Reuse, Mawson Lakes, SA 5095, Australia) has been used by Australian waterworks to adjust coagulant dosing from 2004 and later applied to predict coagulant doses in a modified jar test protocol by Murshed *et al.* (2014). Following the footsteps of the Australian experience, and although not implemented yet (into the practical operation of waterworks), the same approach was later applied and developed further by Wang *et al.* (2012) and Xie *et*

al. (2012) using data from 13 raw water types in China, where besides describing pH changes for four coagulants in more detail, an attempt has also be made to characterize NOM properties by physical and chemical fractionation techniques.

A different approach was described by Xianjun *et al.* (2013), who analyzed data from the Chinese Xining Water Treatment Plant No. 8. In their case, raw water turbidity (T) seemed to be the main factor affecting coagulant demand, and the following simple empirical equation was derived by regression analysis from more than 400 datasets covering one full year:

$$\text{Coagulant Dose (mg/l)} = c_1 + c_2T + c_3T^2 \quad (3.6)$$

It was claimed by the plant that the correlation is now applied for dosing control purposes, but further data were not provided.

A relatively new approach for feed-forward dosing control for the removal of dissolved organic matter is based on *fluorescence spectroscopy*, where organic substances are characterized by excitation-emission matrices that are analyzed by multi-way techniques, in particular the so-called *Parallel Factor Analysis* (PARAFAC). Sanchez *et al.* (2014) have demonstrated the suitability of this approach in evaluating the performance of two coagulants by evaluating data from a full-scale drinking water treatment plant. Shutova *et al.* (2014) identified a number of fluorescence parameters that were useful to characterize the properties of organic substances with respect to treatability. They derived the following empirical relationship between DOC removal and the ratio between three of the PARAFAC components (C1,C2 and C4, respectively):

$$\text{DOC removal(\%)} = 26.161(C1 : C2) + 3.171(C1 : C4) - 12.003 \quad (3.7)$$

According to their analysis, fluorescence spectroscopy may be a more robust method to monitor organic substances than using UV absorbance or the SUVA value. Ongoing studies will demonstrate whether this approach can be applied for online dosing control in the future.

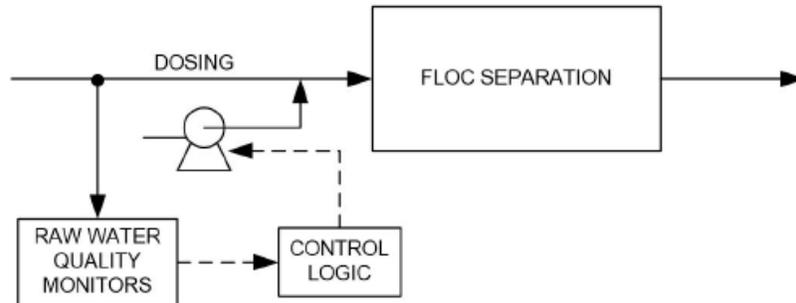


Figure 3.1: Sketch of a feed-forward control scheme based on raw water quality. (Source: from [21])

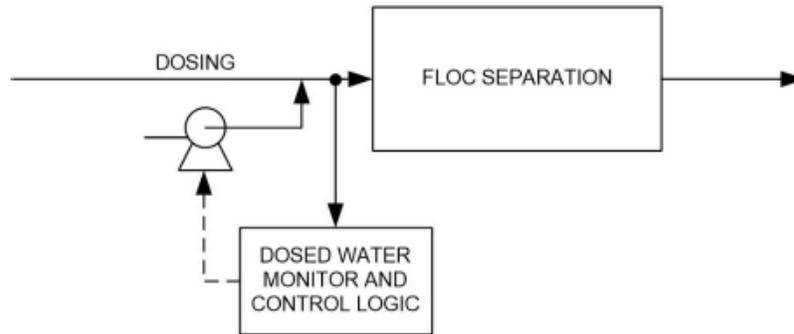


Figure 3.2: Sketch of a feed-backward control scheme based on dosed water quality. (Source: from [21])

3.1.2 Feed-backward control based on dosed water quality

There are two different ways to characterize dosed water quality: direct methods focus on measuring (by means of physical sensors) a few quantifiable properties of destabilized particles and freshly formed flocs, whereas indirect methods aim to evaluate the effects of coagulation on subsequent separation processes (e.g. sedimentation and filtration). The working principle of direct methods for feed-backward control is schematically depicted in fig. 3.2. Among the direct methods, the most common application is probably the *Streaming Current Detection* (SCD). The critical points of streaming current monitoring are mainly related to pH stability, rapid fluctuations in flow or water quality and poor mixing of chemicals, that can lead to unstable signals (as demonstrated by Critchley *et al.* in 1990). In a pilot plant study, Adgar *et al.* (2005) developed an advanced control strategy that accounts empirically for the relationship between

streaming current and pH. They concluded from their data that this strategy appeared to be less susceptible to disturbances. Yavich and de Wege (2013) compared the performance of SCD and the application of feed-forward control (mainly based on UV absorbance and turbidity) in full-scale treatment of the raw water from Lake Michigan, finding that the optimum SCD set point needed to be adjusted regularly (in some cases even daily), depending on the raw water conditions. They demonstrate that SCD may be limited when raw water turbidity changes rapidly or NOM concentrations increase. Surprisingly, the feed-forward control scheme developed for the waterworks (that was not operated automatically, but gave real-time recommendations to the operators) was instead able to predict proper dosages for all raw water conditions, suggesting that to improve both the effectiveness of the treatment and the reliability of the coagulant dosage control, SCD should be used as a feedback trim in a feed-forward control scheme.

Oh and Lee (2005) determined optimum coagulant dosages for the treatment of river water by pre-coagulation and microfiltration by just looking at the maximum removal efficiency and the minimum flux decline in the membrane stage; they concluded that coagulation was dominated by charge neutralization and that the streaming current signal monitored simultaneously could have been used to indicate whether the optimum dosage was applied or not. Nam *et al.* (2013) combined a rapid titration procedure for turbid waters with SCD measurements, with satisfactory results when poly-aluminium chloride (PAC) was used, but not so when using a mixture of PAC and a cationic polymer (poly-DADMAC). The authors claimed that their technique could be used to replace jar testing for the determination of the optimum coagulant dosage, even though further development of an online tool was not discussed. Byun *et al.* (2013) investigated the application of SCD for feed-backward control of polymer dosing for the conditioning of alum sludge from a waterworks in Korea. They showed with a pilot-plan study that there was a linear relationship between SCD readings and zeta potential measurements after dosage where the zero points were not identical, concluding that SCD could be successfully used for online dosing control. Streaming current detectors are available from a number of suppliers, though it should be noted that according to the *Manual of Water Supply Practices* M37, 3rd ed. of the American Water Works Association (AWWA, 2011), no standard calibration procedures have been developed so far. A proper set-point of the instrument has to be determined (and compared frequently to jar test results and/or plant performances) according to the application, water quality

and site characteristics (altitude, temperature, etc.).

Online sensors focused on measuring the size and shape of the aggregates formed in the flocculation stage are mainly applied in research, they have not been practically implemented in yet. A commercially available instrument is the so-called *Photometric Dispersion Analyzer* (PDA), which measures turbidity and turbidity fluctuations after coagulant addition via a fiber-optical monitor. The principle developed by Gregory and Nelson (1986) relies on the determination of a flocculation index (derived from the mean turbidity value and its fluctuation) to estimate optimum coagulation conditions. The applicability of this flocculation index has been proven in several jar test studies by Huang & Chen (1996), Staaks *et al.* and Zouloubis & Tzoupanos (2010) to compare different coagulants, optimize new coagulants or to evaluate the impact of turbulent heterogeneity during flocculation (as shown by Hopkins & Ducoste (2003)). Finally, Huang and Liu (1996) showed that PDA could also be used to adjust dosing to inflow turbidities varying between 20 and 2000 NTU, although there were some concerns about the technical applicability of the instrument to online dosing control. A somewhat similar approach was described by Cheng *et al.* (2008 and 2011), who evaluated turbidity data measured continuously 3 cm below the water surface during the slow mixing stage in jar tests. They observed a linear relationship between the square root of the average floc diameter and the standard deviation of the measured turbidity fluctuations. The technique, called *Nephelometric Turbidimeter Monitoring System* (NTMS), is supposed to determine optimum operational parameters for coagulation/flocculation, as demonstrated again by Cheng (2012) at a full-scale drinking water treatment plant in Taiwan (although a combination with an online control system has not been fully developed yet).

Sangu *et al.* (2012,2015) described an automatic coagulant dosage control method relying on the measurement of residual aluminium concentrations in the rapid mixing stage. Although this techniques may seem particularly suitable especially for rapid changes in raw water turbidity, it might be questioned, however, whether in this case coagulation pH was not the determining parameter that was indirectly measured via residual aluminium under the sweep floc regime ([21]).

With some limitations, particle size analysis can be used for floc characterization ([21]). Small-angle light scattering devices have shown some potential, even though

for low-density flocs the effective size might be underestimated, as shown by Gregory (2009). Nevertheless, there have been several applications of particle size analyzers that have contributed to a better understanding of the flocculation kinetics, like the one proposed by Lin *et al.* (2008) and Xiao *et al.* (2009). From the measured data, not only mean floc diameters but also their fractal dimensions could be derived.

Another promising technique to estimate floc properties *in situ* is the so-called *photoanalysis*, for which high-resolution images of the flocs are taken and fed to an image processing software. Jin (2005) investigated the effect of temperature and dosage on floc growth and floc size in the mixing phase of jar test experiments. Yu *et al.* (2012) determined the size and fractal dimension of aggregates during the settling period in batch tests and demonstrated the influence of dosage on these parameters. Lin *et al.* (2008) published data obtained by floc image analysis via wet scanning electron microscopy and confirmed that there is a distinct effect of dosage on floc size. Raspati *et al.* (2013) investigated the influence of two coagulants under different mixing conditions on the fractal dimension of flocs formed via in-line coagulation of NOM prior to microfiltration. They concluded that these floc parameters could be helpful in membrane fouling prediction.

Another property that could be exploited and was already investigated is floc strength (i.e. their resistance to external forces), which is somewhat closely related to their structure. As described in detail by Jarvis *et al.* (2005), there are several established techniques to determine floc strength. However, because of the costs of the equipment and the qualified manpower needed to operate it, an application of these advanced floc characterization methods for real-time dosing control would be difficult. The only sensing device available that is capable to detect the status of coagulated flocs on a real-time basis uses *laser light scattering*, and its output is used for dosing control (2014). Another concept using online techniques for image analysis has been reported in (2014).

In recent years, indirect methods for optimizing coagulation and flocculation based on the effectiveness of the floc separation stage have gained more interest. Tse *et al.* (2011) developed a flocculation residual turbidity analyzer (FReTA) that monitors the turbidity of floc suspensions undergoing quiescent settling and records the residual turbidity. The settling velocity distribution of flocs, characterized by mean value

and variance, could be determined. This technique could theoretically be used for feed-backward control, but has not been applied for that purpose so far. Blankert *et al* (2007) studied the combination of in-line coagulation and ultrafiltration with raw water from a Dutch river on a pilot scale and found a correlation between coagulant dosage and the initial resistance of the last dead-end filtration (before the chemical cleaning phase) that could be used for feed-backward control. This approach was later applied by Futselaar *et al* (2013) at technical scale resulting in up to 75% coagulant savings. Finally, Wei *et al.* (2013) and Manamperuma *et al.* (2013) have both reported successful full-scale applications of feed-backward control of coagulant dosing for drinking water and wastewater treatment plants based on raw water quality and coagulated pH measurements after dosing.

3.1.3 Feed-backward control based on treated water quality

Although final water quality is the primary criterion for optimum coagulation estimation, there are very few approaches that use finished water parameters for real-time control purposes. The main reason behind this lack of interest is related to the time lag between dosing and effluent coming from the separation stage, which is generally on the order of at least 30 min for direct filtration and one or few hours for a conventional settling stage. Considering the high variability and fluctuating behavior associated with raw water quality and influent flow rate, this time lag greatly degrades the performance of the control system, making it highly unstable and unreliable. In this respect, ballasted flocculation like the Actiflo[®] process (Veolia Water Technologies, Saint Maurice, France) or in-line coagulation followed by membrane filtration have clear advantages. In drinking water treatment, finished water quality is closely monitored as required by the respective guidelines. The data are normally used to adjust coagulant dosage manually by the operating staff based on empirical knowledge. Although software tools predicting both dosages and finished water quality based on feed-forward control techniques have been developed, automatic processing of finished water parameters for direct feed-backward control has not yet been realized ([21]). For raw water sources with quality variations observed only after weeks or months (like some Scandinavian lakes), feed-backward control based on finished water quality could be employed, although many of these plants do not actually really see the

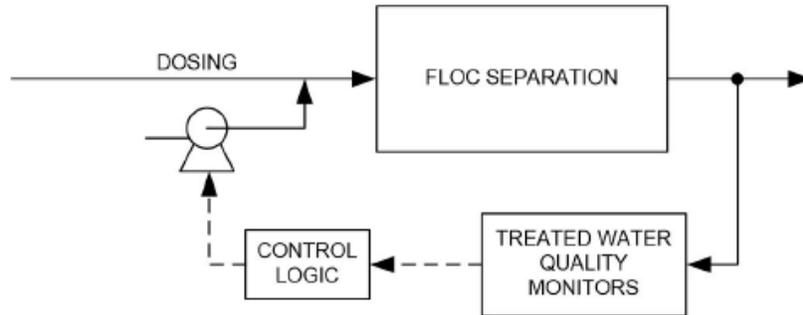


Figure 3.3: Sketch of a feed-backward control scheme based on finished water quality. (Source: from [21])

need for an online dosing control. Similarly, the direct use of finished water data for dosing control in wastewater treatment is not common either. Although many municipal treatment plants measure particles and ortho-P concentrations in the effluent, according to Storhaug, (2009) this information is not used for automatic dosing control purposes in a feed-backward control scheme. Another study on this regards that is worth mentioning is that of Tik and Vanrolleghem (2012), who proposed a control scheme for a full-scale chemically enhanced primary clarifier that uses a turbidity sensor at the effluent. While the coagulation/flocculation process is designed based on jar tests data, a layer approach is used to model the clarifier. According to their observations so far, jar tests tend to overestimate suspended solids removal, meaning that the control scheme generally has to be adapted accordingly.

3.2 Indirect Use of Online Parameters for Dosing Control

Multivariate Regression Analysis

AlGhazzawi and Lennox (2009) have given a comprehensive overview of the use of multivariate statistical process control methods like *Principal Component Analysis* (PCA) and *Partial Least Squares* (PLS) in the framework of *Model Predictive Control*, which is now an industry standard for many chemical plants and installations all over the world. They developed a monitoring tool that enables users to detect abnormal controller operation and resolve abnormalities when necessary, but they also

ponited out that process knowledge and experience are mandatory to analyze model predictive control or process abnormalities correctly. Franceschi *et al.* (2002), through an extensive analysis of jar test data obtained with synthetic raw water containing both particles and NOM coagulated with alum, accounted for eight water parameters, namely: type of particles, initial turbidity, initial UV absorbance, temperature, pH, conductivity, calcium concentration and sulfate concentration. They applied a two-level fractional factorial design method to determine the importance of each parameter on the minimum dosage required and the residual turbidity after settling. They found that there was an antagonistic influence of the different input parameters on the two responses studied (dosage minimization and residual turbidity), so it was not possible to simultaneously optimize the process with respect to both. When applied to data from a water treatment plant in France, the results were in good agreement except for conditions when the raw water had a low concentration of organic matter. An online dosing control tool was not developed based on this approach [21].

