VOLUME 1, NUMBER 1, MAY 20, 2016

Application of Engineering and Technology

ISSN Pending

Editor-in-Chief: Bob Lew.

1088 EMAIL PRESS

Application of Engineering and Technology

ISSN Pending

Volume 1, number 1, May 20, 2016

Impact Factor: Pending

Contents

157 East ELM Street, Unit A, Greenwich, CT 06830-6614, USA

E-mail: aet@1088.email

Online first

Experimental and Modelling Study of Low Velocity Impacts on Composite Sandwich Structures for Railway Applications

Adnene Sakly [1, *], Abdelouahed Laksimi [1], Hocine Kebir [1], Salim Benmedakhen [1, 2]

[1] Laboratoire Roberval UTC-CNRS, UMR 6253, Université de Technologie de Compiègne, BP 20529-60206 Compiègne Cedex, France

[2] Cybernetix: Parc Technologique des Rives de l'Oise, BP 50149-60201 Compiègne Cedex, France

* Corresponding Author. E-mail: : adnene.sakly@utc.fr

Abstract. This paper aims to evaluate the damage resistance of sandwich composites used in train structures subjected to low velocity impacts. These impacts can be induced by railway ballast projection incidents underneath the floor of the train. A high-speed and low weight test bench to simulate such impacts was developed in our laboratory. After the impact, the damage mechanism evolution as a function of solicitation conditions was inspected. Three-dimensional finite element models were implemented to analyze the impact response. Results obtained from numerical simulations have confirmed experimental observations. Shock wave propagation localized around the impact point confined the damage. Visual inspection identified delamination in composite skin, fiber breakage, and indentation. Skin/core debonding, core crushing and shear failure were analyzed. Damages intensified proportionally with the number of successive projectile launches. Correlations between impact energy and damage mode are discussed.

Keywords: Sandwich structures; Railway; Impact behavior; Failure analysis; Finite element modeling.

1. Introduction

Composite materials are inevitably used to construct structures where the weight is a major consideration. In order to save energy, less weight is primordial for all kinds of transport industries: aircraft, cars and trains. The low ratio of weight/hardness of foam-based sandwich panels, thermal isolation characteristics, as well as good energy-absorbing capacity, make these materials attractive for railway cars [1], [2] and [3]. In fact, sandwich composite materials are currently used for superstructures, doors, partition walls and chassis/train floors [4].

A major concern with the mechanical performance of these sandwich structures is their susceptibility to localized manufacturing defects and/or impact damages. As a result, a multitude of damage mechanisms such as skin/core debonding, core crushing and shear failure, in addition to skin delamination, matrix cracking, and fiber breakage, could occur. Studies on possible modes of failure and impact response of sandwich structures are quite numerous, e.g. [5] and [6].

In general, damage can be fatal for the sandwich structure, especially when it appears in invisible zones or is covered by paint. For trains, low velocity impacts generated by ballast projection are common. Ground, crosswind, and atmospheric conditions affect ballast displacement. In extreme climatic conditions, foreign objects such as ice can fall off the train wheelset and lead to ballast motions. Statistical studies revealed that for an average daily temperature of -3 °C, and 3 cm of snow falling, ice bulk can range between 5 and 10 kg [7]. As the train travels in cold and freezing climates, ice can build up in the covered region of the wheelset. When a train goes through a tunnel, the ambient temperature changes, causing the ice bulk to melt and fall off the wheelset, generating a ballast projection. The French rail company (SNCF) has conducted some experiments to simulate the falling of ice from a moving train [8].

The ballast is a stone granulate with an average weight of 60 g; it can reach the bottom of a railcar floor, therefore affecting the structure. Thus, factories require technical specifications to design and build train structures that are resistant to low velocity impacts [9] and [10]. The NF F07-101 standard defines foam-based composite damage resistance for these types of impacts [11]. It allows classification of materials according to their impact resistance and therefore their class protection.

In order to understand and predict this kind of damage, it is important to reproduce real impact tests that conform to the standard. This paper presents the impact parameters, technical specifications of the test bench, and selected materials used in this study. Results of original impact tests to predict damage on face sheet and foam materials, as well as on sandwich structures, are also presented. In order to allow a better understanding of the role of the core materials a finite element model was developed to simulate the three-dimensional state of stress during successive low velocity impacts.

2. Materials and manufacturing

The selected materials are composite/foam sandwich panels with two different configurations. Their properties are presented in Table 1. The closed-cell foam type used for both materials significantly improved their compressive strength and absorption capacity [12]. The glass transition temperature was between 68 °C and 114 °C for PET, and was 224 °C for PES [13].

The first material, Comp1, consists of glass fiber epoxy composite skins with a PES foam core. The skins were made by stacking sequences of 24 plies of unidirectional glass fibers to make a quasi-isotropic laminate with an epoxy resin. They are the same for both sides of the sandwich panel, with an overall midplane symmetry. The second material, Comp2, is made of glass fiber acrylic composite skins with a PET foam core. The skins were made by stacking sequences of 20 plies of unidirectional glass fibers with an acrylic resin. They are the same for both sides of the sandwich panel and preserve an overall midplane symmetry.

A different manufacturing process was used for both materials. Comp1 was manufactured by simultaneously infusing epoxy resin from two borders. For Comp2, acrylic resin was infused from only one side. To achieve infusion, the core structure was perforated in order to allow the resin to pass to the other side. The perforations had a 3 mm diameter and were spaced 50 mm apart. This process created visual difference without significantly influencing the density, thickness, or mechanical properties of the skins.

Panels measuring 610 × 70 × 1220 mm were created through the thermal bonding of 55 mm-wide sandwich panels. The panels were then cut to obtain new dimensions of 200 × 1220 × 70 mm.

3. Experimental testing

3.1. Low velocity test conditions

At climatic changes or vibrations, ice formed on the wheelset falls off, generating ballast projection (Fig. 1). This type of collision is very complex. To estimate the relative impact speed, this phenomenon was divided into two parts: (a) the impact between the ice and ballast layer and (b) the ballast taking off in the direction of the train floor. Only some of these ballasts can reach the train structure due to the height requirement of 33 cm. Using the law of conservation of energy, and overlooking all non-conservative forces, the speed of the ice impacting the ballast layer can be considered equal to the speed of the train [14]. Knowing the train speed and the ice weight, it is possible to estimate the number and the height of the projected ballast after a collision between the ice and the ballast layer. This analytical method, developed in [14] suggests that the projected ballast will have the same initial launch speed for 3 kg of ice.

Kawashima et al. [15] evaluated the ice/ballast impact phenomenon using an experimental approach. They confirmed that ballast took off the layer starting from an ice speed of 80 km/h (22.2 m/s). Nevertheless, only 25% of the projected ballasts reached the necessary height to damage the structure of the train. In the current case study, the train speed reached 120 km/h. According to tests carried out in [15], the angle α of the ballast path, from the horizontal axis, is situated between 0° and 70°. If we consider that the speed of the ballast is equal to the speed of the train, we obtain probable values ranging from 0° to 35° (Fig. 1b). Additionally, the NF F07-101 standard used to describe the impact test method for simulating ballast projections recommends the same values [11]. Moreover, it requires a minimum number of four successive launches at the same zone for each test case. As for recommended test temperatures, two operating conditions were considered: − 25 °C and 20 °C.

For this case study, we developed a test bench in order to simulate low velocity impacts on sandwich panels using a gas gun (Fig. 2a). A metallic projectile was fired by releasing a large volume of air in a relatively short period of time. The pressure difference, before and after the jet, created a sufficient force that propelled the projectile with a modifiable speed. In order to respect standards, projectile zero acceleration was performed by depressurizing the system on the last part of the cannon. Projectile speed was controlled before impact and, in order to minimize losses due to friction with air, the distance between the cannon and the sample did not exceed 25 mm. The projectile weighed 60 g and was manufactured by machining from a 45SiCrMo6 steel bar following specifications given in Fig. 2b.

The instrument was calibrated before starting each experiment. Therefore, several launches at different pressure values were conducted. Measurements were compared with the analytical results for different friction coefficients. Hence, good reproducibility was achieved and the results were in better agreement with theory.

Depending on the impact energy, different protection classes can be defined using the standard. In our study, we considered the K4 class corresponding to an energy of 35 J and a projectile velocity of 34.2 m/s. This speed corresponds to a nominal pressure of 0.7 bar. Table 2 presents the different test conditions considered in this study.

3.2. Composite skins and core mechanical properties

In order to identify the mechanical properties of sandwich panels, a number of experimental characterizations of skins and foam were achieved. For skins, biaxial tensile, interlaminar shear and three-point bending tests were performed. All tests were normalized and the results are shown in Table 3. The difference in number of plies and stacking sequences for each sandwich skin explains the superiority of mechanical properties identified for Comp1 compared to Comp2 [16].

For foam characterization, we performed compression tests on samples cut out from the sandwich panel. Tests were completed at 20 °C with a loading speed of 4 mm/min and an unloading speed of 20 mm/min. Table 4 presents elastic modulus and compressive strength limits for both materials in three directions (thickness H, longitudinal L, and transverse T).

Comp1 foam material is anisotropic: elastic modulus and compressive strength limit values in the thickness and transverse directions are quite different. However, Comp2 foam is transversely isotropic: longitudinal and transverse elastic modulus and compressive strength limit values are similar. If our experimental measurements are compared with data provided by material suppliers, one can notice an increase in those values. The manufacturing process used to produce the panels explains this difference: for Comp2, the presence of holes perforated to canalize acrylic resin was identified; and for both materials, the presence of interfaces formed by thermal bonding was inspected. Elsewhere, those interfaces placed in the normal direction expounded the higher values of Young's modulus and compressive strength limits in thickness direction.

For a uniaxial compression test, materials evolved through three phases: an initial linear elastic phase, a stress plateau, and a final densification stage, characterized by a steep increase in stiffness. Both core materials were made using closed-cell foam; depending on the level of compression, they exhibited either a stiffening or softening by compression, or a perfectly plastic behavior [17].

Fig. 3 shows that both foams have a perfectly plastic behavior in the thickness direction. However, in the longitudinal and transverse directions, compression resulted in stiffening. A peak load for the samples was identified at the transition between the linear phase and the beginning of the plateau during compression in the thickness direction. The appearance of this peak is much like the behavior of a honeycomb: it corresponds to the local buckling of the vertical edges of hexagonal cells [18]. Table 4 shows that, in the thickness directions, the mechanical properties of PES and PET closed-cell foams are higher. The honeycomb structure analogy explains the difference in structure of our composite foams in thickness direction. Macroscopic analysis after impact primarily allowed us to observe the panel damage.