Juntunen *et al.* (2012, 2013) applied both multiple linear regression analysis and a *Multilayer Perceptron* feedforward neural network to laboratory and full-scale data from a lake water treatment plant in Finland. The input water quality parameters were permanganate index, color, pH, hardness, turbidity, conductivity and silicate concentration. Flow rate, coagulant dosage, coagulation pH and temperature were the most important process variables accounted for. The purpose of the model was to predict final turbidity and residual aluminium concentrations. The results showed that both models could be used for this purpose, although the goodness of the nonlinear model was slightly better than that of the linear one. However, it was concluded that the linear model was more suitable for applications requiring explicit correlations or fast calculations, as is the case with adaptive soft sensors ([21]). Ratnaweera *et al.* (2002, 2005) used technical-scale data from few mechanical-chemical wastewater treatment plants in Norway to derive a relationship between coagulant dosage, input parameters (flow rate, temperature, pH, conductivity, turbidity and orto-P), process parameters (settling time) and effluent parameters (turbidity) via partial least square regression analysis. The first results were promising, and an online control tool called DOSCON[®] (DOSCON AS, Oslo, Norway) was developed in subsequent studies. The concept was documented and proved after several years of full-scale operations that have enabled a reduction of coagulant consumption by over 30%, compared with flow proportional

dosing, as reported by Manamperuma *et al* (2013).

Annadurai *et al.* (2004) applied the response surface methodology to determine optimum operational conditions from jar tests using synthetic water containing both turbidity and NOM to simulate the coagulating behavior of stormwater. They used a quadratic polynomial equation to relate some input parameters (pH, UV absorbance, turbidity and alkalinity) and the process variable (the coagulant dosage) to residual turbidity and final UV absorbance, respectively, as responses. The coefficients were determined by multiple regression analysis and the proposed strategy was successfully applied to predict the required dosage to treat stormwater with extremely high turbidity (1650 NTU)[21]. In a similar approach, Ghafari *et al.* (2009) applied the response surface methodology in order to evaluate data from the coagulation of landfill leachate with PAC and alum, but using a different set of parameters. Only pH and coagulant dosage were defined as input data, while COD removal, turbidity removal, color removal and TSS removal were used as response parameters. The coefficient of the quadratic polynomial relationships were still determined via multiple regression analysis. A similar approach was later applied by Adlan *et al.* (2011) to leachate coagulation with iron chloride with floc separation by dissolved air flotation where pressure, flow rate and injection time (proportional to the gas-to-liquid ratio) in the flotation unit were used as additional input data. The method was able to identify the underlying meaning of the different parameters involved in the model and to optimize removal efficiencies. Finally, Guida *et al.* (2007) studied the coagulation behavior of real wastewater from five different municipal plants in Italy in a series of jar tests. They applied different statistical tools to the collected data in order to determine whether there were any correlations between the the parameters. While COD removal was seen to be significantly dependent on the origin of the wastewater, influent COD, influent TSS concentration, pH and alum dose, TSS removal seemed to be only affected by influent TSS concentration, pH and alum dosage. According to Ratnaweera and Fettig [21], these findings were not used further for online control purposes.

3.2.1 Artificial Neural Network Models (ANN)

A neural network is a powerful yet simple machine learning tool (and a nonlinear regression technique) that is capable of recognizing patterns among big data.

Neural networks are thought to be *universal approximators* and can be successfully implemented as a predictive tool provided a large set of historical data. Depending on the architecture of the model, the following types are distinguished:

- Multilayer Perceptrons (MLP);
- Time Delay Neural Networks (TDNN);
- Radial Basis Function (RBF).

A comprehensive overview of their applications in different fields has been given by Paliwal and Kumar (2009), who further emphasized the different prediction capabilities of ANN and statistical techniques when applied to dynamic processes. In recent years, ANN models have also been proven to be useful as software sensors for several water quality parameters in wastewater treatment plants (2007). Zhang & Stanley (1999) and Baxter *et al.* (1999, 2002) developed and validated a process control system based on an ANN model for NOM removal by enhanced coagulation in a waterworks in Edmonton, Canada. According to their results, this type of dosing control technique is feasible and it can help to improve cost savings in water treatment. Valentin & Denoëux (2001) and Bloch & Denoëux (2003) described in detail the building of MLP and RBF neural network models and applied the first one on a coagulation stage of a river water treatment plant in France. They collected data from jar tests covering a period of one year, and the training of the MLP was done using turbidity, conductivity, pH, temperature, dissolved oxygen and UV absorbance of the raw water as input variables and the optimum coagulant dosage as the output parameter. A total of 1600 datasets were available for building and validating the model. After its implementation for process control, a significant savings in coagulant usage was achieved.

In a comparative study of different modeling approaches, Yu *et al.* (2000) built each model using raw water pH, conductivity, turbidity and settled water turbidity from a waterworks in Taiwan as input variables and coagulant dosage as the output parameter. They found that their ANN model gave overall better predictions than regression or time series methods. Joo *et al.* (2000) pointed out the sensitivity of ANN models to redundancies on the training dataset, mostly arising from noisy and/or corrupted operating data accumulated over time. To improve the reliability of ANN models, the authors developed a data pre-processing tool to prepare both the training and the test

datasets before model calibration. Using jar tests data taken from Australian water sources, Maier *et al.* (2004) built and tested two different MLP neural network models. The first one used pH, turbidity, color, UV absorbance, DOC and alkalinity of the raw water in addition to the alum dose in order to predict residual turbidity, color and UV absorbance. The second model used instead the same water parameters, as well as the desired treated water quality (described by turbidity, color and UV absorbance) for the prediction of optimal dosage, final pH and residual aluminium concentration. It was concluded that further testing was required before the tools could be deployed for full-scale, real-time control purposes, probably because of the large number of process variables involved (which typically requires an even higher dataset dimensionality in order to build sufficiently robust and generalizable models). The approach of Maier *et al.* (2004) was later applied by Kriti & Smita (2013) to water treatment plant data in Bopal, India. The MLP neural network was built using raw water parameters (pH, turbidity, conductivity, TSS and dissolved oxygen) and settled water parameters (conductivity and turbidity) as model inputs, and the required alum dosage as output.

A MPL neural network model was also implemented by Zhang and Luo (2004) using full-scale operation data from a Chinese waterworks. They used raw water temperature, pH and turbidity as input parameters and coagulant dosage and the streaming current set point as model outputs. Thanks to their model, the treatment plant could achieve significant coagulant savings.

A similar approach was illustrated in Lamirini *et al.* (2005), who built a neural software sensor based on a large dataset covering four years of operation of a large drinking water treatment plant in Morocco. They used temperature, pH, conductivity, dissolved oxygen and turbidity (or TSS) as input parameters. Although their MLP neural network model was first intended to be applied as an online tool for dosing control, it was mostly used by the treatment plant as a support device, where the information generated by the madel was integrated into plant operation as an input to a diagnosis system (2006).

Another MLP neural network model was proposed by Wu and Lo (2008) using data from full-scale treatment of surface water in Taiwan. Temperature, pH, turbidity and color of both raw and treated water were fed into the models as input variables, whereas the dosage of PAC was treated as the model output. The authors compared the

MLP performance with that of an *Adaptive Neuro-Fuzzy Inference System* (ANFIS), which showed slightly better predictions than the former. Their ANFIS model was thus recommended to be applied by operators to assist dosing control. In a later study (2010), the same authors developed another method to build an improved ANN model that had better predictive capabilities. An ANFIS model was proven to be more accurate than an ANN model by Kumar *et al* (2013), who trained both models using jar test data and applied them for real-time dosing control in a pilot plant study.

3.2.2 Fuzzy Logic Models (FLM)

During the past 30 years, fuzzy controllers have found a wide range of applications in technical processes. Real-time data are used as an input for a multiparametric controller characterized by a set of logic (if-then) rules to generate an output signal aimed to keep the parameters as close as possible to a predefined set-point.

Liu and Wu (1997) conducted jar tests and continuous bench-scale experiments with synthetic raw water and ferric chloride as coagulant. They implemented a fuzzy logic feed-backward control system with streaming current (set point -0.05) and pH (set point of 8.0) as input, and coagulant and base dose as output parameters. They found that an artificially induced initial turbidity of 110 NTU could be steadily reduced to less than 10 NTU after coagulation and sedimentation. A full scale application of their approach has not been realized yet.

Ratnaweera *et al.* (1998) have demonstrated the capability of fuzzy logic modeling in water treatment to manage raw waters with time varying concentrations. A feed-forward fuzzy logic dosing system was successfully developed to treat raw water with color up to $225 \text{ mg} - \text{Pt}/\text{L}$ and turbidities up to 16 NTU to below $7 \text{ mg} - \text{Pt}/\text{L}$ and 0.1 NTU, respectively. No full-scale applications were reported after repeated tests on continuous pilot-scale demonstrations.

A feed-forward control system with a fuzzy feed-backward component was developed by Chen and Hou (2006) for a drinking water treatment plant in Taiwan. The model was based on full-scale data collected over a period of four years, with fuzzy control rules based on pH and turbidity of the settled water. The effectiveness of the proposed strategy was successfully demonstrated in several field tests.

Heddam *et al.* (2012) proposed a grid partition-based and subtractive clustering-based approach to build an adaptive neuro-fuzzy inference system (ANFIS). The main difference between the two approaches used was the determination of the optimum number and form of fuzzy rules, resulting in 729 rules for the grid partition-based model and only 15 rules for the subtractive clustering-based model. Temperature, pH, conductivity, turbidity, UV absorbance and dissolved oxygen concentration of the raw water were used as input parameters, whereas alum dosage was treated as the output parameter. The models were built, validated and tested using a total of 725 datasets collected at the Boudouaou drinking water treatment plant in Algeria. It was found that the subtractive clustering-based model outperformed the other due to its simplicity in parameter selection and a more accurate and reliable prediction of the optimal coagulant dosage [21].

Finally, an integral neuro-fuzzy process controller was developed in China by Wan *et al.* (2010) wastewater with poly-aluminium chloride. The model was fed with flow rate, raw water COD and coagulant dosage as input parameters, and effluent COD as the output parameter. Altogether, 100 datasets were available for model development and testing. The performance of the new (advanced) process controller was compared with a simple feed-backward controller based on effluent COD data, showing a saving of chemicals of almost 25% for the more advanced control scheme.

3.3 Conclusions and future developments

The lack of a universally accepted mathematical description for water and wastewater coagulation has prompted researchers over the years to gradually shift interest from process modeling to data-driven approaches and software sensors design. Although reliable mathematical models derived from first principles generally offer greater capabilities in terms of predictability, generalizability, control design and physical understanding of the process, their analytical representation, usefulness and/or implementation in the context of coagulation/flocculation control has proven to be too challenging a task for the following reasons:

- The inherent nonlinearity of the coagulation-flocculation process, which derives from the combination of multiple destabilization mechanisms acting

simultaneously according to the values assumed by the key process variables (and working conditions) and by the fact that the growth kinetics of flocs appearing in the population balance equation is a pseudo-second order reaction with varying coefficients.

- The computational effort required to implement such an accurate mathematical description in a process simulator.
- The time-varying characteristics of raw water, both in terms of flow rate and quality, and in terms of unpredicted and/or unmeasured substances that may affect treatment performance.
- The wide spectrum of primary coagulants and flocculant aids available on the market, each with secret formulas designed for specific applications according to the destabilization mechanism(s) exploited, raw water properties, working conditions and treatment objectives.

Even if such an accurate mathematical description for the coagulation/flocculation process could theoretically be derived, there is no guarantee that it could be applied for online control purposes because of well-known inherent issues related to physical sensors and online analyzers. The physical-mathematical model of the process could in fact involve parameters that are somewhat impossible to measure, at least directly; on the other hand, even if such parameters could theoretically be measured (or at least indirectly estimated), it may not be technically feasible to install the corresponding physical sensors/analyzers where needed (or where it is "best") due to fouling and/or interference problems, logistic challenges and/or practical issues, or again because of incompatibilities between the response time of the instruments and the characteristic timescales of the process (which ultimately account for the temporal fluctuations in both water/wastewater quality and flow rate). The growing interest in statistical analysis and machine learning techniques applied to water and wastewater treatment (in general) thus stems from the fact that there are processes - such as coagulation - that are perhaps too complex to be modeled from first principles, or for which there is no real interest in doing so considering all the above concerns (it might not be worth it, at least from a control standpoint). An accurate physical model, when properly calibrated, could however be of extreme importance for the design and management of a process;

it could also be useful for data preprocessing or to produce dummy data to further enhance software sensor training.

In order to derive empirical and easily implementable relationships between key process parameters (explanatory variables) and control/response variables (mainly coagulant dosage or residual turbidity and/or residual UV absorbance), the development of software sensors first drew attention to well-known and consolidated statistical methodologies and subsequently to machine learning techniques. The not so great success of multivariate and principal component analysis applied to coagulation control may be justified by the fact that even though it is technically possible to include nonlinear terms (in the form of products, ratios and/or powers of multiple explanatory variables) as additional explanatory variables to improve the match of the model with experimental data, the least square regression behind both the techniques is still linear, can be prone to overfitting and is highly sensitive to outliers. Most of these issues could be attenuated by using *sparse regression* techniques like LASSO or *Elastic Net Regularization*, which are the new standard for model discovery algorithms like SINDy or *Robust Principal Component Analysis* (RPCA) applied to Dynamic Mode Decomposition for example, but studies conducted using these methodologies have not yet been developed.

The growing interest in artificial neural networks over the past decades stems from the fact that they have been mathematically proved to be universal approximators, and are relatively easy to implement and train. ANNs can be used for data-driven control, model discovery, future prediction and system diagnostics, or even for model calibration. However, the main issues of neural networks are their "black box" nature and their high sensitivity to corrupted data, which require the utmost attention to data preprocessing especially when ANNs are fed with online full-scale measurements. Because the training and validation steps of an ANN is highly data demanding, there have always been concerns about research costs, which often discouraged investors from considering chemical savings and process reliability as aspects worth investigating. Although jar tests data are very useful and easily obtainable (though very time consuming), an ANN model would require spending at least a few months on a pilot test equipped with online analyzers before being successfully implemented on a large scale, which usually requires a huge amount of money and technical expertise.