3.3. Test results and damage inspection

In order to quantify the impact energy, an in-situ high-speed camera (Olympus i- speed III) was used. After impact, the projectile deceleration was estimated to 15.104 m/s2, and it presented an impact force of around 9000 N for a contact period of 0.8 ms (Fig. 4). The measured projectile rebounding velocity was 12 m/s. The energy transferred to the samples after impact was evaluated to 31.5 J, which represents 90% of the initial energy of 35 J.

Absorbed energy is defined by the difference between total energy at the end of an impact and energy at maximum impact force. This energy is mainly absorbed by elastic/plastic deformations and different damage modes [19]. It is divided between skins and foams of sandwich structure according to the materials used for the panel configuration [20].

Four impact tests following experimental conditions shown in Table 2 were conducted for each sample. First, a comparative analysis based on visual inspections between Comp1 and Comp2 sandwich materials during and after four successive impacts was performed. From the first projectile impact, four common damage modes were identified: delamination, matrix cracks, fiber breakage and indentation.

Delamination identified in view of whitening appears on the surface of the sandwich panel. It is common for all glass fiber reinforced composites [5]. The impacted zone has an elliptic form; two parameters were defined to delimit this zone: adel and bdel, corresponding to the major axis and minor axis. For all samples, it has been verified that the delamination zone expanded uniformly with each successive projectile impact. At the same time, the elliptic axis ratio remained constant.

Fig. 5 shows the results of the fourth projectile impact experiment, conducted for all test cases and for both materials, Comp1 and Comp2. Similarities in the damage shape and area were observed for the same impact angle: normal impact (cases I and III) and oblique impact (cases II and IV).

For the first test case (34.2 m/s; + 20 °C; 0°) and for both materials, an indentation depth of 3 mm was measured.

For the second test case (34.2 m/s; + 20 °C; 35°), a fiber pull-out belonging to the top layer was identified. It was oriented at − 45° for Comp1 and 0° for Comp2. The impact angle of 35° generated an asymmetrical shape of the delaminated zone. Therefore, this elliptical shape is a direct result of the stacking sequence used for each material. Unlike Comp1, the top of Comp2 sandwich panel corresponded to resin injection face. Consequently, impacts on the bottom of the panel generated another damage mode. The de-cohesion of top plies is the direct result of poor adhesion caused by an insufficient amount of resin during manufacturing and accented by a dynamic local buckling at impact. Thus, the tangential component of the impact force was quite large, which explains the relatively small size of the delaminated area for this material.

For the third test case (34.2 m/s; − 25 °C; 0°) and for both materials, a 3 mm indentation depth was measured. The same value was measured in the first test case. However, for Comp1, shear cracks appeared in the foam samples tested. The shape of these cracks is symmetrical with respect to the direction of the impact and the present trapezoid profile with a small base oriented to the impacted face. This shear spreads over the entire thickness of the sandwich, for a length of 400 mm (Fig. 6). This failure mode was induced essentially by thermal effect and mechanical behavior of Comp1 foam was negatively influenced by cold temperatures.

For the fourth test case (34.2 m/s; − 25 °C; 35°), the temperature decrease intensified the damage without generating an additional failure mode.

Overall, for Comp1, an analysis of the damaged surfaces during successive projectile launches revealed a polynomial evolution for normal impact (cases I and III) and a linear evolution for the inclined impact (cases II and IV) (Fig. 7). The failure size and propagation due to delamination are very similar for this panel, even with a significant temperature gap. Therefore, delamination phenomena were insufficiently influenced by this important difference in temperature. However, for Comp2, the damaged surface evolution was polynomial for all the test conditions and colder sample temperatures decreased these damage phenomena.

In order to better understand the damage mechanism generated by low velocity impacts, a macroscopic and microscopic analysis of the Comp1 material was performed. First, experiments were conducted on Comp1 panels, with a dimension of 440 × 200 × 81 mm, under the conditions of the first test conditions: 34.2 m/s, 20 °C, 0°. Two test architectures were distinguished: the first sample, denoted A1, tightened against the board edges and the second sample, denoted A2, tightened by a rigid support placed behind the panel. The first, unlike the second, allowed flexion movement. From the impact point, material A1 was cut with a 3 mm thick diamond blade, which generated eight new samples of 5 × 200 × 81 mm, four for each architecture.

When cutting was performed in the direction of the skin plies oriented at 90°, delamination failure was identified at 45°/90° interfaces. Fig. 8 shows that delamination sizes increased during successive launches for all 45°/90° interfaces, except the one closest to the foam. In fact, delamination size remained constant; regardless of the number of projectile launches or of the experimental architecture. From the impact point, within a circle radius 5 mm, no delamination failure was identified. This phenomenon was a direct result of compressive stresses created during the impact of sandwich composite by projectile launching force, and compressive strength of the foam. From a known threshold, impact force is not able to initiate delamination failure. Hou et al. [21] have discussed this phenomenon in detail.

Fig. 9 gives an overview of post-test results with cross-sections through impact points for the Comp1 sandwich composite. Micrographs were taken at representative positions directly under the impact point in order to analyze the number and extant of foam cracks, skin delamination interfaces, and matrix cracks, which appeared to be the main damage modes. For visualization reasons, different colors highlight foam cracks and both skin damage modes.

Cross-sections illustrate the damage modes previously identified for Comp1 skins: delamination, matrix cracks, fiber breakage, and indentation. By increasing the impact energy for low velocity impacts, the composite laminate reacted in this sequence [22]. The number of matrix cracks was very large and increased with successive launches. The lateral expansion of delamination and the local matrix cracks became increasingly occurring damage modes, absorbing the impact energy. Matrix cracking, the first damage mechanism introduced upon impact, was responsible for the initiation of delamination and fiber breakage, even though it did not significantly change the stiffness of the composite [23]. In addition, Olsson [24] showed that merging matrix cracks at interfaces with high interlaminar shear stresses initiate delamination from the impact point.

Damage modes in sandwich materials are different from those in laminated composites. The introduction of foam and the presence of a second skin change the energy propagation upon impact [25]. Fig. 9 shows the presence of foam cavitations (colored areas), which occurred because of successive projectile launches. Those cavitations are the result of the plastic behavior of the PES foam, as opposed to the elastic behavior of the glass fiber/epoxy skins. In fact, when the projectile impacted the sandwich panel, the skin flexed, compressing the foam. A 3 mm indentation depth at the impact point for the first architecture, A1, was measured. Immediately after impact, tensile stresses occurred, in a small area of foam under the projectile, due to a rising up of the skin. The skin/core interface did not crack in view of the perfect adhesion between those sandwich panel elements. Consequently, a failure occurred in the foam at a small distance from this interface, forming cavitation. Compressive stress generated by successive projectile impacts widened the cavitation in a perpendicular plane of impact. Foam residual stresses initiate this failure and rapid skin recovery after projectile rebounding causes it to spread [26] and [27]. This case study confirmed cavitation appearances, unlike static indentation, which only appears in compression zones of sandwich foam-cores [28] and [29].

In order to understand the damage influence of the sample boundary conditions, Fig. 9 presents four successive launch results for both experimental conditions, A1 and A2. By comparing the two damages, it is clear that the cavitation size was more significant in the case of impact with a rigid support, A2. Therefore, cavitation width largely exceeds skin delamination size. In the case of the first experimental architecture, A1, the impact energy is devised into contact energy, generating local damages, and bending energy. However, for A2, the impact energy only transformed into contact energy, creating a larger defect. In addition, this phenomenon was explained by the fact that the compressed foam lost tensile strength [30].

4. Modeling and numerical simulations

A numerical study was performed to identify mechanical stress evolution during successive impact tests on Comp1 sandwich composite. The commercial finite element (FE) code Abaqus 6.13 was used. The composite material was modeled with 8-node continuum 3D elements of the type C3D8R, with reduced integration and hourglass control. These solid elements were utilized to model the [−45/90/45/0]6 laminated skins and sandwich core foam. The laminated skins were modeled by five solid elements in the thickness direction with homogenized properties. The crushable foam model with volumetric hardening, available in Abaqus, simulated the inelastic behavior of the foam, whereas the hardening behavior was presented in terms of uniaxial compressive stress versus plastic strain. A skin/core interface was set up with a fixed boundary condition. Moreover, no failure criteria for skins and foam core were introduced, and no cohesive elements for delamination interfaces were used. Many studies are available in the literature and present finite element modeling of damages induced by low velocity impacts on sandwich composites, e.g. [31], [32] and [33].

Two numerical models corresponding to both boundary conditions, A1 and A2, were defined. Projectiles were modeled as a rigid body with a lumped mass of 60 g and an impact velocity of 34.2 m/s. Fig. 10 presents the finite element model used to predict the successive impact behaviors.

Table 5 presents FE predictions of maximum values of foam equivalent to plastic strain PEEQ. Starting from the second projectile launch, significant PEEQ differences appeared between both boundary conditions. PEEQ values during and at the end of the analysis explained the forming and propagation of cavitation. In fact, sandwich panel bending was permitted for the first architecture, A1: the impact energy divided into contact energy, generating foam crushing, and bending energy, generating bottom skin flexion. Indeed, the bottom skin of the second architecture blocked the projectile, while the core foam and top skin absorbed all the energy. Considering that compressed foam lost tensile strength, the experiments confirmed those results and explained the appearance and expansion of defects.

Equivalent plastic strain distribution showed that maximum values appeared near to the Skin/Core interface (Fig. 11). The cross-section view of post-test sandwich samples showed foam failure cavitations at a small distance from this interface. Residual stresses initiated and expanded this failure at projectile rebounding and skin recovery.

5. Conclusions

This paper presents a test bench developed to simulate low velocity impact behavior of ballast projectiles on sandwich composite materials. Considering standard rail, four impact conditions were tested to study the damage resistance of two different materials. For both materials, four successive perpendicular impacts generated a 3 mm indentation depth, fiber ruptures, and delamination damage. In the same experimental conditions, oblique impacts generated delamination and fiber blowout damages due to the tangential component of launch force. Cooling Comp1 sandwich panels to −25 °C before impact negatively affected the global mechanical behavior of the material. After four perpendicular impacts, shears appeared in the foam core samples.

Micrographs proved delamination and matrix cracks, which appeared to be the main damage modes. Successive impacts caused interlaminar damage by matrix cracking, especially in −45°/90° interfaces. Consequently, delamination propagated from the skin surface in conical form and stopped at the foam's closest interface. Macroscopic analyses of impacted samples with different boundary conditions revealed the presence of cavitation in the foam at a small distance below the impact point. This damage mode is one of the most critical for sandwich composite structures. It has been proven that successive projectile launches widened foam cavitation. In addition, this defect intensified in the case of blocked sandwich samples, eliminating bending energy. The distribution of the equivalent plastic strain in the foam core by numerical simulations confirmed these results.

Results presented in this paper will be used in future studies to identify static and dynamic bending stress damages that could appear on impacted sandwich composites.

Acknowledgments

The authors are thankful to the Conseil Régional de Picardie for supporting this project.