To overcome the semantic limitations of ANNs due to their black box nature (i.e. even if an ANN works, it is not known exactly how it works or where to look at when it is necessary to slightly modify the model due to minor changes in plant operation, coagulant type or plant layout for example, without the need to entirely rebuilt it from scratch), some hybrid architectures combining the learning capabilities of neural networks with the semantic power of fuzzy logic (like ANFIS) were proposed, and have proved equally valid - sometimes even better than ANNs - in terms of performance. Fuzzy logic has also been successfully applied for dosage control and its strength lies in its underlying simplicity and ease of communication even to non-expert operators; from a mathematical point of view, the intrinsic nonlinearity of fuzzy models makes them very desirable and versatile in terms of descriptive capacity, even if their biggest flaw could be precisely the high arbitrariness with which the parameters of the model (number of fuzzy rules and form of membership functions) can be chosen. Sometimes, in order to approximate highly nonlinear processes, fuzzy logic models may require a very large number of tuning parameters (which is greater than or equal to the number of fuzzy rules multiplied by the number of parameters describing the shape of the membership functions, which is usually 2 or 3). Deriving such a number of unknowns requires having a lot of data and fast but reliable optimization algorithms. Despite appearances, optimization and tuning of a fuzzy logic controller require a lot of experience.

In recent years, thanks to the development of sparsity promoting algorithms, a new era of machine learning has emerged, with very promising results and a renewed sense of scientific purpose. The possibility of integrating the concept of sparsity to famous and reliable algorithms such as the Singular Value Decomposition (SVD) or the Fast Fourier Transform (FFT) has opened the door to a whole new field of fast and reliable data-driven techniques for model discovery like the Dynamic Mode Decomposition (DMD, 3.4) or the Sparse Identification of Nonlinear Dynamics (SINDy, 3.5). These modeling techniques, unlike those mentioned so far, are entirely data-driven (meaning that there is little to no arbitrariness in terms of modeling choices), extremely easy to implement (sometimes just a few lines of code) and have a tremendous semantic power in terms of physical meaning of the terms appearing in the discovered model. Since there has always been an interest in deriving a dynamic model for coagulation anyway, it might be interesting to apply such data-driven techniques in this field. Indeed, it should be noted that to design a control in an optimal way it is always better to have

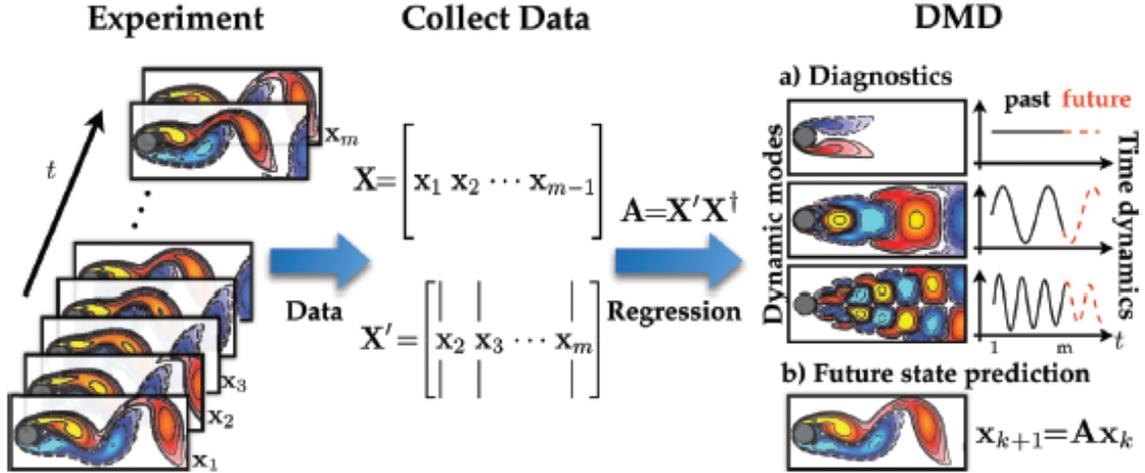


Figure 3.4: Overview of DMD illustrated on the fluid flow past a circular cylinder at Reynolds number 100. 1) Data (images in this case) are collected over time within the entire spatial domain and stacked into extremely tall column vectors x_i . 2) A *matrix of snapshots* X is built by placing the different instantaneous column vectors x_i side by side. 3) A second matrix of snapshots X' is built by moving each column vector one step forward in time. 4) The best fit linear operator \hat{A} describing the temporal evolution of the state $X' = AX$ is derived by performing a Singular Value Decomposition of the pseudo-inverse $A = X'X^\dagger$ without actually computing it. 5) Eigenvalues and eigenvectors of \hat{A} extracted this way represent the dominant modes and coherent structures of the discovered dynamical system (in terms of energy and frequency content) (*Source: from [23], Reproduced from: [24]*)

a dynamic model of the system to be controlled.

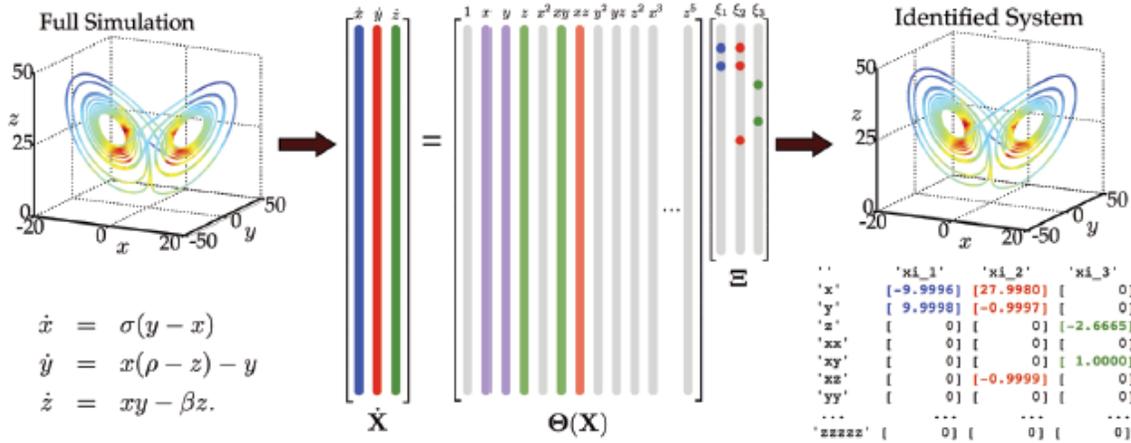


Figure 3.5: Schematic of the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm applied to the Lorenz attractor. Parsimonious models (i.e. sparse coefficients of the Ξ matrix) are extracted using a library $\Theta(X)$ of possible candidate nonlinear functions using sparse regression (e.g. LASSO). The $\Theta(X)$ library may be constructed purely from measurement data. (Source: from [23], Reproduced from: [25])

Chapter 4

Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

4.1 Site Inspection and First Observations

Although the wastewater and discharges managed by the client generally differ from plant to plant, the chemical-physical treatment units all work almost in the same way, namely:

1. Oil/grit separation stage (not always present), with possible addition of a demulsifier agent to further improve oil-water separation (if necessary);
2. Eventual pH correction (found only in very few cases), followed by coagulant addition (generally Ferric Chloride or a $FeCl_3$ -based product);
3. Bentonite addition whenever necessary (before or after coagulant addition - not all plant managers agree on this point);
4. Base addition (generally lime), to bring about metal precipitation;
5. Flocculant addition (typically an anionic polyelectrolyte);
6. Floc sedimentation on circular or lamella clarifiers;

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

In some instances (such as in the Mirafiori plant), bentonite is added after both ferric chloride and lime. The use of ferric chloride (or other coagulants containing a high percentage of ferric chloride) stems from the need on one hand to make the treatment more resilient to pH fluctuations (which are ordinary and to which alum or aluminium-based products are more sensitive), and on the other to improve floc settleability without producing too much sludge due to extensive hydrolysis (the molecular weight of iron is almost twice that of aluminium). Bentonite is added in suspended form on one hand to adsorb dissolved organic substances (mostly nonionic surfactants and emulsified oil droplets eventually present), on the other to further improve floc settleability, which might potentially be lightened by the extensive adsorption of oily substances eventually passed undisturbed from the oil-water separation stage.

After a careful morphological and management analysis of the pilot plants selected by the customer, as well as of the scarce technical documentation made available, the following criticalities and observations have emerged:

- Poor communication between the producer and the manager of the discharges, which requires the utmost attention to equalization and storage systems to mitigate the frequency and extent of extraordinary events.
- Qualitative and quantitative discontinuity of wastewater, which significantly impacts both the treatment efficiencies and operating costs; the automatic dosage of reagents should ideally be calibrated according to both the flow rate and the quality of wastewater, and not only on the former (an aspect that, to date, has always required experience and human intervention, especially in the field of industrial wastewater treatment).
- Time consuming and unsatisfactory search for self-monitoring data and registers, often not (fully) computerized or with unusable and/or deficient content for the purposes of this study.
- Careful identification and sorting of all possible parameters and corresponding online sensors/analyzers somewhat related to the coagulation-flocculation process available on the market, with problems deriving from multiple reasons, namely:
 - Poor measurement reliability due to fouling and interference phenomena,

with particular regard to the incoming and coagulated wastewater;

- Considerable dead time between the beginning of flocculation and the clarified water, which makes it difficult to successfully implement any sort of feedback control based on treated water quality (furthermore, the process dynamics is unknown, as is the aforementioned time delay, which depends on flow rate);
- Poor availability on the market of specific instrumentation for monitoring and control of the coagulation stage (e.g. streaming current sensors), which jeopardizes the possibility of a punctual dosage of reagents; the reliability of such analyzers can also be severely compromised by the known complexity of industrial wastewaters.
- To be useful for control purposes, both the sampling frequency and the response time of analyzers and sensors should be somewhat compatible with the qualitative and quantitative variability of the raw wastewater (e.g. TOC-meters and COD-meters may not be recommended); this design constraint is fortunately mitigated by the equalization tank(s) upstream the physical-chemical treatment units.

4.2 Jar tests and chemical analyses: materials and methods

The laboratory activity, which began very late - and lasted just as briefly - due to the COVID pandemic, consisted in a jar test campaign conducted on the concentrated wastewaters of the so-called "IDAC" unit of the wastewater treatment plant of Mirafiori (Turin, Italy). IDAC is the Italian acronym for "Impianto di Depurazione Acque di Carrozzeria" and is essentially a treatment unit for the washing waters coming from the FCA automotive painting plant of Mirafiori. These wastewaters are characterized by a high content of heavy metals (iron and zinc in particular, with traces of aluminium and lead) and surfactants (mainly nonionic), and are more or less rich in settleable and suspended solids depending on the degree of supersaturation of heavy metals. During the jar test campaign, the following data have been acquired:

- pH, conductivity and turbidity measurements, using portable analyzers;
- COD, anionic and nonionic surfactants, metals (iron, zinc, aluminium, lead), nitrogen (ammonium, nitrate, nitrite) and phenols concentrations using LCK[®] cuvette test systems (Hach), before and after treatment;
- Solid content (TS, TSS, TDS) of the raw water samples and total solids (TS) of the dehydrated sludges, using the Standard Procedure (0.45 μm membrane filters, a precision scale and a lab oven);
- UV absorbance spectra between 190 and 280 nm of samples before and after treatment using a DR6000 UV/VIS laboratory spectrophotometer with RFID technology (Hach) and high precision quartz cuvettes for quantitative analyses (chamber volume of 3.5 ml and optical path length of 10 mm).

4.2.1 Materials and Chemicals used

- Cogagulant: Drewo8016[®] (a mixture that, according to its Safety Sheet, is composed by more than 50% of ferric chloride and by less than 5% of cationic polyamines); given its high effectiveness, it was always used a dilution ratio of 1:50 with demineralized water in order to have a better dosing control.
- Flocculant aid: Dreflo975[®], a high molecular weight anionic polyelectrolyte that was prepared with a concentration equal to 0.5 mg/L because of its considerable density and viscosity.
- Highly concentrated lime solution (of unknown concentration);
- HCl (0.01 M) prepared from a 33% concentrated bottle of HCl.
- Bentonite suspension, prepared with an average total solids concentration of 17 g/L.
- A 4 slots jar test apparatus.
- Beakers (1L, 2L), 2 adjustable micropipettes (1-5 ml and 0-1 ml);
- Portable conducimeter, pH-meter and turbidimeter.

4.2. Jar tests and chemical analyses: materials and methods



Figure 4.1: Jar test apparatus, portable turbidimeter, beakers and magnetic stirrer (upper left), portable turbidimeter and pH-meter and quartz cuvettes (upper right), desiccator and membrane filters (lower left), cuvet kit tests for chemical analyses, adjustable micropipette, vacuum pump glass chamber and spectrophotometer DR6000 (lower right)

- A Lab oven, a desiccator, membrane filters and a vacuum pump for the gravimetric estimation of solids.
- 2 magnetic stirrers.
- A DR6000 UV-VIS lab spectrophotometer (with a lower detection limit of 190 nm).
- 4 High precision quartz cuvettes for quantitative analyses with chamber volume of 3.5 ml and optical path length of 10 mm.
- LCK[®] cuvette tests for chemical analyses (ammonium, nitrite, nitrate, COD (high and low range), anionic and nonionic surfactants, aluminium, lead, total iron, zinc, phenols).
- A COD digester.