References

1. A. Zinno, E. Fusco, A. Prota, G. Manfredi, Multiscale approach for the design of composite sandwich structures for train application, Compos. Struct. 92 (2010) 2208–2219.
2. G. Belingardi, M.P. Cavatorta, R. Duella, Material characterization of a composite–foam sandwich for the front structure of a high speed train, Compos. Struct. 61 (2003) 13–25.
3. J.S. Kim, S.J. Lee, K.B. Shin, Manufacturing and structural safety evaluation of a composite train carbody, Compos. Struct. 78 (2007) 468–476.
4. Savey C. Superstructure for railway or road vehicles, European Patent EP0306584 A1, Mars 1989.
5. S. Abrate, Localized impact on sandwich structures with laminated facings, Appl. Mech. Rev. 50 (1997) 70–82.
6. S. Abrate, Impact on composite structures, Cambridge University Press, Cambridge, 1998.
7. S. Iikura, K. Kawashima, T. Endo, T. Fujii, Evaluation of snow accretions under the car body using the digital pictures, RTRI Report Aout, 16 2002, p. 8.
8. G. Saussine, E. Allain, A. Vaillant, M. Ribourg, O. Neel, High speed in extreme conditions: ballast projection phenomenon, International Workshop on Train Aerodynamics, Birmingham, 2013.
9. RailCorp. Xplorer/endeavour refurbishment, Part E — technical requirements, 2006.
10. State Rail Authority of NSW, Part E — Technical Specifications, Annual Report, 2001.
11. AFNOR NF F07–101: Railway applications — shock test by throwing up ballast simulation, Mars 2002.
12. U.K. Chakravarty, An investigation on the dynamic response of polymeric, metallic, and biomaterial foams, Compos. Struct. 92 (2010) 2339–2344.
13. J.E. Mark, Polymer Data Handbook. s.l. édition internet Oxford University Press, 1999.
14. Nistor DA, Caractérisation du comportement de nouveaux matériaux composites sous sollicitations d'impact type ballastPhD Thesis UTC, October 2010.
15. K. Kawashima, S. Iikura, T. Endo, T. Fujii, Experimental studies on ballast-flying phenomenon caused by dropping of accreted snow/ice from high-speed trains, RTRI Report, 17 2003, pp. 31–36.
16. Sawadogo CAH. Comportement en fatigue des compositesmonolithiques et sandwichs: détection et suivi de l'endommagement par techniques non destructives. PhD thesis, Université de Lille1; December 2009.
17. E.A. Flores-Johnson, Q.M. Li, R.A.W. Mines, Degradation of elastic modulus of progressively-crushable foams, J. Cell. Plast. 44 (2008) 415–434.
18. Y. Aminanda, B. Castanié, J.J. Barrau, P. Thevenet, Modélisation de l'indentation des structures sandwichs à peaux métalliques, Mécanique & Industires, 6 2005, pp. 487–498.
19. M.V. Hosur, A.A. Mohammed, S. Zainuddin, S. Jeelani, Processing of nanoclay filled sandwich composites and their response to low-velocity impact loading, Compos. Struct. 82 (2008) 101–116.
20. L. Aktay, A.F. Johnson, M. Holzapfel, Prediction of impact damage on sandwich composite panels, Comput. Mater. Sci. 32 (2005) 252–260.
21. J.P. Hou, N. Petrinic, C. Ruiz, S.R. Hallet, Prediction of impact damage in composite plates, Compos. Sci. Technol. 60 (2000) 273–281.
22. M.V. Donadon, L. Iannucci, B.G. Falzon, J.M. Hodgkinson, S.F.M. Almeida, A progressive failure model for composite laminates subjected to low velocity impact damage, Comput. Struct. 86 (2008) 1232–1252.
23. G.A. Schoeppner, S. Abrate, Delamination threshold loads for low velocity impact on composite laminates, Compos. A: Appl. Sci. Manuf. 31 (2000) 903–915.
24. R. Olsson, Analytical prediction of large mass impact damage in composite laminates, Compos. A: Appl. Sci. Manuf. 32 (2001) 1207–1215.
25. C.L. Wu, C.T. Sun, Low velocity impact damage in composite sandwich beams, Compos. Struct. 34 (1996) 21–27.
26. V. Koissin, V. Skvortsov, A. Shipsha, Stability of the face layer of sandwich beamswith sub-interface damage in the foamcore, Compos. Struct. 78 (2007) 507–518.
27. G.J. Dvorak, A.P. Suvorov, Protection of sandwich plates from low velocity impact, J. Compos. Mater. 40 (2006) 1317–1331.
28. V. Koissin, A. Shipsha, Residual dent in locally loaded foam core sandwich structures — analysis and use for NDI, Compos. Sci. Technol. 68 (2008) 57–74.
29. V. Koissin, A. Shipsha, V. Skvortsov, Compression strength of sandwich panels with sub-interface damage in the foam core, Compos. Sci. Technol. 69 (2009) 2231–2240.
30. D. Zenkert, A. Shipsha, K. Persson, Static indentation and unloading response of sandwich beams, Compos. Part B 35 (2004) 511–522.
31. D. Feng, F. Aymerich, Finite element modelling of damage induced by low-velocity impact on composite laminates, Compos. Struct. 108 (2014) 161–171.
32. D. Feng, F. Aymerich, Damage prediction in composite sandwich panels subjected to low-velocity impact, Compos. A: Appl. Sci. Manuf. 52 (2013) 12–22.
33. T. Besant, G.A.O. Davies, D. Hitchings, Finite element modelling of low velocity impact of composite sandwich panels, Compos. A: Appl. Sci. Manuf. 32 (2001) 1189–1196.

A Meshless Solution of Two Dimensional Multiphase Flow in Porous Media

Karel Kovářík [1,*], Soňa Masarovičová [1], Juraj Mužík [1], Dana Sitányiová [1]

[1] University of Žilina, Faculty of Civil Engineering, Department of Geotechnics, Univerzitná 8215/1, 010 26 Žilina, Slovakia

* Corresponding Author. E-mail: kovarik@fstav.uniza.sk

Abstract. Multiphase fluid flow problems are of importance in many disciplines including hydrology and petroleum reservoir engineering. Standard methods such as the finite differences, finite volumes and expanded mixed finite elements methods use very general unstructured grids and need different grid adaptation strategies to ensure optimal solution of this non-linear problem. The meshless methods seem to be quite a good alternative to these classical mesh-based methods. In our work we used the meshless Petrov–Galerkin local method based on the pressure-saturation formulation.

Keywords: Solute transfer; Multiphase flow; Meshless method; Radial basis functions.

1. Introduction

A multiphase flow appears mainly in problems related to the environment and the energy. This paper is focused on the modelling of two-phase flow, for example the flow of a wetting phase like the groundwater and a non-wetting phase like dense non-aqueous liquids through the porous medium. The problem is non-linear and therefore the simulation requires usually large meshes and also too much computational time even for simulations of testing examples. Typical numerical methods used to solve these problems are based on different formulations of the finite differences, volume and element methods [1] and [2] or the discontinuous Galerkin method [3] and [4].

Meshless methods are widely used in the last decades due to their flexibility in solving various boundary value problems and possibility to reduce a problem with generation of different meshes. Therefore these methods are considered as a powerful approach to solve partial differential equations of various kinds. The large number of meshless methods have been developed by different authors as several types of least square collocation meshless method [5], [6], [7], [8] and [9], meshless local Petrov–Galerkin method (MLPG) [10] and [11], local boundary integral element method (LBIEM) [11], radial basis integral equation method (RBIEM) [12], etc. The least square collocation methods require no integration but they have deficiency with formulation of boundary conditions and singularities as pumping wells. The MLPG, LBIEM, and RBIEM are the local weak methods and they can easy deal with different boundary conditions but evaluation of integrals is needed.

In our paper we try to present a meshless numerical method based on local Petrov–Galerkin formulation (MLPG). This method uses a local symmetric weak form to solve the problem of multiphase flow. The most important advantages of this method are simple computation of all needed integrals as they are regular and also very easy setting of boundary conditions of the second kind. This property results from the weak formulation of the solved problem.

The MLPG method has been introduced by Atluri et al. [10] and [11]. It is characterized as meshless since distributed nodal points, covering the domain, are employed. These nodal points can be randomly spread over the domain but it is well-known that using completely randomly distributed nodes may lead to less accurate results [13]. Therefore a certain effort should be invested into the positioning of the points or more sophisticated algorithms for selection neighbourhood nodes used for interpolation can be also used [14]. All needed integrals are carried out on the local subdomain centred at every point. All unknown variables are approximated by some interpolation method to obtain a system of non-linear equations. Solving this system of equations leads to a numerical solution of the problem. Atluri et al. [10] used the moving least squares (MLS) approximation scheme but nowadays the radial basis functions (RBFs) interpolation can be used instead (see e.g. [15] and [16]). An important advantage of RBF interpolation is an existence of the delta property and therefore the boundary conditions of the first kind can be easily defined.

In this paper the solution of two-phase flow through porous medium based on the MLPG-RBF method is presented. Our main goal was to investigate the robustness and the ability of this meshless method to solve this non-linear and heterogeneous problems.

2. Governing equations of multiphase flow

A two-phase flow through porous media can be usually described by the mass balance equation and Darcy's law for each of the fluid phases (see also [3])

equation(1)

View the MathML source∂(ραϕSα)∂t+∇.(ρανα)=ραfα

equation(2)

View the MathML sourcevα=−Kkrαμα(∇pα−ραg)

where α=wα=w is the wetting phase (e.g. water), α=nα=n is the non-wetting phase (e.g. oil or air), ϕ and K are the porosity and the absolute permeability of the porous media. Sα,pα,vα,ρα,μαSα,pα,vα,ρα,μα and krαkrα are, respectively, the saturation, pressure, volumetric velocity, density, viscosity and the relative permeability of the α-phase.

In addition to Eqs. (1) and (2) the following relations should be also fulfilled

equation(3)

View the MathML sourceSw+Sn=1pc=pn−pw

where pc is the capillary pressure.

To simplify the Darcy's law equation (2) we define phase mobilities

equation(4)

View the MathML sourceλα=krαμα,α=w,n

and (2) can be written as

equation(5)

vα=−λαK(∇pα−ραg)vα=−λαK(∇pα−ραg)

We have focused on the incompressible fluid flow, i. e. the densities ραρα are constant. Furthermore, we assume that the porosity ϕ remains constant over the whole domain of interest. Then Eq. (1) can be simplified

equation(6)

View the MathML sourceϕ∂Sα∂t+∇.vα=fα

The mass balance equation (6) and the Darcy's law (5) is the basis to description the multiphase incompressible flow. The pressure and saturation can be coupled using (3). Several formulations of the two-phase flow problem are possible (see e.g. [3] or [17]). In our paper we focused on the formulation based on the saturation and pressure of the wetting phase (Sw and pw).

equation(7)

−∇[λtK∇pw+λnK∇pc−(ρw+ρn)g]=fw+fn−∇[λtK∇pw+λnK∇pc−(ρw+ρn)g]=fw+fn

View the MathML sourceϕ∂Sw∂t−∇(λwK∇pw−ρwg)=fw

where λt is the total mobility defined as λt=λw+λnλt=λw+λn. These two equations are coupled because the mobilities and the capillary pressure are functions of the effective wetting phase saturation Se defined as

equation(8)

View the MathML sourceSe=Sw−Srw1−Srw−Srn

where Srw and Srn are residual wetting and non-wetting phase saturations. In our paper the Brooks–Corey model [18] is considered

equation(9)

View the MathML sourcekrw=Se2+3mmkrn=(1−Se)2(1−Se2+mm)

where m is the dimensionless pore size distribution index. The capillary pressure is defined as

equation(10)

View the MathML sourcepc=pdSe−1m

where pd is a constant entry pressure.

3. Meshless local Petrov–Galerkin formulation of the problem

The entire domain Ω is covered by nodes located inside the area and also on the global boundary Γ (see Fig. 1). The local weak formulation of the multiphase flow is formulated over a local sub-domain, created around every node. This sub-domain can be any simple geometry (rectangular or circle in 2D).

The mutual relationship of particular nodes is based on some interpolation algorithm. The local radial basis functions (RBFs) are used to approximate unknown pressure and saturation of the wetting phase pw and Sw in the neighbourhood or support of a reference point i. Multiquadric functions are one of the most popular radial functions used for this purpose and they have been used in our paper. They can be defined as

equation(11)

View the MathML sourceR(rij)=rij2+ϵ2

where rij is a distance between points i and j and ϵ is the so-called shape factor of multiquadric function. The formula of Hardy [19] with a slight modification is applied to the local RBFs (see [15]) to find the optimal value of the shape factor, which can be computed in point i as

equation(12)

View the MathML sourceϵ=0.815N∑j=1Nrij

The interpolation of the unknown variables can be written using the basis functions φij in the form (more details can be found in e.g. [16] or [15])

equation(13)

View the MathML sourcepwi=∑j=1N ϕ ijpwj,Swi=∑j=1N ϕ ijSwj

where N is the number of neighbourhood points. There are several different strategies to find local supporting points (see [7] and [14]). Here we used the following two strategies:–

Prescribed radius: The user can define the maximum distance from point i, and a support is created by the points inside this virtual circle. This strategy has been used for all regular meshes in this paper.

– Segmentation: The neighbourhood of the point i is divided into a given number of segments, and a support is created by the closest point in every segment. This is a useful strategy, particularly for randomly placed points [14] and it has been used in this paper for irregular distributed points.

The five supporting points for point i inside of domain Ω (including the point i itself) were used for all regular networks and 8 segments (i.e. 9 supporting points including the point i itself) were used for irregular networks using randomly distributed points in all numerical examples in our paper (Fig. 2).

To obtain a local weak form of the multiphase flow in the reference point i, (7) is weighted by test functions wij and integrated over the local sub-domain Ωi and the following equations are obtained

The sub-domain Ωi was divided into several triangles and all integrals were computed numerically using Gauss quadrature formula for triangles. As all integrals were regular the integration was straightforward.

Eqs. (14) and (15) create a system of non-linear equations in the general form of

equation(17)

View the MathML sourceF(pwi,Swi)=[Fp(pwi,Swi)FS(pwi,Swi)]=0

where pwi and Swi are the pressure and the saturation of the wetting phase in the time step i, respectively. To solve (17) we apply the Newton–Raphson algorithm [17]

equation(18)

View the MathML sourceJF(pwi,r,Swi,r)δr+1=−F(pwi,r,Swi,r)(pwi,r+1,Swi,r+1)=(pwi,r,Swi,r)+δr+1

where the subscript r means the r-th iteration and JF denotes the Jacobian of the system (17). The Jacobian can be written as

equation(19)

View the MathML sourceJF=[∂Fp∂p∂Fp∂S∂FS∂p∂FS∂S]

The iteration is stopped when the relative error is less than the prescribed value ϵ, i. e.

equation(20)

View the MathML source|δr+1(pwi,r+1,Swi,r+1)|<ϵ

The initial guess for the Newton–Raphson iteration for the time step i is chosen as

equation(21)

View the MathML source(pwi,0,Swi,0)=(pwi−1,SWi−1)

For the first time step the saturation View the MathML sourceSw0 is given as an initial condition and the initial value of the pressure View the MathML sourcepw0 is given as the solution of (14) based on the known value of View the MathML sourceSw0 (see also [17]).

4. Numerical examples

4.1. Simulation of the Buckley–Leverett 1D problem

The Buckley–Leverett (BL) problem is one of the most popular test used for two-phase flow models (see e.g. [20], [21] and [22]). This is a 1D problem describing the wetting phase (water) intrusion into a homogeneous aquifer which was initially saturated with non-wetting phase (oil). The capillary pressure is neglected. The equation has the form of a hyperbolic conservation law

equation(22)

View the MathML sourceϕ∂Sw∂t+vt∂f(Sw)∂x=0

where vt=vw+vnvt=vw+vn is the total velocity and f(Sw)=λw/λtf(Sw)=λw/λt is the fractional function.

The geometry of the problem is shown in Fig. 4. The boundary conditions for flow consist of two impermeable parts in the horizontal boundaries of the aquifer. The right vertical part of the boundary is considered as the oil production part and a constant pressure View the MathML sourcep0=1×106Pa is prescribed along it. A constant inflow of water (wetting phase) into the solved area is assumed along the opposite vertical part of the boundary with specific discharge View the MathML sourceq0=0.0201m2s−1. The relevant data of this problem are provided in Table 1. The relative permeabilities are given as

equation(23)

View the MathML sourcekrw=Se2krn=(1−Se)2,

where Se is defined using (8). We solved the BL problem using several types of points networks (see Fig. 4). Basically, five regular networks of 301×31,201×21,151×16,101×11301×31,201×21,151×16,101×11 and 61×7 points have been used. To demonstrate possibility of irregular network we present also the same example using randomly distributed 1515 points (Fig. 4). Support points for RBF interpolation have been chosen in regular networks by the prescribed maximum distance; in the random distributed points the segmentation strategy [14] has been used.

These aforementioned five regular networks have been used to define the order of MLPG method. Table 2 and Fig. 6 present comparison of absolute errors E1 for all these solutions in time intervals 300 and 600 days. The error E1 defined as

equation(24)

View the MathML sourceE1=1N∑i=1N|Swi−S˜wi|

has been applied where View the MathML sourceS˜wi are the values of saturation of water acquainted by analytical method. The order n has been computed from the definition E(h)=E0hnE(h)=E0hn, where E0 is the error for h=1h=1. Here h means the distance from the nearest neighbourhood point. It can be seen that the order of MLPG method is decreasing if the node number decreases but also it is slightly decreasing if the node number increase up to the optimum number. This interesting phenomenon is described also in [24] and it is caused probably by numerical instabilities like increasing condition number of the coefficient matrix.

A distribution of errors in the direction X is also interesting. The values of E1 are influenced mainly by errors around the saturation front as is illustrated in Fig. 7. We can also see that random distribution of points can cause higher errors. The origin of these errors consist mainly in the RBF interpolation using irregular distributed supporting points.

4.2. Five-spot problem

The five-spot problem is very often used as a test example of 2D two-phase problem (see e. g. [25], [3] and [17]). The classic five-spot configuration contains one injection well in the middle of a square domain and four production wells in the corners (Fig. 8). To compare the results with [3] we choose the same input data. Owing to symmetry of groundwater flow, only one quarter of the whole domain is used in our simulation. The injection well is situated in the left bottom corner and the production well in the right upper corner of the square area View the MathML source300×300m. The whole area is covered by 1675 regular distributed points (Fig. 8).

4.3. Two-phase flow through heterogeneous area

The last two presented examples serve as a study of two-phase flow through heterogeneous media. The first example presents the regular area covered by regular network of points; the second one is an example of irregular area with random distributed points.

4.3.1. Regular domain

The area of interest is a square 100×100 m and the regular network of 41×41 points has been used. The water is injected along the left boundary, the oil is pushed out through the right boundary, upper and lower boundaries are impermeable. The less permeable area is a rectangle View the MathML source37.5×50m inside this domain (see Fig. 14). The co-ordinates of the lower left corner of this area are x=37.5x=37.5 and y=25y=25 m.

xThe values of pressure pw are not so sensitive and no oscillations occur (see Fig. 16).

4.3.2. Irregular domain

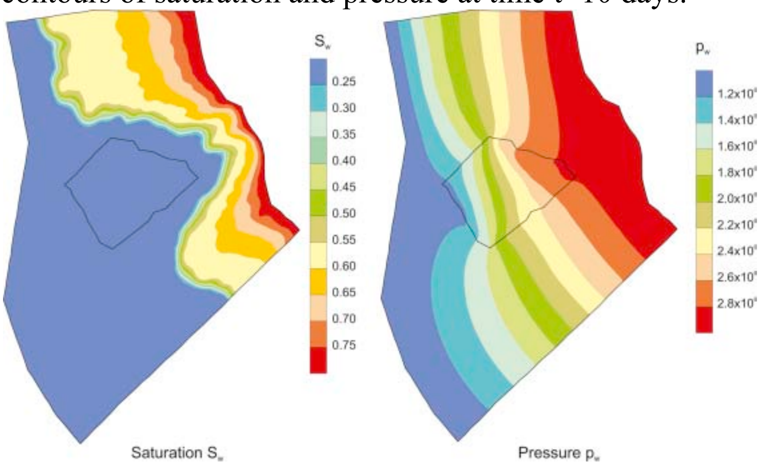
This example should demonstrate a possibility of MLPG method to solve two-phase flow through the area of arbitrary shape. The solved domain has the area of 54,801 m2. Its shape and boundary conditions are present in Fig. 18. The properties of porous media are the same as in the previous example (see Table 5). The less permeable part also has an irregular shape and it is located approximately in the centre of the domain of interest. The relative permeabilities and the capillary pressure are again described using the Brooks–Corey model (9) and (10).

The boundary conditions remain similar as in the previous example; the input pressure is prescribed constant View the MathML sourcepw=3×106Pa together with input saturation value Sw=0.8. The output pressure is also constant, View the MathML sourcepw=1×106Pa.

The whole area is covered by 1670 irregular distributed points. The support points are found using the segmented strategy (see [14]) of 9 points (including the solved point).

Results are presented in the form of contours of saturation Sw and pressure pw in Fig. 19 at time t=10t=10 days. As in the case of regular domain the contours of saturation Sw are more sensitive than those of pressure not only to the vicinity of impermeable zone but also to the shape of the input boundary condition. Some small oscillations indicate that the values of saturation are also responsive to random distribution of support points.