Figure 4.2: 1:50 dilution ratio of Drewo8016[®] with demineralized water (left) and freshly prepared bentonite suspension (right)

4.2.2 Sample Preparation for UV/VIS Spectroscopy

As can be seen from fig. 4.4 and fig. 4.5, after a few preliminary tests it was clear that demineralized water was not a suitable solvent for UV absorbance analyses of wastewater samples. First of all, since the ultimate purpose of this segment of research was to find a quantitative correlation between dissolved COD and UV absorbance, it can be seen from fig. 4.4 that using demineralized water as a solvent for quantitative analyses can lead to spectral UV absorbances frequently greater than 1. According to Beer-Lambert law,

$$T = \frac{I_t}{I_0} = e^{-k_\lambda l}; \quad A = -\ln(T) = \varepsilon_\lambda l M, \quad (4.1)$$

where I_t/I_0 is the ratio between the intensity of transmitted and incident beam, respectively, k_λ is the attenuation coefficient (which depends on the crossed medium properties and on the wavelength λ of the incident ray, T is the transmittance, A is the absorbance, ε_λ is the molar attenuation coefficient (or absorptivity) of the species in solution, M is its molarity (i.e. the molar concentration of the attenuating species of interest) and l is the optical path length; hence, for absorbances greater than 1-1.5 (i.e. for transmittances lower than 22-36%) measurement uncertainties due to the

4.2. Jar tests and chemical analyses: materials and methods

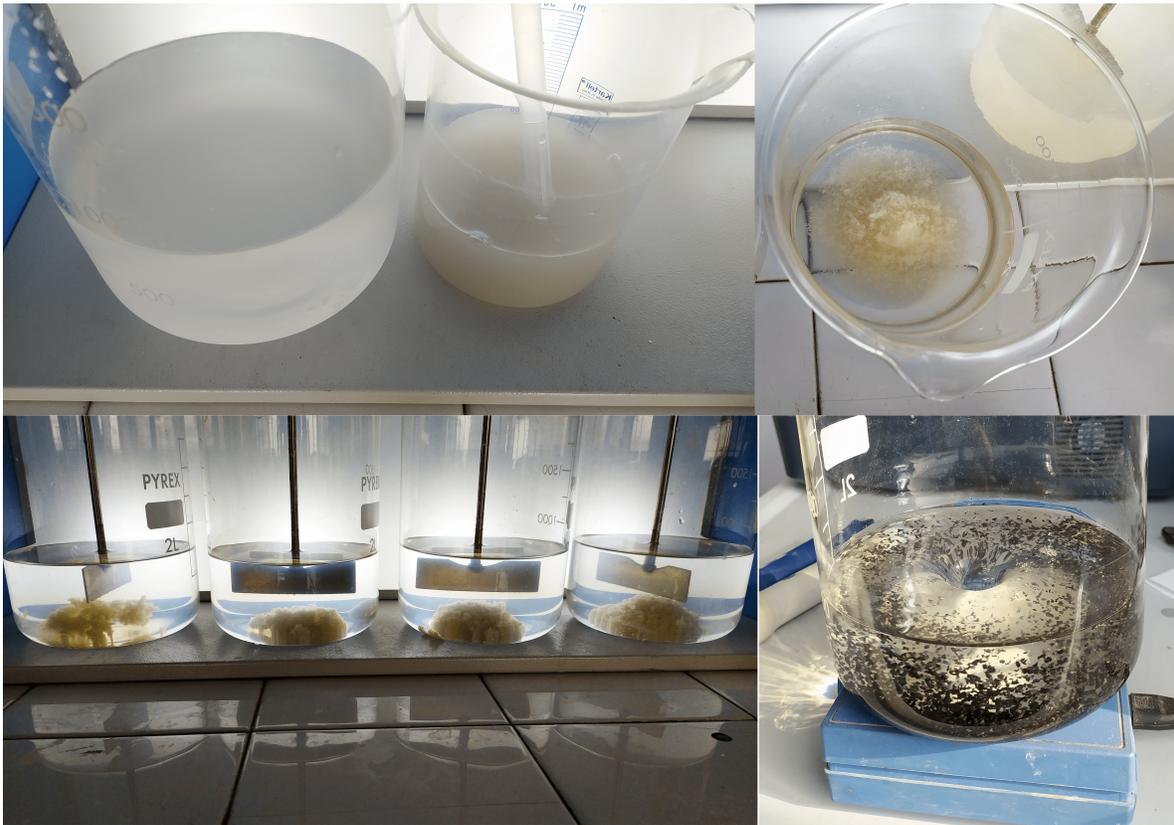


Figure 4.3: Sample before treatment (settled and unsettled)(upper left), treatment with bentonite and a slight excess of Drewo8016® (upper right), settled flocs after Drefo975® addition using slow mixing for 15-20 minutes (lower left), treatment of a clarified sample with Granular Activated Carbon (GAC) (lower right).

sensitivity of the spectrophotometer become more and more relevant, thus affecting accuracy. In addition, it is well known that the optimal absorbance range within which Lambert-Beer law is generally quite accurate is $0.1 < A < 1$; for absorbance values near the edges of this interval, slight deviations from linearity (between the spectral absorbance and the solute molarity) may be observed. Reducing the absorbance of the sample to fall within the range of validity of Beer-Lambert law would necessarily imply a greater dilution ratio, which could therefore become too large to give reliable results. Another issue related to the use of water as a solvent for quantitative spectrophotometric analyses of organic substances is that it tends to homogenize and dampen the absorbance peaks, which instead are highlighted when apolar solvents (like hexane or heptane for example) are used. As a result, absorbance spectra are often not clearly distinguishable (poor resolution). To solve this problem, an attempt was done using hydrochloric acid (0.01 M) instead of demineralized water, with the same dilution ratio (1:10). As can be seen in fig. 4.5, the absorbance spectra of the same sample appeared more distinguishable, with an absorbance peak below 1 for the settled raw water sample (compared to 1.3 using demineralized water). This behavior could be explained by the protonation of specific functional groups on the dissolved organic substances, which are slightly delocalized by the acid that makes them respond more to certain wavelengths (unlike with demineralized water, the peak wavelength turned out to be greater than the lower limit of detection of the spectrophotometer, i.e. 190 nm). The molarity of the HCl solution (0.01 M) and the dilution ratio of 1:10 were appropriately selected after a careful trial and error procedure carried out on several representative samples in order to find the better compromise between the following (often conflicting) needs:

1. Use the same dilution ratio and acid strength for all the samples along the same treatment chain;
2. Try to fall as much as possible within the range of validity of Beer-Lambert law anywhere along the treatment chain (paying attention to the fact that absorbances are generally high at the inlet and low at the outlet);
3. Use enough sample for the analyses (dilution ratios not greater than 1:10);
4. Try to save on acid consumptions.

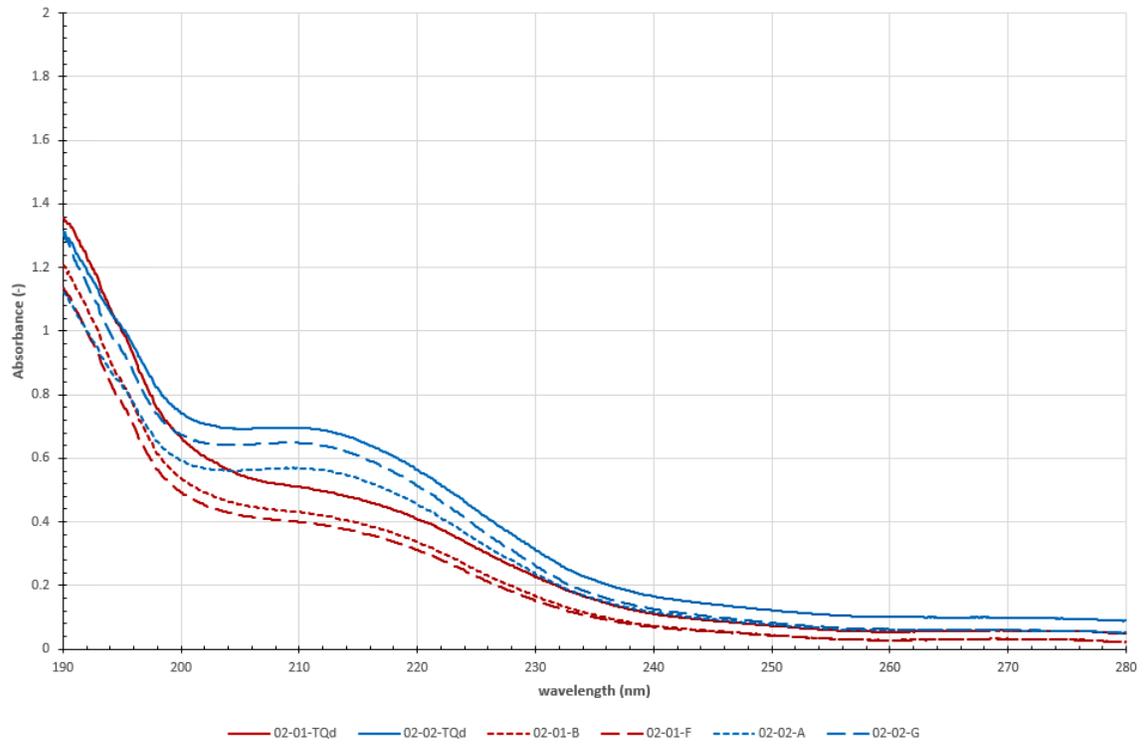


Figure 4.4: absorbance spectra of a settled, raw wastewater sample before (TQd) and after several treatment options. 1:10 dilution ratios with demineralized water

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

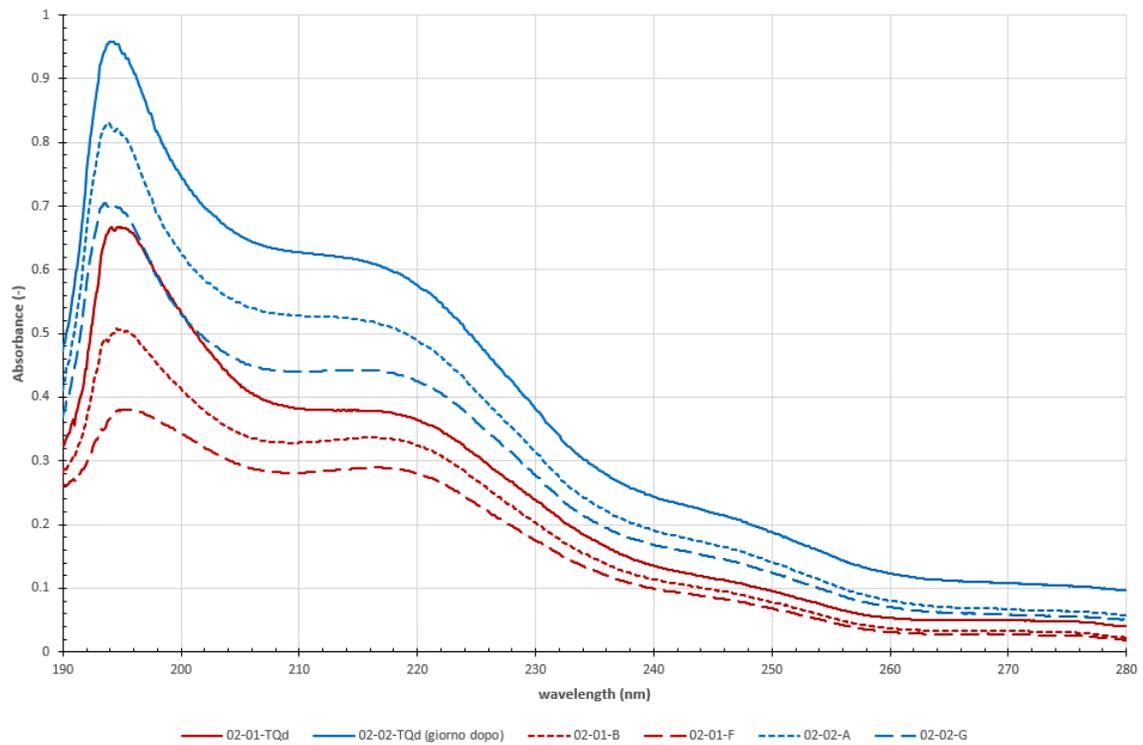


Figure 4.5: absorbance spectra of the same settled, raw wastewater sample before (TQd) and after several treatment options. 1:10 dilution ratios with HCl 0.01 M

4.2.3 The jar test procedure

Jar tests were conducted on raw wastewater samples left to settle for about 45 minutes - 1 hour, to simulate the effect of a primary settling stage (which does not actually exist in the IDAC treatment unit, but it is a design choice that will be justified later in more detail). Indeed, the secondary purpose of these laboratory tests was to eventually improve (if possible) the treatment currently performed in the IDAC unit, which consists in a two-stage flocculation with an intermediate flotation between the first and the second stage, followed by a final settling in a lamella clarifier. Bentonite is currently added only in the second flocculation stage, because otherwise it would negatively affect the performance of the dissolved air flotator. Dosing of chemicals is currently performed as follows:

- First flocculation stage: 120 ppm of $FeCl_3$ and 4 ppm of Dreflo975[®], with a pH value after lime addition of 9.5-10.0;
- Second flocculation stage: 80 ppm of $FeCl_3$ and 3 ppm of Dreflo975[®], with a precipitation pH of 8.2-8.6 after lime addition. The suspended bentonite added in this stage has an average concentration of 18 g/l.

Jar test experiments were structured as follows:

1. Definition of the daily relationship between the added (freshly prepared) bentonite suspension and turbidity of the settled sample added with bentonite, using a linear regression with 5 data points (each recorded as the average of multiple turbidity measurements).
2. Inversion of the above relationship to define the amount of bentonite to add in each jar in order to bring the turbidity of each sample to the desired value.
3. Bentonite addition in each jar but one, mixed for about 10 minutes at 45-60 rpm.
4. Coagulant addition (Drewo8016[®]), mixed at 45-60 rpm for about 5 minutes, without pH correction.
5. Lime addition, to rise pH to a value between 8.2 and 8.6 to bring about metal precipitation (primarily zinc, but also iron(II and III)).
6. Flocculant addition (Dreflo975[®]), mixed for about 15 minutes with gradually

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

reducing speed (from 30 rpm to 10 rpm).

7. Storage on a USB key of the UV absorbance spectra between 190 and 280 nm (0.1 nm resolution) of the settled raw samples and of all treated samples using a 1:10 dilution ratio with HCl 0.01 M as solvent.
8. Use of the UV absorbance spectra (plotted on an Excel sheet) to define the optimal bentonite dosage (i.e. the one for which the yield in the abatement of absorbance was considered as no longer economically sustainable).
9. Chemical analyses performed with LCK[®] cuvette tests and the DR6000 UV/VIS spectrophotometer. Given the limited budget and the different amount of boxes available for each parameter, the following analyses were performed:
 - Raw sample: COD, iron, lead, aluminium, zinc;
 - Settled sample: COD, iron, lead, aluminium, zinc, anionic and nonionic surfactants, phenols, N-NH₄, N-NO₂, N-NO₃.
 - Treated sample (no bentonite, henceforth called as "NB" sample): COD, iron, zinc, anionic and nonionic surfactants.
 - Treated sample ("optimal" dosage of bentonite, henceforth called as "OB" sample): COD, iron, lead, aluminium, zinc, anionic and nonionic surfactants, phenols.
10. Weighing of the dried sludges (TS) of the two samples treated with (OB) and without bentonite (NB).
11. Estimation of the solid content (TS, TSS, TDS) in the raw water sample and bentonite suspension (only TS and TSS).
12. Batch treatment of OB and NB samples with Granular Activated Carbon (GAC) for 10 minutes at 90 rpm, to obtain two additional UV absorbance spectra to be further correlated with the corresponding COD measurement.

The choice of adding bentonite prior to the coagulant stems from two reasons: on one hand it allows a better adsorption of dissolved organic substances (longer contact time), also providing a greater surface area for the heterogeneous nucleation of metal precipitates (first adsorbed as hydrated cations and then converted into metal

hydroxides after lime addition); on the other, the increase in particle concentration provided by the suspended bentonite allows a slight saving of coagulant whenever turbidity is too low by increasing the flocculation rate, which could potentially be reduced after primary settling (lower collision frequencies). It could be argued that lime could have also been added before the coagulant for example, because there may be instances where the dissolved metal content is so high that coagulation may occur spontaneously without coagulant addition through a sweep floc action by the freshly precipitated hydroxides. However, considering that any $FeCl_3$ -based coagulant is strongly acidic, this choice of reagent sequencing would have required a second pH control loop after coagulant addition, in order to maintain the precipitation pH between 8.2 and 8.6 (where zinc is less soluble). Since it is not possible to predict whether the dissolved metal content is sufficient or not to effectively destabilize the suspension via a sweep floc action by the corresponding precipitates, and that process reliability always has priority over cost savings (the complete precipitation of zinc must be ensured), if one wants to carry out the flocculation-precipitation process in single-stage without the need for a second pH control loop (with further addition of lime), this is perhaps the safest (and perhaps best) way to do it. Otherwise, although there may be instances in which some coagulant could be saved by avoiding primary settling, there may be no guarantee that the concentration limits of zinc (and perhaps iron) in the effluent would be respected after a single precipitation step.