Irregular domain, contours of saturation and pressure at time t=10 days.



5. Conclusions

This paper presents a possible use of MLPG-RBF meshless method for modelling the two-phase groundwater flow. This method seems to be quite effective and useful to model this non-linear problem. The research is at the beginning and a follow-up study should focus on the modification of existing algorithms to enable adaptivity. H-adaptable scheme is now preparing to increase an accuracy of the method. This approach is now easily adapted to the direct collocation schemes as no evaluation of integrals is needed. H-adaptable MLPG scheme will require a new integration in newly added points and the computational time will increase. Therefore some effort should be spent to prepare H-adaptable schemes in multi-phase models.

Acknowledgments

This contribution is the result of the project implementation: “Support of the Research and Development for Centre of Excellence in Transport Engineering” (ITMS: 26220120031) supported by the Research and Development Operational Programme funded by the ERDF and Scientific Grant Agency of Slovak Republic (VEGA) No. 1-0789-12.

References

1. Bastian P, Helmig R. Efficient fully-coupled solution techniques for two-phase flow in porous media: Parallel multigrid solution and large scale computations. Adv Water Resour 1999;23:199–216.
2. Huber R, Helmig R. Node-centered finite volume discretizations for the numerical simulation of multiphase flow in heterogeneous porous media. Comput Geosci 2000;4:141–64.
3. Klieber W, Rivière B. Adaptive simulations of two-phase flow by discontinuous Galerkin methods. Comput Methods Appl Mech Eng 2006;196:404–19.
4. Epshteyn Y, Rivière B. Analysis of discontinuous Galerkin methods for incompressible two-phase flow. J Comput Appl Math 2009;225:487–509.
5. Kansa EJ. Multiquadrics—a scattered data approximation scheme with application to computational fluid dynamics. Comput Math Appl 1990;19:127–45.
6. Shu C, Ding H, Yeo KS. Local radial basis function-based differential quadrature method and its application to solve two-dimensional incompressible Navier–Stokes equations. Comput Methods Appl Mech Eng 2003;192:941–54.
7. Divo E, Kassab AJ. An efficient localized radial basis function meshless method for fluid flow and conjugate heat transfer. Trans ASME 2007;129:124–36.
8. Sanyasiraju Y, Chandhini G. Local radial basis function based gridfree scheme for unsteady incompressible viscous flows. J Comput Phys 2008;227:8922–48.
9. Kosec G, Šarler B. Local RBF collocation method for Darcy flow. Comput Model Eng Sci 2012;36:310–21.
10. Atluri SN, Zhu T. A new meshless local Petrov–Galerkin (MLPG) approach in computational mechanics. Comput Mech 1998;22:117–27.
11. Atluri SN, Kim H-G, Cho JY. A critical assessment of the truly meshless local Petrov–Galerkin (MLPG), and local boundary integral equation (LBIE) methods. Comput Mech 1999;24:348–72.
12. Popov V, Bui TT. A meshless solution to two-dimensional convection–diffusion problems. Eng Anal Bound Elem 2010;34:680–9.
13. Trobec R, Kosec G, Šterk M, Šarler B. Comparison of local weak and strong form meshless methods for 2-D diffusion equation. Eng Anal Bound Elem 2012;36:310–21.
14. Kovářík K, Mužík J. A meshless solution for two dimensional density-driven groundwater flow. Eng Anal Bound Elem 2013;37:187–96.
15. Sellountos EJ, Sequeira A. An advanced meshless LBIE/RBF method for solving two-dimensional incompressible fluid flows. Comput Mech 2008;41:617–31.
16. Kovářík K, Mužík J, Mahmood MS. A meshless solution of two dimensional unsteady flow. Eng Anal Bound Elem 2012;36:738–43.
17. Epshteyn Y, Rivière B. Fully implicit discontinuous finite element methods for two-phase flow. Appl Numer Math 2007;57:383–401.
18. Brooks RH, Corey AT. Hydraulic properties of porous media. Hydrol Pap 1964;3.
19. Hardy RL. Theory and applications of the multiquadrics-biharmonic method (20 years of discovery 1968–1988). Comput Math Appl 1990;19:163–208.
20. Helmig R. Multiphase flow and transport processes in the subsurface.New York, Berlin: Springer-Verlag; 1997.
21. Epsteyn Y. HP Primal discontinuous galerkin finite element methods for twophase flow in porous media [PhD thesis], University of Pittsburgh, 2007.
22. Hoteit H, Firoozabadi A. Numerical modeling of two-phase flow in heterogeneous permeable media with different capillarity pressures. Adv Water Resour 2008;31:56–73.
23. Laevsky YM, Kandryukova Ta. On approximation of discontinuous solutions to the Buckley–Leverett equation. Numer Anal Appl 2012;5:222–30.
24. Chandhini G, Sanyasiraju Y. Local RBF-FD solutions for steady convection–diffusion problems. Int J Numer Methods Eng 2007;72:352–78.
25. Chen CY, Meiburg E. Miscible porous media displacements in the quarter fivespot configuration. The homogeneous case, Journal of Fluid Mechanics 1998;371(Part 1):233–68.

Proposing a Framework to Combine Geological and Geotechnical Information for City Planning in Sanandaj (Iran)

Behrooz Samadian [1,*], Ali Fakher [2]

[1] Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Department of Civil Engineering, University of Tehran, Tehran, Iran

* Corresponding Author. E-mail: b_samadian@yahoo.com, b_samadian@iausdj.ac.ir

Abstract. A proposed framework that combines geological and geotechnical characteristics is very useful in planning and designing construction activities in a large area, such as a city. A step-by-step process of the proposed framework is presented for the city of Sanandaj, Iran. Sanandaj was developed on Cretaceous rock units as well as Quaternary alluviums. Identifying engineering variables related to these rock units and alluvial deposits is essential for determining their behavior in construction projects. In the presented research, first, all rock types and alluvial layers in Sanandaj were determined and their properties, including their formation and origin, were described, and then, the area was geologically classified. A database including the results of 211 geotechnical boreholes at Sanandaj was compiled and data were completed and validated through the excavation of 9 pits in different locations around Sanandaj. Finally, a geological and geotechnical framework was proposed for Sanandaj to identify the city's underlying geological layers.

Keywords: Sanandaj city soil; Urban geology; Iran; Geotechnical tests.

1. Introduction

Identifying the ground geotechnical properties in a large area by engineering tests and borehole drilling is very difficult because it is expensive and time consuming. The identification of large areas initially requires geological studies followed by the inclusion of geotechnical information. Finally, a geological and geotechnical classification is prepared for the entire area. This type of classification might not only be used in urban planning and regional development, but it might also be applied in the quick selection of geotechnical variables and initial designs in small-scale projects. Furthermore, these types of studies provide useful information on the effects of construction projects, on the possible changes in environmental conditions and on the surrounding landscape. In this paper, an example of such a classification process is presented for Sanandaj, the capital of Kurdistan province. The city of Sanandaj is situated in the west of Iran covering an area of 80 km2 and is economically important due to its vicinity to Iran–Iraq border. From 2006 to 2011, the population of the city has grown from 330,000 to 400,000. The Sanandaj site is a combination of rock and soil layers, which is an interesting and complicated site for such a study. In Fig. 1, the location of Sanandaj and its main roads are shown. Construction in this city is rapidly expanding due to population growth and urban development. Because Sanandaj is located in a large valley and surrounded on all four sides by high mountains, there is limited potential for the development of the city surface area, which has resulted in an increase in the number of buildings' floors as well as the construction of deep basements. An increase in the heights of buildings and the depths of excavations increases the importance of identifying the ground geotechnical properties of this city. To ensure that the urban development in Sanandaj is conducted with an in-depth understanding about the city's geotechnical conditions, it is very important that a geological and geotechnical classification be prepared. In this research, major geotechnical studies on Sanandaj are presented.

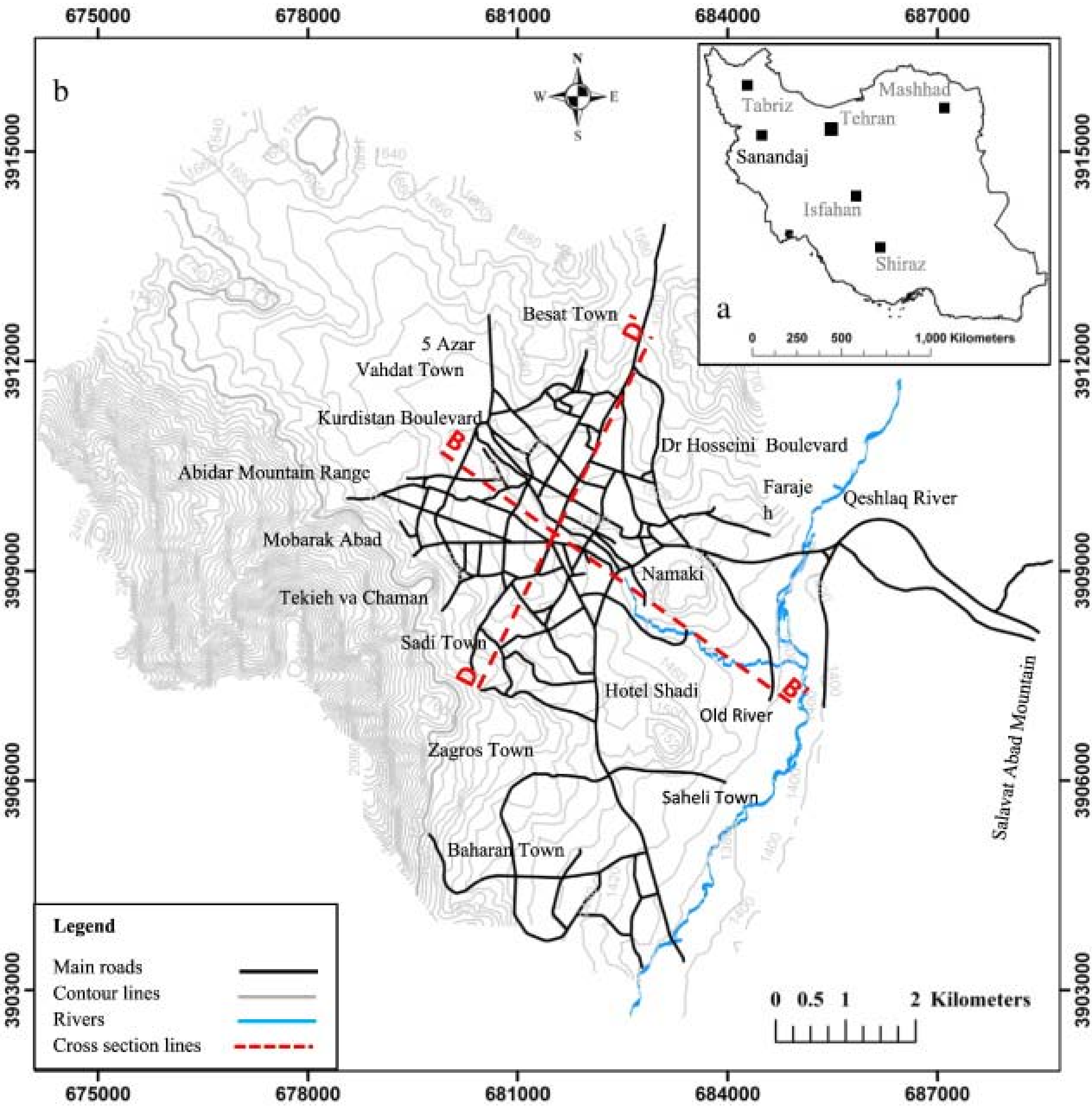


Fig. 1. a) Map of Iran, including main cities. b) Map showing the location of roads in Sanandaj and trace of the sections presented in Fig 7.

2. Similar research

Geological and geotechnical classification of urban areas has been proposed at several sites. Here we provide a brief overview on some of these researches, which were used as examples in this study. Tehran's coarse-grained alluvia have been divided into four geologic categories (Rieben, 1966). In the Rieben (1966) the geological classification system of the age and general geological characteristics of alluvia are considered. Fakher et al. (2007) added geotechnical characteristics to the Rieben geological classification system, and consequently, a new geotechnical–geological classification was presented for Tehran's soils, which is widely used in Tehran. Cheshomi et al. (2009) investigated the influence of grain shape factors on the friction angle of Tehran's soils and recommended empirical relations for Tehran's soils in various geological zones. Skipper et al. (2005) identified geological properties of the Quaternary deposits in Dublin and presented the characteristics and hazards of clay in this city. In another similar study, Maharaj (1995) combined the geological and geotechnical information to identify the physical and mechanical properties of Jamaican soil and evaluated the geotechnical hazards of Jamaican soil. Geotechnical studies on the development of Suez were conducted in Egypt by Arnous (2011) to develop the selection and construction of appropriate sites. Fuchu et al. (1994) conducted an engineering geological study on Tong Chuan in China, whereby the research, geotechnical, and geological problems of the city were investigated along with the problems induced by construction activities, such as the constructed slopes. Additionally, several urban geological-geotechnical studies have been conducted for some cities, aimed at understanding the interplay of natural processes and urban development (Haworth, 2003, Raspa et al., 2008, El May et al., 2010, Touch et al., 2014 and Jannuzzi et al., 2015).

3. Steps of the proposed framework

The proposed framework for combining geological and geotechnical information of the Sanandaj site was conducted in six steps, as presented in Table 1. These steps are described in the following sections.

3.1. Step one: studying the general geology of the considered area

The study area is located in a highly active seismic zone, not far from the Main Recent Fault in the Zagros fold and thrust belt. Floods and slope processes occur, too. Based on the Sanandaj map, as shown in Fig. 1, some parts of Sanandaj were established on both sides of an Old River, with some parts being located on the side of the Qeshlaq River. The Old Sanandaj River originated in the western highlands of the city, spanning from the west to the east of Sanandaj and finally joining the Qeshlaq River. The altitude of the central part of Sanandaj is 1450 m above sea level. The maximum height of the mountains surrounding Sanandaj is also approximately 2600 m. Iran is divided into three main sedimentary-structural zones (Stocklin, 1968): northern, central and southern Iran or Zagros. Central Iran is divided into four subzones: (1) Sanandaj–Sirjan, (2) Alborz, (3) the central Iranian micro-continent and (4) the Iranian east mountains (Nabavi, 1976, Eftekharnejad, 1980, Berberian and King, 1981 and Alavi, 1994). In Fig. 2, the subzones of Central Iran are shown. The Sanandaj–Sirjan zone is a metamorphic belt in the southwest of central Iran, which is located in a prolongation of the major Zagros thrust. The length of the Sanandaj–Sirjan zone is approximately 1500 km, and its width is 150–200 km, extending from northwest Sanandaj to southwest Sirjan in the western part of Iran (McCall, 2002). All of the rocks of the Sanandaj–Sirjan zone are located within three tectonic-stratigraphic units, namely, (a) late Precambrian–middle Triassic, (b) late Triassic–Cretaceous and (c) Tertiary set (Ghasemi and Talbot, 2006). The geological characteristics of northern-mid and south of this zone are not the same. In the northern-mid of this zone, where Sanandaj city is located, the middle Cimmerian events, especially the Laramide orogeny, are Plutonism and Metamorphism components (Aghanabati, 2010). In the late Triassic–Cretaceous, this zone was characterized by massive volcanic activities with deep Flysch–Turbiditic deposits. The late Cretaceous deposits in the Sanandaj area are primarily composed of 1700 m of dark gray shale (Sanandaj shale) with minor sandstone and micritic limestone (Sahandi, 2005). Its geomorphology is represented as large and small dome-shaped hills, in which some parts of the hills are covered with a thin layer of clay soil (Zahedi, 1990). According to Zahedi (1990), the units of the Sanandaj region are shale, limestone, volcanic rock and Quaternary alluvial deposits. Fig. 3 shows the main units of the Sanandaj region. The alluvial deposition is enhanced by the presence of the Qeshlaq River. Other factors accelerating the formation of soil layers in Sanandaj are the topography and climate conditions in the area. Flooding events and seasonal rivers rapidly produce sedimentation, forming alluvial layers with a thickness of several meters.

3.2. Step two: field reconnaissance

Many field visits were extensively conducted for this project at different times from 2012 to 2014. The apparent evidence and primary rocks observed in the Sanandaj region along with the soil layers is described in this section.

3.2.1. Rock units

The location of Sanandaj in a valley and its topography is such that all of the rock units have apparent outcrops on the walls of the valley. The units found in the Sanandaj region consist of shale, limestone and andesite, each of which is described separately.

Shale (late Cretaceous)

This rock unit is formed of quartz, feldspar and clay minerals according to XRD and microscopic studies (Sahandi, 2005). The discontinuities of the unit are sometimes high and significant on the surface. However the fractures reduce with an increase in the depth. The rock unit is highly weathered and crushed due to exposure to air and moisture. The spacing of discontinuities in this rock is usually less than 20 cm. This rock outcrops in many places of the Sanandaj region, such as the Abidar Mountain Range, Abidar–Hassan Abad Belt, Sadi Town, Zagros Town, Vahdat Town, Baharan Town, Payam Town, Besat Town, Ghopal, area behind the Shadi Hotel, Degayran, Sheikh Salam Hill and many other locations and hills in the city. Fig. 4a shows an example of the observed shale. Shale is considered to be the dominant rock in Sanandaj. Based on the observations in this study, it is estimated that 70% of the city was formed by shale rock. Loading and unloading processes during the geological periods play an important role in shaping properties of the shale. Shale expands during unloading, creating joints and cracks (Terzaghi and Peck, 1976).

Limestone (late Cretaceous)

The rock unit is gray and dark gray and consists of detrital limestone and sand limestone. The thickness of the rock unit is 200 to 300 m, and its age dates back to late Cretaceous (Zahedi, 1990). The spacing of discontinuities in this rock varies between 1 and 2 m and the separating of joints is also less than 1 mm. The limestone can be found in the mountains of the Sanandaj region, such as Abidar Mountain and its mountain range and some other locations. Fig. 4b shows an outcrop of limestone in Sanandaj. This rock is typically located above shale rock; however, it can also be found within the layers of shale. The limestone gradually created an alluvial fan at the foot of Mount Abidar. However, calcareous rubble was also found in places farther away from Mount Abidar towards the center of Sanandaj.

Andesite (late Cretaceous)

The areas in Sanandaj with the most extensive amount of Cretaceous volcanic rock outcrops include an area between Sanandaj and Divandarah (in 90 km north of Sanandaj), recognized by Zahedi (1990) as an inter-stratigraphic Cretaceous volcanic rock. Field studies indicate that volcanic rocks with a greenish-gray to dark color are located within the layers of coal shale and the light to gray limestone. In the aforementioned region, volcanic rocks have been found in large and scattered masses along a north-south direction. The thickness of this rock is approximately 300 to 400 m, which belongs to the late Cretaceous (Zahedi, 1990). Only some parts of this rock on the ground surface are weathered. The spacing of discontinuities in this rock varies between 0.5 and 2 m and the separating of joints is less than 1 mm. Fig. 4c shows an outcrop of andesite rock in Sanandaj.

3.2.2. Soil layers

Relatively deep soil layers are present in different parts of Sanandaj, especially in the axial part of the valley. Here, the soil layers are described.

Two types of clays, namely alluvial and residual, can be found in Sanandaj. The distinction is based among all on the shape of coarse aggregates, which are more rounded in alluvial clays and more angular in residual ones.

Alluvial clay

The presence of this clay in Sanandaj is the result of floods that occurred from west to east, especially in the center of Sanandaj. The thickness of this layer varies between 5 and 40 m, whereas the thickness in the central part of Sanandaj (Azadi Square) is approximately 35 m. Fig. 5a and b show the outcrops of alluvial clay in the excavation. The alluvial clay of Sanandaj has a relatively uniform gradation and a small percentage of coarse aggregates (less than 30%), with a maximum size of 2 cm. Most of the aggregates have rounded corners and/or a flaky shape. In some areas, the amount of aggregates increases and the alluvial clay is renamed from CL to SC or GC, according to the unified soil classification system. In Fig. 5c, the increased amount of aggregates in the clay soil can be seen. However, the aggregates are not in contact with each other, thus they do not play an important role in the mechanical behavior of the clay. The wandering limestone cobbles with diameters of 7 to 15 cm, which have flooded from Mount Abidar and the western mountain ranges of Sanandaj to the central part of the city, can also be seen within the clay layers. Moving farther from the surrounding mountains, the diameter of aggregates in the clay becomes smaller. The accumulation of alluvial deposits and the flooding from Sanandaj's western mountains have caused the formation of alluvial fans.

Residual clay

This clay resulted from the weathering of rocks, especially shale, and is usually seen in the western highlands and some shale hills in Sanandaj, which have been highly weathered. The thickness of residual clay usually does not exceed a few meters. The unit contains rock fragments that are 3–4 cm in size. These fragments have sharp corners, evidence of weaker water erosion in respect of alluvial clays. Fig. 6 shows an example of residual clay with rock fragments embedded in the clayish unit. Table 2 also compares two types of Sanandaj clays.

Sand gravel

Clean sand and gravel can be seen along the Qeshlaq River (Kurdistan Boulevard) and its surrounding areas as well as along the Old River with a thickness of a few meters. In the deeper part of the Sanandaj, such as Shebli Boulevard, Chamran Boulevard, South Mardoukh and Azadi Square, alluvial clay there are also sand and gravel interlayers.