To get a better estimate of the concentration of bentonite introduced into each wastewater sample, the corresponding TS and TSS content was measured daily instead of using the dry weight of the bentonite powder used to prepare the suspension. In this way it was possible to minimize errors related to evaporation or poor dissolution (with formation of lumps) of the bentonite powder.

4.3 Results and Discussion

In fig. 4.6 is reported the solid content of 10 raw wastewater samples expressed as a percentage of the relative total solids (TS), whereas fig. 4.7 shows a comparison between 11 wastewater samples in terms of concentrations of the main pollutants of interest, measured before and after primary settling (i.e. COD, iron, zinc, lead and aluminium). To give a better overall picture, each concentration has been dimensionless with respect

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

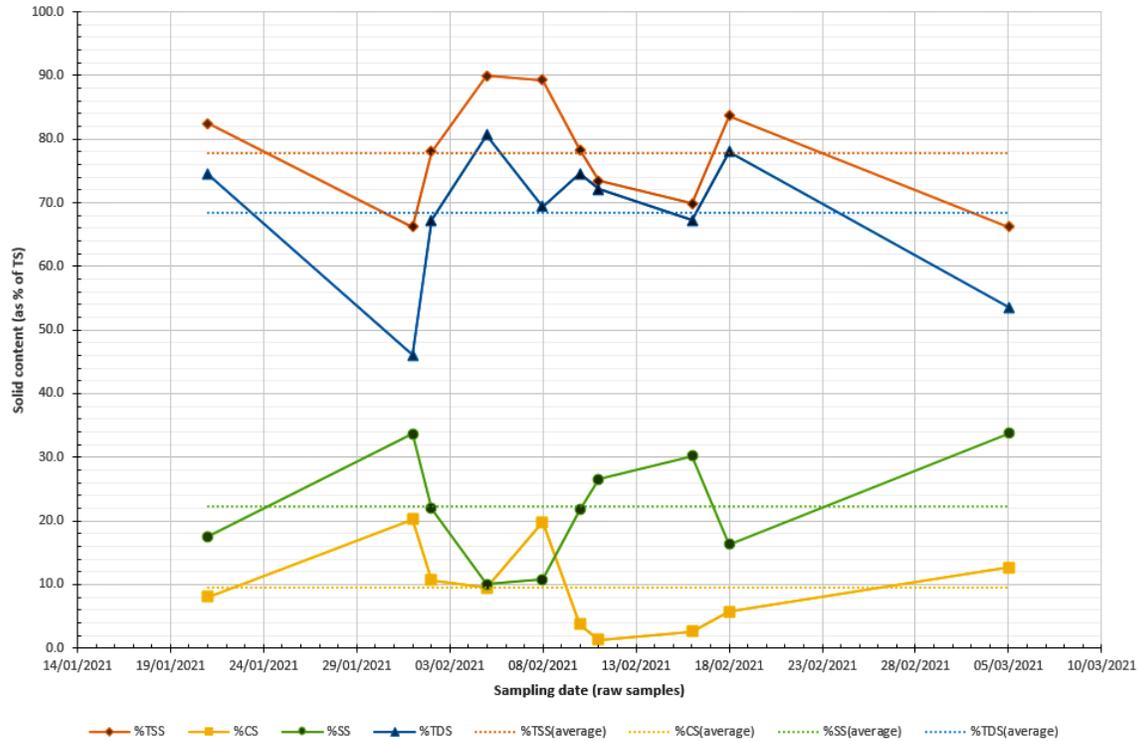


Figure 4.6: Solid content of raw samples as percentage of the Total Solids (TS): Total Suspended Solids (TSS), Colloidal Solids (CS), Settleable Solids (SS), Total Dissolved Solids (TDS)

to the corresponding Italian law limit for discharge in surface water bodies (TAB.3 all. 5 D.Lgs 152/06), which is: 160 mg/l for COD (as O₂), 2 mg/l for iron, 0.5 mg/l for zinc, 0.2 mg/l for lead and 1 mg/l for aluminium. Even though the number of samples is not sufficient for a satisfactory statistical analysis, it is still useful to infer some reasonable hypothesis.

Looking at samples 10/02/2021, 11/02/2021, 16/02/2021 and 05/03/2021 in both fig. 4.7 and fig. 4.6, it could reasonably be argued that most of the sediment present is attributable to supersaturation conditions of one or more of the dissolved metals (mostly iron and zinc), which proves the beneficial effect that a primary settling could have on the overall metal removal yield without the need for a second precipitation stage. It can also be seen that primary settling has also a slightly beneficial effect in terms of COD removal. Particle settling is one of the cheapest technologies in wastewater engineering - perhaps the cheapest for particle removal, alongside granular

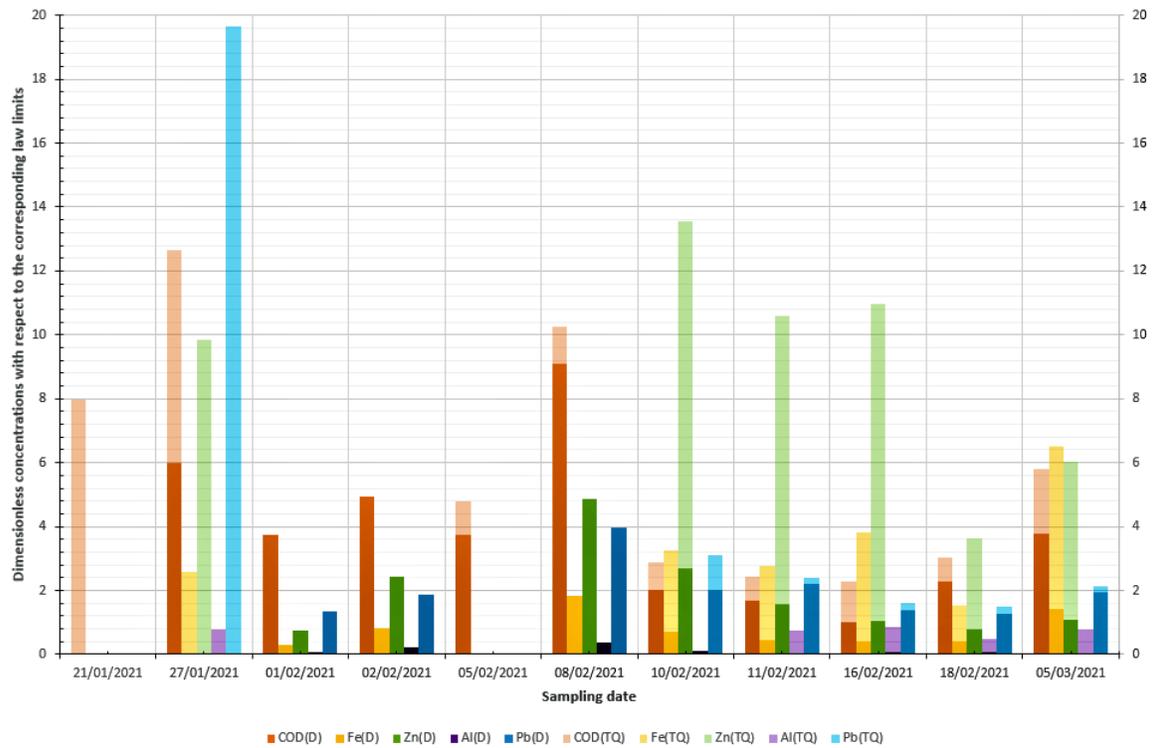


Figure 4.7: Comparison between raw (TQ, "Tal Quale") and settled (D, "Decantato") wastewater samples with respect to the main pollutants of interest. To give a better overall picture, each concentration has been dimensionless with respect to the corresponding Italian law limit for discharge in surface water bodies (TAB.3 all. 5 D.Lgs 152/06). Each concentration greater than 1 represents a value above the corresponding discharge limit

media filtration -, which also has an important and beneficial equalizing effect on the incoming flow (concentration peaks are smoothed out and hence control is less "stressed"). Considering that the observed settleable solids account for an average 22% of the total solids content (fig. 4.6) and that their removal could potentially improve measurement reliability of any online physical sensor/analyzer eventually installed prior to coagulation (which may particularly suffer from fouling and/or interference phenomena), a primary sedimentation seems more than justified.

The number of collected samples is by no means sufficient to derive a relationship between coagulant dosage and wastewater parameters. However, considering that the coagulant was (and currently is) added without initial pH correction and that pH was observed to vary from an average value of 7.2 (in raw samples) to a value never lower than 6.8 after coagulant addition, being an iron(III)-based coagulant (which is highly insoluble in water) it is generally safe to say that the main destabilization mechanism exploited by Drewo8016[®] is almost always the one identified as *sweep coagulation* (as can be seen in Figure 1.15 for the pH of interest, which ranges between 6.8 and 8.6 after coagulant and lime addition, respectively). From Chapter 1 and 2 it was seen that for this particular destabilization mechanism such a relationship does not exist, because coagulation via *sweep floc* is essentially governed by particle concentration (i.e. the flocculation rate depends almost solely on the collision frequency). Provided enough coagulant to trigger Fe(OH)₃ precipitation (which ultimately depends on iron(III) concentration and the solution pH), collisions are almost always effective thanks to precipitate enmeshment: if particle concentration is "low" the collision efficiency is practically 1.0, because eventually all particles become nucleated/coated by hydroxides and are seen by other particles as if they were hydroxides themselves; however, for appreciable floc formation, a generous amount of ferric chloride may still be required to raise the collision frequency (which is limited at low particle concentrations). Overdosing does not negatively affect collision efficiency whatsoever because restabilization with ferric chloride never occurs in the pH range of interest due to its very poor solubility. On the other hand, if particle concentration is "high" (meaning that there is a significant disproportion between the volume of hydroxides and the total surface area of particles), collisions are almost as effective, because particles/flocs would have just to reorient themselves to show each other surfaces that are covered with similar precipitates in order to "stick" together.

Table 4.1: Comparison between the two-stage treatment currently performed in the Mirafiori IDAC unit and the proposed alternative treatment

Chemical	1st stage	2nd stage	Alternative
Coagulant (ppm)	120 (FeCl ₃)	80 (FeCl ₃)	45 ± 5 (Drewo8016 [®])
Dreflo975 [®] (ppm)	4	3	2 ± 0.5
Bentonite (ppm)	-	N.D. (18 g/l)	250-900

Due to the addition of bentonite, it is rare that flocculation rate is limited by low concentrations of particles, therefore the study of coagulation with Drewo8016[®] should have focused more on the search for interfering adsorbates that could potentially reduce its performance.

Experimentally it was observed (table 4.1 and table 4.3) that the proposed solution (primary sedimentation followed by a single coagulation-flocculation-precipitation stage) proved to be equally effective but significantly less intensive in terms of chemicals used compared to the treatment currently performed in the Mirafiori IDAC unit (two coagulation-flocculation-precipitation stages in series, table 4.1). Indeed, post-treatment concentrations of heavy metals (Fe, Zn, Pb, Al) were all found to be lower than the corresponding law limits in each of the tests carried out (as can be seen in table 4.3). It was also found a moderately beneficial effect of bentonite on the abatement of nonionic surfactants, whereas for anionic surfactants such an effect was not observed. The little surprising effect of bentonite on COD could be justified by the fact that it was later discovered that the bentonite powder used for the analyses was actually an alternative prototype to be tested than the one currently used in their plants (which most likely has better adsorbing properties). In this regard, it was later found that the TSS/TS ratio of the prepared bentonite suspensions had an average value of 0.18, meaning that they were not really colloidal suspensions, hence the reduced adsorptivity.

Both the relationships "Bentonite concentration (in terms of TS) vs. Turbidity" and "Total Solids concentration (of the *settled* raw water samples added with bentonite) vs. Turbidity" shown in fig. 4.8 and fig. 4.9 respectively seem to be linear. The difference in intercepts in both graphs can be justified by the fact that the initial turbidity is probably not influenced solely by particle concentration, as can be seen from table 4.2. To give an example, despite their lower TSS content, samples 02-02-TQd, 02-11-TQd,

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

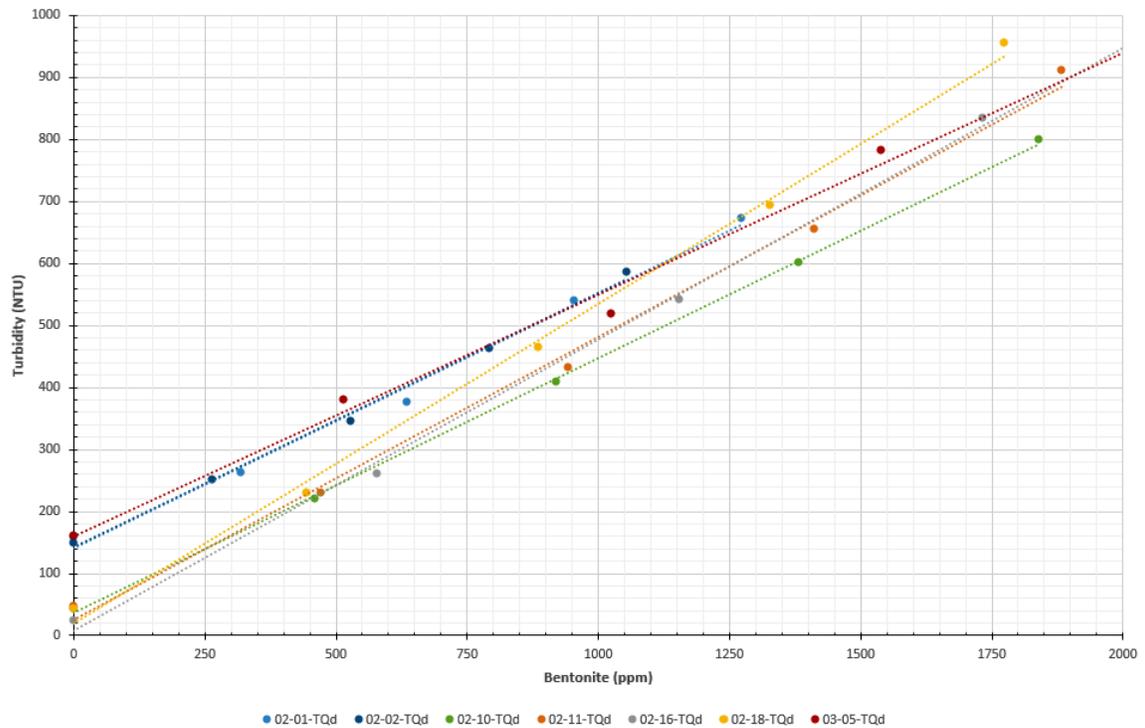


Figure 4.8: Relationship between Bentonite concentration (TS) and Turbidity of the *settled* raw water samples added with bentonite (1 ppm = 1 mg/l)

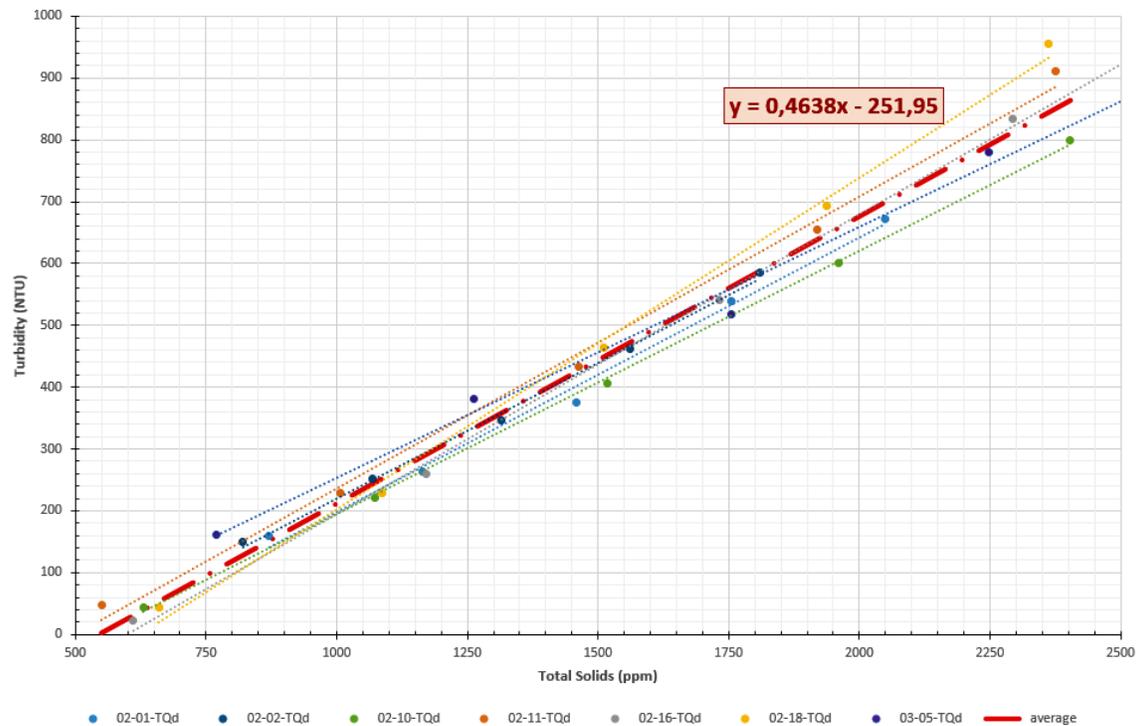


Figure 4.9: Relationship between Total Solids (of the *settled* raw water samples added with bentonite) and Turbidity (1 ppm = 1 mg/l)

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

Table 4.2: Initial Total Suspended Solids, COD, UV absorbance (integral between 190 and 280 nm) and Turbidity in settled raw water samples

Sample	TSS _{in} (mg/l)	COD _{in} (mg/l)	UV _{TOT,in} (-)	T _{in} (NTU)
02-01-TQd	872	578	21.41	159.0
02-02-TQd	815	780	29.11	149.0
02-10-TQd	629	313	21.88	43.5
02-11-TQd	554	264	20.44	47.3
02-16-TQd	609	160	14.46	22.3
02-18-TQd	660	366	21.65	44.2
03-05-TQd	769	602	27.90	160.5

03-05-TQd showed higher (or somewhat similar) turbidity values compared to samples 02-01-TQd, 02-16-TQd and 02-01-TQd respectively, which means that there are most likely other factors affecting turbidity measurements (like COD for example). It might be interesting to further investigate whether or not it is possible to derive a relationship between the Initial Turbidity and the initial COD (in terms of UV absorbance) and TSS concentrations. Considering the non-negligible measurement uncertainty associated with gravimetric solids estimation procedures, much more data would have been needed to develop such a relationship through multivariate regression (for example). However, considering that most of the TSS online sensors available on the market use a Multi-Beam Pulsed Infrared Light technology for their estimation, it could theoretically be possible to invert the above relationship to further strengthen the influent COD estimation via both infrared TSS and UV absorbance measurements. Indeed, it should be noted that particulate and soluble COD measurements may differ significantly from one another, so that a relationship of the form $COD_{in} = f(UV_{TOT,in}, T_{in}, TSS_{in})$ is likely to be nonlinear.

As an example, in fig. 4.10 and fig. 4.11 are reported the UV absorbance spectra of two samples subjected to the proposed treatment during jar testing (primary sedimentation followed by a single flocculation-precipitation stage, with ("-xNTU") and without ("-NB") bentonite addition). Once it was understood that there was a relationship between COD and UV absorbance, the UV spectra were used to discern in each experiment which treated samples chose to undergo chemical analyses due to the long time required to get a COD measurement (generally 30-40 minutes). With respect to fig. 4.10, the initial COD of the settled raw sample (which was equal to

Table 4.3: Concentrations and Turbidity in raw water samples ("-TQ"), raw *settled* samples ("-TQd"), and *settled* samples *treated* with ("-xNTU") and without bentonite addition ("-NB"). Final pH between 8.3 and 9.0

Sample	T (NTU)	Fe (mg/l)	Zn (mg/l)	Pb (mg/l)	COD (mg/l)
02-10-TQ	258.0	6.51	6.78	0.622	463
02-10-TQd	43.5	1.42	1.34	0.402	313
02-10-A(NB)	5.8	0.123	0.439	-	275
02-10-D(200NTU)	6.0	0.142	0.313	0.138	260
02-11-TQ	269.0	5.57	5.29	0.479	386
02-11-TQd	47.3	0.92	0.782	0.441	264
02-11-A(NB)	5.4	0.116	N.D.	-	217
02-11-C(200NTU)	6.7	0.096	N.D.	0.171	208
02-16-TQ	279.0	7.65	5.49	0.275	365
02-16-TQd	22.3	0.816	0.525	0.324	160
02-16-A(NB)	2.2	0.057	N.D	-	168
02-16-B(200NTU)	3.0	0.074	N.D	0.149	126
02-18-TQ	196.0	3.06	1.82	0.297	487
02-18-TQd	44.2	0.843	0.391	0.256	366
02-18-A(NB)	5.7	0.167	N.D.	-	308
02-18-D(400NTU)	5.7	0.210	N.D.	0.165	280
03-05-TQ	742.0	13	3.02	0.427	928
03-05-TQd	160.5	2.84	0.55	0.392	602
03-05-A(NB)	22.9	0.332	-	0.196	510
03-05-E(300NTU)	19.3	0.453	N.D.	0.197	479

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

602 mg/l) reduced to 510 mg/l after treatment without bentonite and to a value of 479 mg/l when treated with bentonite mixed until a turbidity value of 300 NTU was reached. Before knowing the COD results, sample 03-05-E(300NTU) was taken for subsequent chemical analyses (along with sample 03-05-A-NB) instead of samples 03-05-D-400NTU or 03-05-B-200NTU, because the relative absorbance spectrum was lower than the other two. The same goes for raw sample 02-01-TQd in fig. 4.11, for which sample 02-18-D-(400NTU) was chosen for further chemical analyses alongside sample 02-18-A-NB instead of the other two. In both examples, the two samples chosen for chemical analyses were subsequently treated with Granular Activated Carbon (GAC) to obtain two additional absorbance spectra to be further correlated with COD, to verify if the same COD vs. UV absorbance relationship could be used also to monitor COD in the downstream process.

From fig. 4.12 and fig. 4.13 it appears that an exponential trendline fits better with the experimental data than a power trendline. In abscissa is reported the integrals of the UV absorbance spectra between 190 and 280 nm (i.e. the area underlying each UV absorbance spectrum acquired during jar tests and for which the corresponding COD was also measured). Such a relationship is obviously meant to be intended as site-specific. Each dot in the figure was depicted with a different color depending on the type of treatment it was subjected to (i.e. settling, settling + coagulation-flocculation without bentonite addition, settling + coagulation-flocculation with bentonite addition, settling + coagulation-flocculation w/ and w/o bentonite addition + GAC). Although the number of collected samples in each group is not sufficient to statistically infer some physical hypothesis, it would be interesting to verify if there is indeed a linear relationship between COD and total UV absorbance within each group, in accordance with the Beer-Lambert law (i.e. concentration of attenuating species proportional to the relative absorbances). Measurement uncertainties related to COD could be reduced by increasing the number of samples. Since COD is an aggregate parameter that returns only an indirect estimate of the quantity of oxidizable substances present in a sample, it is not guaranteed that such a linear relationship can be observed mainly for two reasons. First, for a given COD value there can be theoretically infinite combinations of mixed substances of different nature in the same wastewater sample (i.e. with different chemical formulas) that evidently cannot be isolated and measured individually and that are likely to exhibit slightly different absorbance spectra. The double and

triple bonds that absorb UV radiation are quantitatively different in molecules of different structures. Second, Beer-Lambert law gets more complicated when considering more than one attenuating species and more than one wavelength at a time, hence deviations from linearity are somewhat kind of expected. Nevertheless, it appears that a relationship between COD and total UV absorbance does indeed exist, as can be seen in fig. 4.12 both from an overall point of view (exponential trendline) and from a local one (it does seem that linear trendlines with different intercepts and slopes can effectively be drawn for each treatment group). This different behavior could also be interpreted in terms of selectivity: different treatment solutions (e.g. bentonite, bentonite + GAC, GAC alone, etc.) are expected to be more or less selective/effective towards specific substances (e.g. nonpolar molecules, fats and grease, surfactants, etc.) in different proportions from one other, hence it is likely that given the same COD value, the corresponding UV absorbance spectra associated with different treatment options are different. Further investigations should be carried out to better understand the aspects mentioned so far.

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

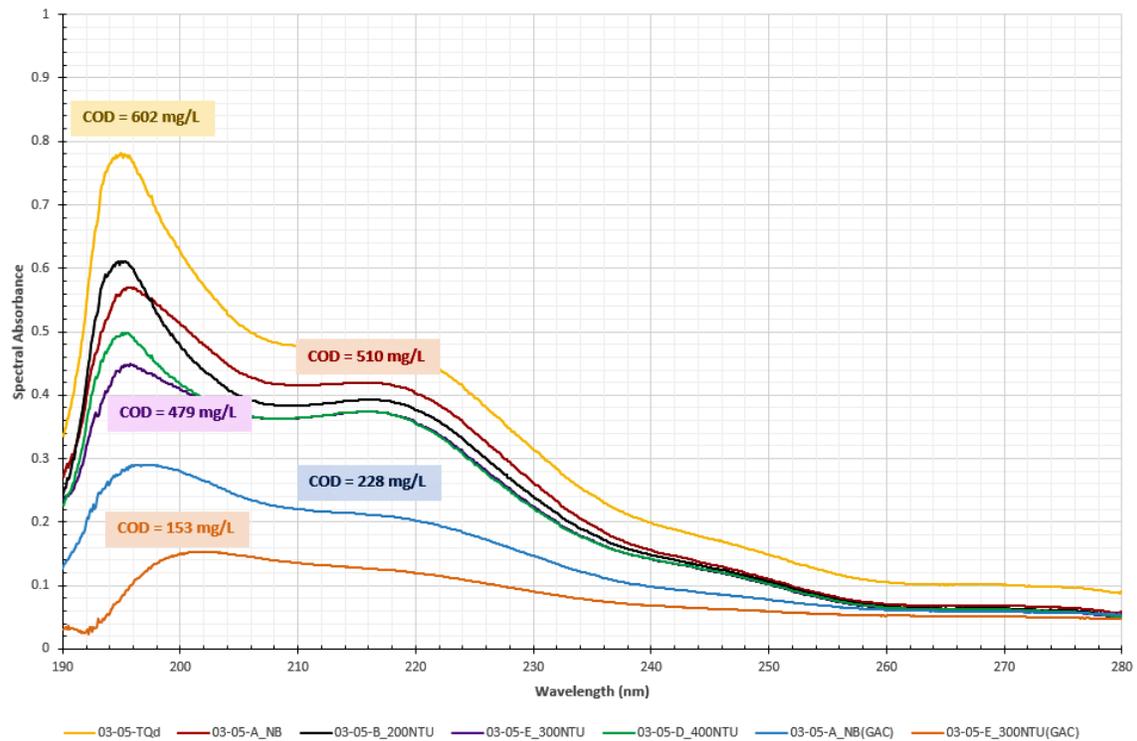


Figure 4.10: UV absorbance spectra of sample 03-05-TQd. The optimal dosage of bentonite was determined as that for which a further addition of bentonite did not appear to be significantly more beneficial in terms of lowering UV absorbance (i.e. COD removal). "-TQd" → *settled* raw wastewater sample, "-NB" → *settled* sample treated without bentonite addition, "-xNTU" → *settled* sample treated adding bentonite until a turbidity equal to "x" is reached, "GAC" → clarified water further treated with Granular Activated Carbon

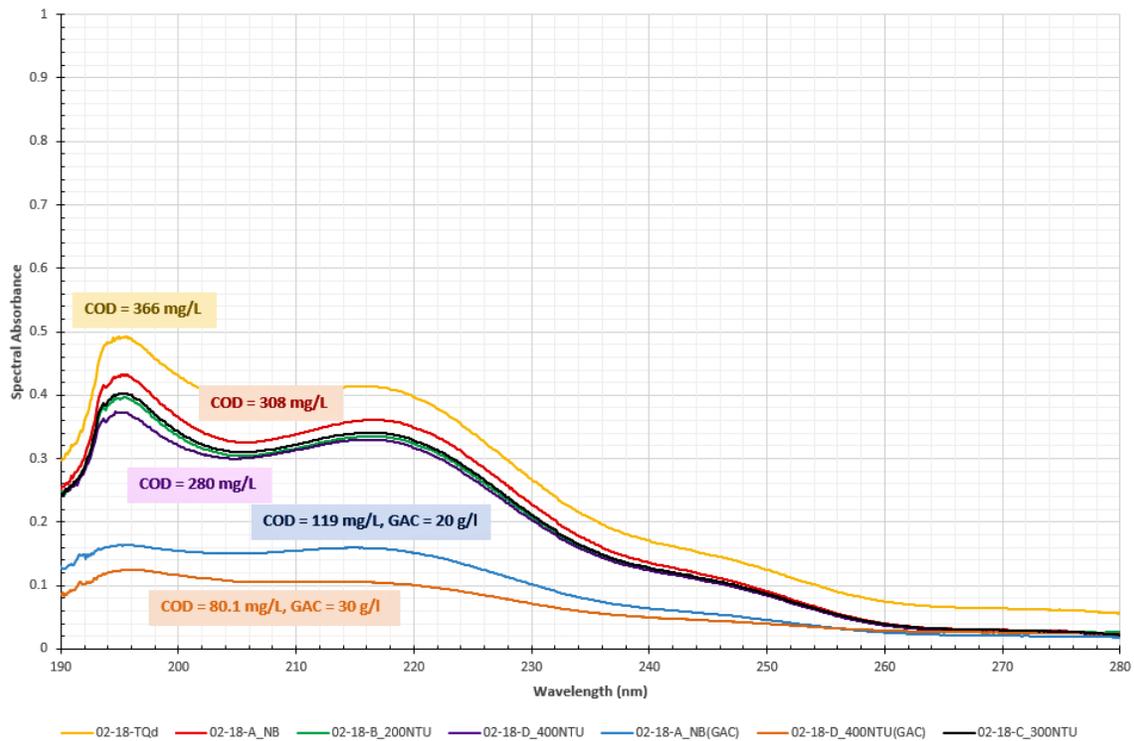


Figure 4.11: UV absorbance spectra of sample 02-18-TQd. The optimal dosage of bentonite was determined as that for which a further addition did not appear to be significantly more beneficial in terms of lowering UV absorbance (i.e. COD removal). "-TQd" → *settled* raw wastewater sample, "-NB" → *settled* sample treated without bentonite addition, "-xNTU" → *settled* sample treated adding bentonite until a turbidity equal to "x" is reached, "GAC" → clarified water further treated with Granular Activated Carbon