3.3. Step three: understanding the subsurface layers and drawing the general profiles

During the third step, it is necessary to image the subsurface stratigraphy. Here, the results of the boreholes must be used. The borehole data include bedrock depth, groundwater level, and the physical and mechanical properties of the layers in different parts of Sanandaj City. Because Sanandaj is located in a valley, all of the units and layers were observed in the field visits and these observations were verified by borings and subsurface identifications. The boreholes used in this study are described in the following steps. Fig. 1 shows the location of the B－B′ cross section along the Old River and the D-D’ cross section perpendicular to the B－B′ cross section, passing from Sanandaj center. Additionally, Fig. 7 shows the cross sections. It must be noted that the Old River is presently dry because of drought.

3.4. Step four: geological classification based on the stages of unit formation

In this step, it is necessary to classify the geological units determined in the second and third steps based on the formation stages, presented in Table 3. Additionally, Table 4 presents the classification of the observed geological units in Sanandaj.

3.5. Step five: determining the geotechnical properties of the geological units

This step should be performed over several phases. First, the available geotechnical information should be collected, and then, the collected information should be checked. Eventually, the missing information will be obtained through field investigations.

3.5.1. Collecting the available geotechnical data

The information obtained from the previous geotechnical site investigations was collected in order to provide the data for the engineering characterization. The information was obtained from 211 boreholes at a maximum depth of 40 m. It must be noted in some areas, the spacing between boreholes in Sanandaj was a few kilometers and the depth of boreholes in some areas was less than 5 m. Therefore, one could not use only the boreholes data for geotechnical evaluation. Some limitations, including the lack of data, can be resolved through the combination of geological and geotechnical data. Fig. 8 shows the location of the boreholes from which the data was collected. The information was implemented in a geographic information system (GIS) program and a database was established.

3.5.2. Verification of the boreholes data and providing additional data

In this study, a number of pits were manually excavated with a diameter of 1 to 1.2 m and a series of tests were also performed. The purpose of preparing new pits and tests was to check and compare the previous data and to provide additional data. Nine locations were selected in different parts of Sanandaj for the excavation of pits (Fig. 8), whose characteristics are presented in Table 5. These locations were selected because of their proximity to the previously drilled boreholes, thus allowing a direct comparison. For the sake of identification, a series of tests were planned according to the

ASTM standard (ASTM D1556–00, 2000, ASTM D2435–03, 2003, ASTM D2487–00, 2000, ASTM D2850 – 03a, 2003, ASTM D3080–03, 2003, ASTM D422–63, 2002, ASTM D4318–00, 2000 and ASTM D6473–99, 1999) (Table 6). It must be noted that there was a limitation regarding the depth of the pits because of the level of the groundwater and bedrock. Additionally, a geological survey was used when deeper pits or boreholes were needed. After completion of the necessary tests, on samples taken from the tips, the collected boreholes data were recognized to be acceptable.

Rock units are strong enough to be used for urban construction. However, in previous projects in Sanandaj, characterization was sometimes overlooked, so there was not much information available on them. In the presented research, rock properties were studied via laboratory investigations. To identify the rock properties in Sanandaj, rock units were classified according to the RMR classification system (Bieniawski, 1989) and the Franklin (1981) system for rating shale rock. Accordingly, in addition to the tests for specific gravity and water absorption, durability index and point load tests were also performed (Franklin and Chandra, 1972). The point load test was selected due to its simplicity, speed and correlation with the uniaxial compressive strength test (Bieniawski, 1975). Because limestone and andesite in Sanandaj are regarded as good rocks and are not degradable, a limited number of durability index tests were conducted on them. However, shale rocks weather quickly during excavation and their strength decreases significantly; therefore, they are regarded as degradable rocks. Table 7 presents the various tests that were conducted on the rock units in Sanandaj.

3.5.3. Summarizing the geotechnical properties

According to the prepared data bank and the complementary test results, the physical and deformation properties as well as strength properties of Sanandaj soil are presented in Table 8 and Table 9, respectively. Additionally, the properties of the rock units are shown in Table 10. In this table, the physical properties of three types of rocks including specific gravity, water absorption, durability index, and RQD are shown. Moreover, the elasticity modulus, point load and uniaxial compressive strength of rocks are determined, and finally the mentioned rocks are classified based on the RMR system.

3.6. Step six: geological and geotechnical classification for the study area

As previously mentioned, the geological–geotechnical classification of an urban area could be used in planning as well as in the rapid and preliminary identification of the ground in a project. In addition to the ground geotechnical properties, this type of classification must also be able to determine the possibility of finding various units in the study area. The following summary is noteworthy in this regard.

Data collected in this research allow to recognize the factors affecting the possibility of finding certain units. These factors include: (a) the altitude level, (b) the distance from Qeshlaq River and Old River (Dareh Bayan River), and c) the distance from fault A shown in Fig. 3. These factors are described in the following section.

(i) Altitude level

Mountains surround Sanandaj with some hills within the city. The closer to these mountains (either the surrounding mountains or the hills within the city), the higher the possibility of finding rock units. Therefore, the altitude can determine the possibility of finding certain units. Additionally, residual clay can be seen at high altitude levels and in mountain ranges, especially in the vicinity of shale rock due to the weakness of these rocks.

(ii) Distance from Qeshlaq River and Old River in Sanandaj

Generally, Sanandaj soil is primarily composed of clay, especially alluvial clay. The thickness of this soil in the central part of Sanandaj, which is located along the river, is greater than in other areas. Near the Old River, some 2–3 m thick inter layers of sand can be observed. Most of the soil near the Qeshlaq River in western Sanandaj is composed of sand and gravel. Therefore, the distance from the river can also be considered a factor affecting the possibility of encountering different soil layers.

(iii) Distance from fault A

According to Fig. 3 and the description in Table 3, considering fault A, from the southeast to north of Sanandaj, the rock units east of this fault are typically andesite rock and the rock units to the west are shale or limestone. It is likely that the presence of the Salavat Abad fault, which is a few kilometers in length and runs in a north-south direction a short distance from and parallel to fault A, caused such a boundary.

(iv) Determining the rock units and layer type at a project location

The workflow in Fig. 9 shows the possibility of finding a rock unit and Sanandaj soil. From this figure, in addition to the bedrock type and rock outcrops, it is also possible to determine the soil type. Once the soil type and possible bedrock at one point is determined according to Fig. 9, its characteristics can be estimated using Table 11.

a) Determining the bedrock type at a project location:

To determine the bedrock type at a project location, refer to the workflow's right part. If the project is located on the east of fault A, the bedrock is andesite. To determine the rock's geotechnical characteristics, see the row 6 in Table 11. However, if the project is located on the west of fault A, the bedrock is limestone or shale. If the altitude of the project location is more than 1700 m, the rock type is limestone, and for its geotechnical properties, see row 5 in Table 11.However, if the altitude of the project location is less than 1700 m, the bedrock is shale, and for its geotechnical properties, see row 4 in Table 11.

b) Determination of the soil type at a project location:

To determine soil type at a project location, refer to the workflow's left part. If the altitude of the project location is more than 1700 m, the soil type is residual clay (CL or SC), and see row 1 in Table 11 for its geotechnical properties. However, if the altitude of the project location is less than 1700 m, the soil type is alluvial clay or gravel and sand. To recognize the soil type, its distance from Old River soil should be determined. If the project site is less than 50 m from Old River, the soil type is alluvial layer (CL or SC or SM).To determine its geotechnical properties, see rows 2 and 3 in Table 11. However, if the distance from Old River is more than 50 m, to determine the type of soil, the distance from Qeshlaq River should be taken into account. If the distance from Qeshlaq River is less than 50 m, the soil type is gravel and sand (SP or SM or GM). To determine its geotechnical properties see row 3 in Table 11. However, if the distance of the project site from Qeshlaq River is more than 50 m, the soil type is alluvial clay (CL or SC).Geotechnical properties of the soil type can be seen in row 2, Table 11.

4. Geotechnical zoning and hazards in Sanandaj

Using Fig. 10 and Table 11, an initial estimation for the properties of the possible layers can be made. Then, to complete the classification, the hazards are determined. Finally, the geotechnical zoning map of Sanandaj is presented and can be used in urban planning.

4.1. Soil layer hazards

Residual clay is minimally thick. As one gets closer to the mountains and hills, the thickness and age of the clay becomes smaller, and therefore, it is expected that the preconsolidation pressure in this soil would be smaller. Additionally, the residual clay in Sanandaj typically has a high water content due to its low thickness and position. Therefore, when encountering this soil type, the settlement issue should be considered. Additionally, vertical excavation in this soil type should be performed with great caution. The other type of Sanandaj clay is alluvial clay. This clay is at great distance from the mountain ranges and is thicker, older and has a larger preconsolidation pressure and smaller settlement rate in comparison with the residual clay. In terms of consistency, this type of clay is medium to stiff. However, a vertical excavation more than 5 m deep in this soil should be performed with great caution. Additionally, in a vertical excavation of a lesser depth and in the presence of water and leakage, the strength of the clay decreases with the possibility of collapse. River sand and gravel exist in the clay in the form of inter layers, but the thickness of the river sand and gravel increases in areas near the Qeshlaq River and Old River. This soil does not have a high density and is not cemented.

4.2. Rock layer hazards

In general, there are three types of outcrops in Sanandaj, shale rock, limestone and andesite rock. As presented in Table 11, according to the RMR rating system, the andesite rock and limestone have fair to good characteristics and generally do not pose geotechnical hazards during normal construction operations; however, the shale rock is considered to be among the very poor to poor rocks according to this classification. The shale easily crushes and weathers during unloading and often shows a high hazard of collapse during trenching and excavation. In the case of high weathering and decomposition of the shale rock, this rock can be considered to be a soil in the design procedure to extract and calculate its geotechnical characteristics. Based on the durability index and point load test, the diagrams proposed by Franklin (1981) and Richardson and Wiles (1990) can be used to obtain some of the geotechnical parameters of the shale rock. The geotechnical hazards of the Sanandaj layers are briefly summarized in Table 12.

4.3. Geotechnical zoning map of Sanandaj

The geotechnical zoning map of Sanandaj, prepared using the collected borehole data as well as the results of field investigation and laboratory tests conducted in this study, was presented in Fig. 10. This map can be used for preliminary geotechnical evaluation and urban planning. Based on the map, Sanandaj was divided into 6 zones. The properties of the layers including geotechnical classification, layers depth, underground water level and the geotechnical hazards of the layers were indicated in the map.