4. Case Study: the Automotive Painting Wastewater Treatment Plant of FCA Mirafiori (TO)

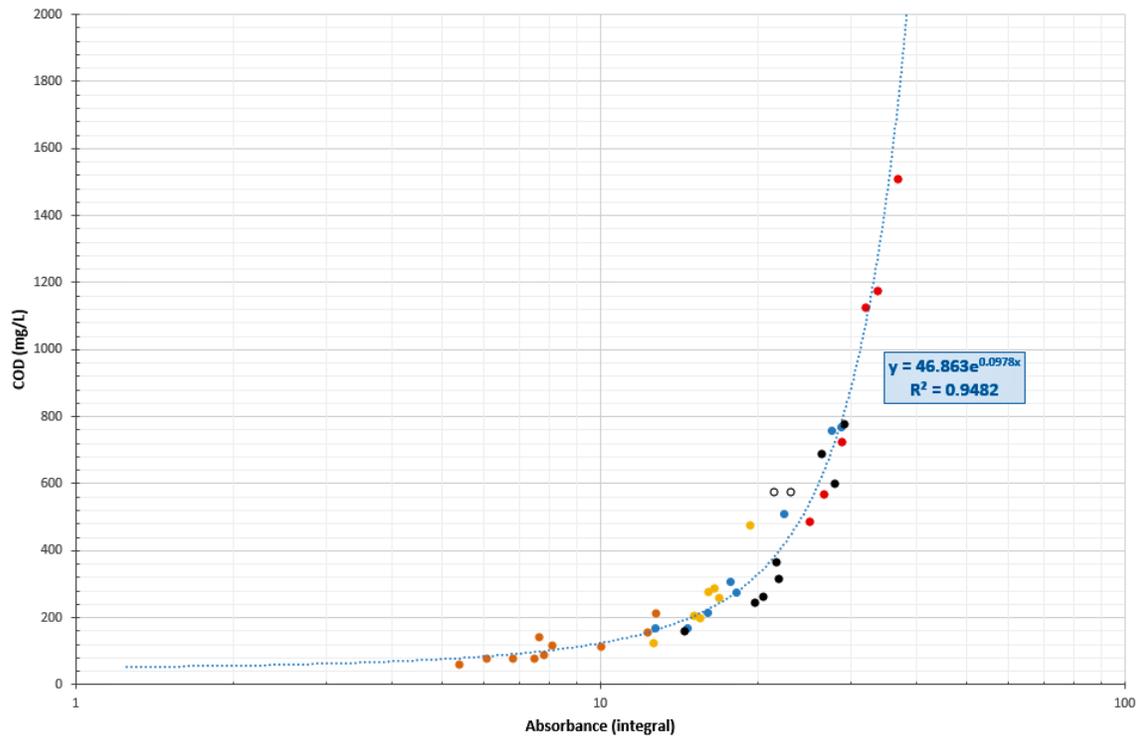


Figure 4.12: Integral of the UV absorbance Spectra vs. COD in semi-log plot (exponential trend). **Black** dots: settled (raw) samples; **Blue** dots: settled samples treated without bentonite addition; **Yellow** dots: settled samples treated with bentonite addition; **Orange** dots: clarified samples treated with GAC; **Red** samples: settled (raw) samples added with concentrated detergents used for car bodywork washing; **White** dots: unrepresentative settled wastewater samples (stagnant water taken from the expansion tank instead of the open-air storage tanks)

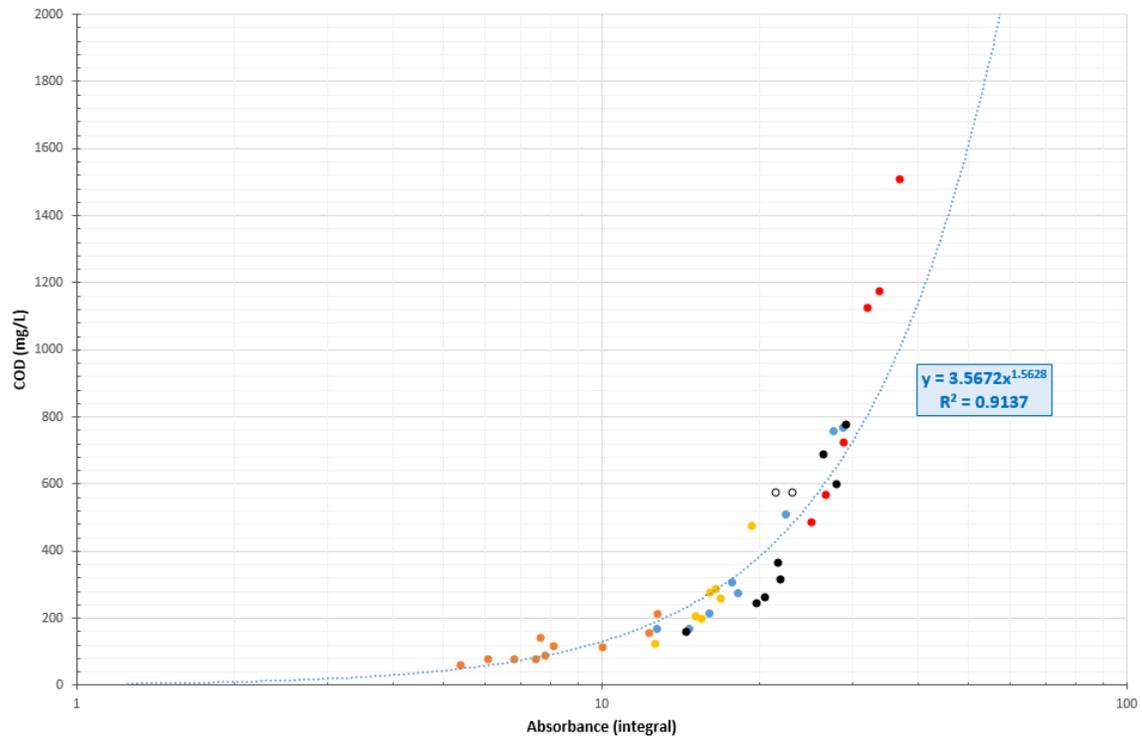


Figure 4.13: Integral of the UV absorbance Spectra vs. COD in semi-log plot (power trend). **Black** dots: settled (raw) samples; **Blue** dots: settled samples treated without bentonite addition; **Yellow** dots: settled samples treated with bentonite addition; **Orange** dots: clarified samples treated with GAC; **Red** samples: settled (raw) samples added with concentrated detergents used for car bodywork washing; **White** dots: unrepresentative settled wastewater samples (stagnant water taken from the expansion tank instead of the open-air storage tanks)

Chapter 5

Conclusions and Future Developments

In this thesis, the topics of flocculation modeling and coagulation control have been explored in sufficient detail to highlight those challenges and criticalities (both theoretical and practical) that place severe limitations on the design, optimization, predictability and controllability of clariflocculation processes, today as in the past.

From a theoretical point of view, despite its ancient origins and its still wide use, this type treatment - which usually relies on empirical design formulas and site-specific dosing protocols - still awaits the development of a quantitative physico-mathematical theory that could be practically implemented for design and control purposes. Although DLVO theory represents probably the simplest, widely recognized and yet sufficiently accurate theory for particle and molecule analysis, it cannot be used in contexts other than academic research because it is almost impossible to implement it within process simulation tools for real systems applications due to its excessive computational demand and dimensionality; the amount of variables and process parameters involved in a DLVO-based coagulation model for industrial wastewater applications may also be too large to allow for effective model calibration, assuming such variables (like zeta potential, ionic strength, etc.) could actually be measured. However, both DLVO theory and other semi-empirical models such as Hogg's model for particle flocculation using polymers are still useful in providing some meaningful insights into those factors that affect coagulation and flocculation the most. According to the scientific literature, flocculation can become a highly nonlinear process depending on the

particular destabilization mechanism exploited during coagulation, which is thought to be the main factor affecting the pseudo-second order flocculation rate of aggregates. To overcome this lack of an intuitive and easy to implement (but sufficiently explanatory) model for the collision efficiency factor, an extension to Hogg's model has been proposed, in order to account for multiple destabilization mechanisms and multiple coagulant species at once using a single expression for the collision efficiency function. The proposed model, which makes use of a Discretized Population Balance Equation of fractal aggregates within a Linear Driving Force framework for adsorption and surface precipitation has yet to be tested, but on paper it looks promising enough to be worth a try.

From a practical point of view, it has been shown that not all the key process variables can actually be measured (directly or indirectly) for multiple reasons. There may be physical or process constraints that do not allow for water sampling at specific points along the treatment line, perhaps because particles in the water sample are still flocculating, thus affecting the reliability of turbidity and other measurements (for example). There may be no physical sensors or analyzers capable of measuring certain key process variables because such variables cannot be measured at all (like ionic strength for example). Perhaps specific instruments related to coagulation already exist on the market (e.g. Streaming Current Analyzers), but cannot actually be installed because they severely suffer from fouling and/or interference phenomena, which are particularly relevant in industrial wastewater treatment applications. Another very important issue is related to dead time, both of physical sensors / analyzers (i.e. how much time they require to perform a measurement) and of the process itself. Usually, when clariflocculation is accomplished using sedimentation as the subsequent separation stage, a significant amount of process dead time is introduced, which could potentially degrade control performance and stability, especially when a feedback control based on treated water quality is performed (time lag between actuation and state observation).

The most interesting results that emerged at the end of the short jar test campaign were the two site-specific relationships found between the integral of the UV absorbance spectra of the treated and untreated wastewater samples (previously diluted with hydrochloric acid 0.01 M) and COD, and between the Total Solids concentration of the settled sample added with bentonite and turbidity. Both the relationships, once

5. Conclusions and Future Developments

strengthened with more data, could theoretically be implemented for online monitoring and control purposes, since a UV absorbance measurement typically takes at most 1 minute to be determined compared to a COD measurement (which usually requires at least 30-40 minutes to be taken). Since the coagulant used worked primarily via the *sweep coagulation* mechanism, it was not possible to find a stoichiometric relationship between wastewater parameters and optimal coagulant dosage. Such a relationship would have required a change of coagulant to be found, with all pros and cons that comes with that choice (direct costs, process reliability, flocs settleability, sludge volume, etc.).

Considering all the challenges and limitations above, researchers over the last two decades have gradually shifted interest towards much more suitable approaches, like machine learning techniques, which usually shine when it comes to complex, unknown and potentially highly nonlinear systems (which are already difficult to control using linear control theory). Artificial Neural Networks and Fuzzy Logic models have already been used in recent years with some success, although they usually require lots of data (and hence large amounts of money for jar testing procedures and pilot plant testing) to be trained and validated. The only issue that comes with these so-called "software sensors" is that they do not provide explicit information to the user on the patterns among the key process variables, making them less generalizable and hence less appealing especially for large companies, which are generally interested in implementing the same technology in all the plants they manage without having to repeat the whole training procedure. With the development in recent times of data-driven algorithms for model discovery like the so-called Dynamic Mode Decomposition (DMD) and Sparse Identification of Nonlinear Dynamics (SINDy) it was possible to overcome this "black-box" (sometimes "grey-box") nature of traditional machine learning algorithms.

As a future research, it would be interesting to implement the proposed extension to Hogg's model within a numerical model to simulate the complex behavior of a real reactor - undergoing coagulation, flocculation, precipitation and adsorption - by producing a sufficiently large dataset of fictitious measurements under different loading scenarios. Once such dataset have been produced, it would be interesting to test if the above mentioned data-driven algorithms for model discovery (i.e. DMD and SINDy) could prove capable of modeling such a complex system. If so, pilot plant tests with

on-line data acquisition for the direct application of these algorithms would then be economically justified.

Bibliography

- [1] Bratby, J. (2016). *Coagulation and flocculation in water and wastewater treatment (third edition)*. London, United Kingdom, IWA Publishing.
- [2] Benjamin, M. M. (2015). *Water chemistry - second edition*. Long Grive, Illinois, Waveland Press, inc.
- [3] Lee, D. G., Bonner, J. S., Garton, L. S., Ernest, N., & Autenrieth, R. L. (2000). Modeling coagulation kinetics incorporating fractal theories: A fractal rectilinear approach. *Water Research*, 34(7), 1987–2000.
- [4] Logan, B. (2012). *Environmental transport processes - 2nd edition*. Wiley.
- [5] Veerapaneni, S., & Wiesner, M. R. (1996). Hydrodynamics of fractal aggregates with radially varying permeability. *Journal of Colloid and Interface Science*, 177(0005), 45–57.
- [6] Happel, J. (1958). Viscous flow in multiparticle systems: Slow motion of fluids relative to beds of spherical particles. *AIChE Journal*, 197–201.
- [7] Chellam, S., & Wiesner, M. R. (1993). Fluid mechanics and fractal aggregates. *Water Research*, 27(9), 1493–1496.
- [8] Neale, G., & Epstein, N. (1972). Creeping flow relative to permeable spheres. *Chemical Engineering Science*, 28, 1865–1874.
- [9] Adler, P. M. (1981). Streamlines in and around porous particles. *Colloid Interface Science*, 81, 531–535.

-
- [10] Farley, K., Dzombak, D., & Morel, F. (1985). A surface precipitation model for the sorption of cations on metal oxides. *Journal of Colloid Interface Science*, *106*, 226–242.
- [11] Benjamin, M. M., & Lawler, D. F. (2016). *Water quality engineering - physical/chemical treatment processes*. London, United Kingdom, IWA Publishing.
- [12] Hogg, R. (1984a). Collision efficiency factors for polymer flocculation. *Journal of Colloid and Interface Science*, *102*(1), 6574–6597.
- [13] Healy, T., & La Mer, V. K. (1964). The energetics of flocculation and redispersion by polymers. *Journal of Colloid Science*, *19*(4), 323–332.
- [14] Smellie, R. H., & La Mer, V. K. (1958). Flocculation, subsidence and filtration of phosphate slimes: Vi. a quantitative theory of filtration of flocculated suspensions. *Journal of Colloid Science*, *13*(6), 589–599.
- [15] Mirville, R. J. (1978). The interaction of blackwater constituents with polymeric flocculants. *MS Thesis, The Pennsylvania State University*.
- [16] Somasundaran, P., Chia, Y. H., & Gorelik, R. (1984). Polymer adsorption and dispersion stability. *ACS Symposium Series, American Chemical Society, Washington D.C.*, (240).
- [17] Dirican, C. (1981). The structure and growth of aggregates in flocculation. *MS Thesis, The Pennsylvania State University*.
- [18] Scheutjens, J., & Fleer, G. J. (1979). Statistical theory of the adsorption of interacting chain molecules. 1. partition function, segment density distribution, and adsorption isotherms. *J. Phys. Chem*, *83*(12), 1619–1635.
- [19] Fleer, O. J., & Scheutjens, J. (1986). Interaction between adsorbed layers of macromolecules. *Journal of Colloid and Interface Science*, *111*(2), 504–515.
- [20] Hogg, R. (1984b). Polymer adsorption and dispersion stability. *ACS Symposium Series, American Chemical Society, Washington, D. C.*, (240).

- [21] Ratnaweera, H., & Fettig, J. (2015). State of the art of online monitoring and control of the coagulation process. *Water*, 7, 6574–6597.
- [22] Hahn, H., Hoffmann, E., & Ødegaard, H. (2004). *Coagulant dosing control—a review*. London, UK, IWA Publishing.
- [23] Brunton, S., & Kutz, J. (2019). *Data-driven science and engineering: Machine learning, dynamical systems and control*. Cambridge University Press.
- [24] Kutz, J., Brunton, S., B.W., B., & Proctor, J. (2016). *Dynamic mode decomposition: Data-driven modeling of complex systems*. SIAM.
- [25] Brunton, S., & Proctor, J. N., J. L. ad Kutz. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15), 3932–3937.

List of Figures

1.1	Size spectrum of waterborne particles and filter pores. <i>Source:</i> from [1] .	2
1.2	Position and orientation of water molecules in the electrical double layer and their role on electrical permittivity. <i>Source:</i> from [1]	9
1.3	Energy changes during adsorption. <i>Source:</i> from [1]	12
1.4	Conceptual representation of the Stern's model of the EDL (a, from Priesing,1962) and counter-ion vs. co-ion adsorption in the Stern Model (b). <i>Source:</i> from [1]	15
1.5	Charge reversal within the Stern layer due to polyvalent or surface-active counter-ions (a) and adsorption of surface-active co-ions increasing effective surface charge (b). <i>Source:</i> from [1]	16
1.6	Schematic of the interfacial structure near an iron oxide surface, as envisioned in the Triple Layer Model. <i>Source:</i> from [2, p. 316]	23
1.7	Potential energy of interaction between two particles. <i>Source:</i> from [1, p. 29]	30
1.8	Shearing of diffuse layer by flow of water around a charged particle. <i>Source:</i> from [2, p. 531]	32
1.9	Compression of the diffuse layer at high ionic strength (particle is assumed to be negatively charged and the Debye-Huckel length is represented by the dotted circles). <i>Source:</i> from [2, p. 535]	34
1.10	Particle destabilization by charge neutralization. <i>Source:</i> from [2, p. 536]	35
1.11	Schematic of possible metal hydroxide-particle interaction (adapted from Dentel & Grosset, 1987 and 1988)	37

1.12	Adsorption of metal ions (the little spheres) onto a particle (the large sphere) leading to the formation of a precipitate. The lines in the right part of the figure indicate the lattice formation of the precipitate. <i>Source:</i> from [2, p. 538]	39
1.13	Destabilization by two different mechanisms by metal addition. Solid line represents a given suspension at a certain pH, where the two mechanisms are clearly distinct; dotted line represents a more concentrated suspension at the same pH (<i>Source:</i> from [2, p. 538])	40
1.14	Alum destabilization diagram. <i>Source:</i> from Amirtharajah and Mills (1982), as redrawn in Edzwald et al. (1998)	41
1.15	Iron destabilization diagram. <i>Source:</i> from Johnson and Amirtharajah (1983), as redrawn in Edzwald et al. (1998)	42
1.16	Interparticle bridging by a high-molecular weight polymer (a), ineffective destabilization at low particle concentration (b) and effects of polymer overdosing (c) (adapted from [2])	45
1.17	<i>Double layer compression</i> as the predominant mechanism, with little effect of applied coagulant concentration on the residual turbidity (a) and <i>adsorption and charge neutralization</i> as the predominant mechanism (b). (adapted from [1])	50
1.18	Destabilization and restabilization regions at a given pH value as related to colloid concentration (S) and coagulant dosage (C_t) (<i>Source:</i> from [1, p. 87], adapted from Stumm & O'Melia, 1968)	52
1.19	Zones of destabilization and restabilization with respect to coagulation pH, colloid concentration (S) and coagulant dosage (C_t) (<i>Source:</i> from [1, p. 90], adapted from Stumm & O'Melia, 1968)	56
1.20	Performance of ferric chloride and aluminium sulfate at 1°C (<i>Source:</i> from [1, p. 225], adapted from Morris & Knocke, 1984)	58
2.1	Particle paths leading to collisions according to the three most important mechanisms of long-range motion (the length of the arrows reflects the particle velocities). (<i>Source:</i> from [2, p. 565])	63
2.2	Conceptual view of collisions by fluid shear. (<i>Source:</i> from [2, p. 572]) .	71

2.3	Collision frequency function for fluid shear ($G = 25 \text{ s}^{-1}$). (Source: from [2, p. 574])	72
2.4	Conceptual representation of the collision by differential sedimentation. (Source: from [2, p. 575])	75
2.5	Collision frequency function for differential sedimentation ($\rho_i = \rho_j = 1.2 \text{ g/cm}^3$). (Source: from [2, p. 576])	75
2.6	Collision frequency function for Brownian motion ($T = 293 \text{ K}$). (Source: from [2, p. 578])	77
2.7	Total collision frequency function according to the long-range force model. (Source: from [2, p. 579])	78
2.8	Dominant regions for each collision mechanism according to the long-range force model, with borders between regions varying with environmental and operational conditions. (Source: from [2, p. 580]) . .	80
2.9	Trajectories of particles undergoing differential sedimentation according to the long-range and short-range models. (Source: from [2, p. 582]) . .	82
2.10	Collision efficiency functions in the short-range force model. (Source: from [2, p. 583])	84
2.11	Total collision efficiency function according to the short-range force model for the specific set of conditions shown in the picture. (Source: from [2, p. 587])	87
2.12	Dominant regions for each collision mechanism according to the short-range force model. (Source: from [2, p. 588])	89
2.13	Euclidean Model (perfectly coalescing spheres, a) and Fractal Model ("Coalesced Fractal Spheres" model (CFS), b). (Source: adapted from [4] and [3])	90
2.14	Streamlines passing through the same porous aggregate for different values of the fractal dimension D , which deeply affects aggregate permeability (left). Radial variation of permeability within an aggregate for a given fractal dimension D (right). (Source: adapted from [5]) . . .	93
2.15	Ω and η as functions of fractal dimensions for aggregates of different sizes (expressed in terms of r/a_p ratios) (a) and Ω vs. η at different fractal dimensions for four aggregate sizes (b). (Source: from [5])	94

2.16	Schematic of a metal adsorbate Me that is first bound to discrete sites on the adsorbent, but then forms a surface precipitate covering much of the solid up to a point where the surface of the final product behaves almost like that of a particle of pure metal precipitate. (<i>Source:</i> from [11, p. 836])	97
2.17	Apparent adsorption isotherms for systems in which both Langmuir adsorption and surface precipitation occur simultaneously. On the left, isotherms for adsorption and surface precipitation separately (a); on the right, apparent isotherms for simultaneous adsorption and surface precipitation (b). (<i>Source:</i> from [11, p. 839])	98
2.18	Collision efficiency factors for equisized particles (a) and for particles of different sizes (b) as functions of the surface coverage θ . <i>Source:</i> from [12]	105
2.19	Schematic of a typical polymer adsorption isotherm, showing the plateau region at intermediate polymer concentration and surface coverage. <i>Source:</i> from [12]	106
2.20	Schematic of the surface coverages of a particle/floc with adsorption sites grouped by type of adsorbent exposed on the surface, i.e. bentonite ("b") or precipitates ("SP" - " <i>surface precipitates</i> "). The surface sites responsible for particle adhesion are circled in purple.	109
3.1	Sketch of a feed-forward control scheme based on raw water quality. (<i>Source:</i> from [21])	128
3.2	Sketch of a feed-backward control scheme based on dosed water quality. (<i>Source:</i> from [21])	128
3.3	Sketch of a feed-backward control scheme based on finished water quality. (<i>Source:</i> from [21])	133

3.4	Overview of DMD illustrated on the fluid flow past a circular cylinder at Reynolds number 100. 1) Data (images in this case) are collected over time within the entire spatial domain and stacked into extremely tall column vectors x_i . 2) A <i>matrix of snapshots</i> X is built by placing the different instantaneous column vectors x_i side by side. 3) A second matrix of snapshots X' is built by moving each column vector one step forward in time. 4) The best fit linear operator \tilde{A} describing the temporal evolution of the state $X' = AX$ is derived by performing a Singular Value Decomposition of the pseudo-inverse $A = X'X^\dagger$ without actually computing it. 5) Eigenvalues and eigenvectors of \tilde{A} extracted this way represent the dominant modes and coherent structures of the discovered dynamical system (in terms of energy and frequency content) (<i>Source: from [23], Reproduced from: [24]</i>)	143
3.5	Schematic of the Sparse Identification of Nonlinear Dynamics (SINDy) algorithm applied to the Lorentz attractor. Parsimonious models (i.e. sparse coefficients of the Ξ matrix) are extracted using a library $\Theta(X)$ of possible candidate nonlinear functions using sparse regression (e.g. LASSO). The $\Theta(X)$ library may be constructed purely from measurement data. (<i>Source: from [23], Reproduced from: [25]</i>)	144
4.1	Jar test apparatus, portable turbidimeter, beakers and magnetic stirrer (upper left), portable turbidimeter and pH-meter and quartz cuvettes (upper right), desiccator and membrane filters (lower left), cuvet kit tests for chemical analyses, adjustable micropipette, vacuum pump glass chamber and spectrophotometer DR6000 (lower right)	149
4.2	1:50 dilution ratio of Drewo8016 [®] with demineralized water (left) and freshly prepared bentonite suspension (right)	150
4.3	Sample before treatment (settled and unsettled)(upper left), treatment with bentonite and a slight excess of Drewo8016 [®] (upper right), settled flocs after Dreflo975 [®] addition using slow mixing for 15-20 minutes (lower left), treatment of a clarified sample with Granular Activated Carbon (GAC) (lower right).	151

4.4	absorbance spectra of a settled, raw wastewater sample before (TQd) and after several treatment options. 1:10 dilution ratios with demineralized water	153
4.5	absorbance spectra of the same settled, raw wastewater sample before (TQd) and after several treatment options. 1:10 dilution ratios with HCl 0.01 M	154
4.6	Solid content of raw samples as percentage of the Total Solids (TS): Total Suspended Solids (TSS), Colloidal Solids (CS), Settleable Solids (SS), Total Dissolved Solids (TDS)	158
4.7	Comparison between raw (TQ, "Tal Quale") and settled (D, "Decantato") wastewater samples with respect to the main pollutants of interest. To give a better overall picture, each concentration has been dimensionless with respect to the corresponding Italian law limit for discharge in surface water bodies (TAB.3 all. 5 D.Lgs 152/06). Each concentration greater than 1 represents a value above the corresponding discharge limit	159
4.8	Relationship between Bentonite concentration (TS) and Turbidity of the <i>settled</i> raw water samples added with bentonite (1 ppm = 1 mg/l) . . .	162
4.9	Relationship between Total Solids (of the <i>settled</i> raw water samples added with bentonite) and Turbidity (1 ppm = 1 mg/l)	163
4.10	UV absorbance spectra of sample 03-05-TQd. The optimal dosage of bentonite was determined as that for which a further addition of bentonite did not appear to be significantly more beneficial in terms of lowering UV absorbance (i.e. COD removal). "-TQd" → <i>settled</i> raw wastewater sample, "-NB" → <i>settled</i> sample treated without bentonite addition, "-xNTU" → <i>settled</i> sample treated adding bentonite until a turbidity equal to "x" is reached, "GAC" → clarified water further treated with Granular Activated Carbon	168

4.11 UV absorbance spectra of sample 02-18-TQd. The optimal dosage of bentonite was determined as that for which a further addition did not appear to be significantly more beneficial in terms of lowering UV absorbance (i.e. COD removal). "-TQd" → *settled* raw wastewater sample, "-NB" → *settled* sample treated without bentonite addition, "-xNTU" → *settled* sample treated adding bentonite until a turbidity equal to "x" is reached, "GAC" → clarified water further treated with Granular Activated Carbon 169

4.12 Integral of the UV absorbance Spectra vs. COD in semi-log plot (exponential trend). **Black** dots: settled (raw) samples; **Blue** dots: settled samples treated without bentonite addition; **Yellow** dots: settled samples treated with bentonite addition; **Orange** dots: clarified samples treated with GAC; **Red** samples: settled (raw) samples added with concentrated detergents used for car bodywork washing; **White** dots: unrepresentative settled wastewater samples (stagnant water taken from the expansion tank instead of the open-air storage tanks) 170

4.13 Integral of the UV absorbance Spectra vs. COD in semi-log plot (power trend). **Black** dots: settled (raw) samples; **Blue** dots: settled samples treated without bentonite addition; **Yellow** dots: settled samples treated with bentonite addition; **Orange** dots: clarified samples treated with GAC; **Red** samples: settled (raw) samples added with concentrated detergents used for car bodywork washing; **White** dots: unrepresentative settled wastewater samples (stagnant water taken from the expansion tank instead of the open-air storage tanks) 171

List of Figures

List of Tables

1.1	Classification of Solid-Liquid Dispersions based on particle size. <i>Source:</i> from [1]	2
1.2	Characteristics of destabilization mechanisms with metal coagulants (<i>Source:</i> from [1, p. 83], adapted from Stumm & O'Melia, 1968)	48
2.1	Collision efficiency for fluid shear. (<i>Source:</i> from [2, p. 598])	83
2.2	Collision efficiency for differential sedimentation. (<i>Source:</i> from [2, p. 598])	85
2.3	Collision efficiency for Brownian motion. (<i>Source:</i> from [2, p. 597])	85
3.1	Online sensors applied for monitoring and control of clariflocculation processes according to experiences and applications in Northern Europe/Scandinavia (<i>Source:</i> from [21])	120
4.1	Comparison between the two-stage treatment currently performed in the Mirafiori IDAC unit and the proposed alternative treatment	161
4.2	Initial Total Suspended Solids, COD, UV absorbance (integral between 190 and 280 nm) and Turbidity in settled raw water samples	164
4.3	Concentrations and Turbidity in raw water samples ("-TQ"), raw <i>settled</i> samples ("-TQd"), and <i>settled</i> samples <i>treated</i> with ("-xNTU") and without bentonite addition ("-NB"). Final pH between 8.3 and 9.0	165

List of Tables