5. Discussion

The presented research shows an approach for combination of the geological and geotechnical information which can provide a good image of the geotechnical conditions in large areas such as cities. The presented approach can be applied, especially in the areas where there is not sufficient borehole data. In many urban studies, sufficient geotechnical borehole data is available. For example, Raspa et al. (2008) used more than 2000 boreholes in their investigations of Roma. Touch et al. (2014) also used 1200 boreholes to study Phnom Penh City. In the presented study, the number of the boreholes was 211 and in some parts, the spacing of the boreholes was more than 5 km so sufficient borehole data was not available. However by using the proposed approach, the geotechnical properties of the layers were identified and the geotechnical zoning map for Sanandaj site was extracted. A similar study was undertaken by Fakher et al. (2007) and Cheshomi et al. (2009) on the Tehran alluvium. They studied geotechnical properties of the alluvial deposits but the Sanandaj site is a combination of rock and alluvial deposits so it can be an interesting case study for the proposed method. The main points of the proposed approach are as follows. Firstly, the approach emphasizes on geological classification rather than geological descriptions. Secondly, a step by step framework is proposed using rational relationship between geological and geotechnical condition rather than a pure geotechnical study. In the proposed approach, all the geotechnical data are sorted according to geological classification. Therefore, the geological classification of the desired area plays an important role in such investigations. Finally, it is even possible to derive simple diagnostic criteria to determine the engineering properties of the ground at each point. Hence, the proposed approach can be optimally applied in the geotechnical identification and mapping of the cities for urban planning.

6. Conclusion

Combining geological and geotechnical information in a large area can provide a useful framework for large-scale development projects. Furthermore, such a framework can quickly provide preliminary geotechnical information in the early stages of all projects. According to the research presented in this article for the Sanandaj site, the following results can be extracted:

(1) In this article, a step-by-step approach was proposed for combining geological and geotechnical information for the Sanandaj site. The proposed stages will result in a geological and geotechnical classification for a large area.

(2) The main geological units in Sanandaj consist of Cretaceous rocks and Quaternary units. The rock units include shale rock, limestone and andesite rock. The Quaternary unit consists of alluvial clay, residual clay and fluvial sand and gravel.

(3) If a region in Sanandaj is selected for urban development or a construction project, Figs. 9 and 10 can be used to predict the type of layers that will be encountered, and then, Table 11 can be used to estimate the geotechnical properties. In the next stage, Table 12 can be used to predict the possible hazards.

(4) Table 3, Figs. 9 and 10 and Table 11 show examples of the combined geological and geotechnical framework.

References

1. Aghanabati, A., 2010. Geology of Iran. third ed. Iran Geological Survey and Mineral Exploration Organization, Tehran (606 pp., (in Persian)).
2. Alavi, M., 1994. Tectonic of the Zagros organic belt of Iran: New data and interpretations. Tectonophysics 229, 211–238.
3. Arnous, M.O., 2011. Geotechnical site investigations for possible urban extensions at Suez city, Egypt using GIS. Arab. J. Geosci. 6, 1349–1369.
4. ASTMD1556-00, 2000. Standard Test Method for Density and UnitWeight of Soil in Place by the Sand-cone Method.
5. ASTM D2435-03, 2003. Standard Test Methods for One-dimensional Consolidation properties of Soils Using Incremental Loading.
6. ASTM D2487-00, 2000. Standard Practice for Classification of Soils for Engineering Purposes.ASTM D2850-03a, 2003. Standard Test Method for Unconsolidated-Undrained Triaxial Compression Test on Cohesive Soils.
7. ASTM D3080-03, 2003. Standard Test Method for Direct Shear Test of Soils Under Consolidated-Drained Conditions.ASTM D422-63, 2002. Standard Test Method for Particle-size Analysis of Soils.
8. ASTM D4318-00, 2000. Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils.
9. ASTM D6473-99, 1999. Standard Test Method for Specific Gravity and Absorption of Rock for Erosion Control.
10. Berberian, M., King, G.C.P., 1981. Towards a paleogeography and tectonic evolution of Iran. Can. J. Earth Sci. 18, 210–265.
11. Bieniawski, Z.T., 1975. The point-load test in geotechnical practice. Eng. Geol. 9, 1–11.
12. Bieniawski, Z.T., 1989. Engineering rock mass classifications. John Wiley & Sons, New York, p. 251.
13. Cheshomi, A., Fakher, A., Jones, C.J.F.P., 2009. A correlation between friction angle and particle shape metrics in Quaternary coarse alluvia. Q. J. Eng. Geol. Hydrogeol. 42, 145–155.
14. Eftekharnejad, J., 1980. Tectonic division of Iran with respect to sedimentary basins. J. Iran. Pet. Soc. 82, 19–28 (in Persian).
15. El May, M., Dlala, M., Chenini, I., 2010. Urban geological mapping: geotechnical data analysis for rational development planning. Eng. Geol. 116, 129–138.
16. Fakher, A., Cheshomi, A., Khamechian,M., 2007. The addition of geotechnical properties to a geological classification of coarse grain alluvium a pediment zone. Q. J. Eng. Geol. Hydrogeol. 40, 163–174.
17. Franklin, J.A., 1981. A shale rating system and tentative application to shale performance. Transportation Research Record 790. TRB, National Research Council, Washington D.C, pp. 2–12.
18. Franklin, J.A., Chandra, R., 1972. The slake durability test. Int. J. Rock Mech. Min. Sci. 9, 325–341.
19. Fuchu, D., Yuhai, L., Sijing, D., 1994. Urban geology: a case study of Tong Chuan city, Shaanxi Province, China. Eng. Geol. 38, 165–175.
20. Ghasemi, A., Talbot, C.J., 2006. A new tectonic scenario for the Sanandaj–Sirjan zone (Iran). J. Asian Earth Sci. 26, 683–693.
21. Haworth, R.J., 2003. The shaping of Sydney by its urban geology. Quat. Int. 103, 41–55.
22. Jannuzzi, G.M.F., Danziger, F.A.B., Martins, I.S.M., 2015. Geological–geotechnical characterisation of Sarapuí II clay. Eng. Geol. 190, 77–86.
23. Maharaj, R.J., 1995. Engineering–geological mapping of tropical soils for land-use planning and geotechnical purposes: a case study from Jamaica, West Indies. Eng. Geol. 40, 243–286.
24. McCall, G.J.H., 2002. A summary of the geology of the Iranian Makran. In: Cliff, P.D., Kroon, D., Gaedecke, C., Craig, J. (Eds.), The Tectonics and Climatic Evolution of the Arabian Sea Region. Geological Society of London Special Publication 195, pp. 147–204.
25. Nabavi, M.H., 1976. An introduction to the geology of Iran. Geological Survey of Iran (109 pp., (in Persian)).
26. Raspa, G., Moscatelli, M., Stigliano, F., Patera, A., Marconi, F., Folle, D., Vallone, R., Mancini, M., Cavinato, G.P., Milli, S., Costa, J.F.C.L., 2008. Geotechnical characterization of the upper Pleistocene–Holocene alluvial deposits of Roma (Italy) by means of multivariate geostatistics: cross-validation results. Eng. Geol. 101, 251–268.
27. Richardson, D., Wiles, T., 1990. Shale durability rating system based on loss of shear strength. J. Geotech. Eng. 116 (12), 1864–1880.
28. Rieben, E.H., 1966. Geological observation on alluvial deposits in Northern Iran. Geol. Organ. Iran Rep. 9 (in French).
29. Sahandi, M.R., 2005. Sanandaj geological Map with scale of 1:100000, No 5460, Iran Geological Survey and Mineral Exploration Organization (in Persian).
30. Skipper, J., Follett, B., Menkiti, C.O., Long, M., Clark-Hughes, J., 2005. The engineering geology and characterization of Dublin Boulder Clay. Q. J. Eng. Geol. Hydrogeol. 38, 171–187.
31. Stocklin, J., 1968. Structural history and tectonics of Iran: a review. Am. Assoc. Pet. Geol. Bull. 52 (7), 1229–1258.
32. Terzaghi, K., Peck, R., 1976. Soil Mechanics in Engineering Practice. John Wiley & Sons, NewYork, p. 729.
33. Touch, S., Likitlersuang, S., Pipatpongsa, T., 2014. 3D geological modelling and geotechnical characteristics of Phnom Penh subsoils in Cambodi. Eng. Geol. 178, 58–69.
34. Zahedi, M., 1990. Describing Sanandaj quadrangle geologicalmapwith scale of 1: 250000. Iran Geological Survey and Mineral Exploration Organization (in Persian).

Requirements Engineering for Safety-critical Systems: A Systematic Literature Review

Luiz Eduardo G. Martins [1, *], Tony Gorschek [2]

[1] Department of Science and Technology, Federal University of São Paulo, São José dos Campos, Brazil

[2] Software Engineering Research Lab, Blekinge Institute of Technology, Karlskrona, Sweden

* Corresponding Author. E-mail: legmartins@unifesp.br

Abstract. *Context* Safety-Critical Systems (SCS) are becoming increasingly present in our society. A considerable amount of research effort has been invested into improving the SCS requirements engineering process as it is critical to the successful development of SCS and, in particular, the engineering of safety aspects. *Objective* This article aims to investigate which approaches have been proposed to elicit, model, specify and validate safety requirements in the context of SCS, as well as to what extent such approaches have been validated in industrial settings. The paper will also investigate how the usability and usefulness of the reported approaches have been explored, and to what extent they enable requirements communication among the development project/team actors in the development of SCS. *Method* We conducted a systematic literature review by selecting 151 papers published between 1983 and 2014. The research methodology to conduct the SLR was based on the guidelines proposed by Kitchenham and Biolchini. *Results* The results of this systematic review should encourage further research into the design of studies to improve the requirements engineering for SCS, particularly to enable the communication of the safety requirements among the project team actors, and the adoption of other models for hazard and accident models. The presented results point to the need for more industry-oriented studies, particularly with more participation of practitioners in the validation of new approaches. *Conclusion* The most relevant findings from this review and their implications for further research are as follows: integration between requirements engineering and safety engineering areas; dominance of the traditional approaches; early mortality of new approaches; need for industry validation; lack of evidence for the usefulness and usability of most approaches; and the lack of studies that investigate how to improve the communication process throughout the lifecycle. Based on the findings, we suggest a research agenda to the community of researchers and advices to SCS practitioners.

Keywords: Safety requirements; Safety-critical systems; Hazard; Accident; Systematic literature review; Requirements engineering.

1. Introduction

Computing systems are becoming ubiquitous and a basic part of human life. They help us in so many activities that it is difficult to imagine modern society without their support. Nevertheless, this ubiquitous presence carries a high level of dependency, which inevitably demands the need for systems that are increasingly available, reliable, safe, and secure [18,119] (for SLR references please see Appendix A). In many situations we rely on computing systems to help us control highly critical activities [81,115], such as in medical procedures [59,79,85,104,117,130,179], human transportation [48a,53,87,88,96,110,164], aerospace and defence systems [23,71,128,131,132,183], and high energy handling [142,173,182,189]. Failures during the control of these systems might cause serious damage to the environment, property and people [8], [17] and [29], impacting companies, the marketplace, as well as the quality of life of people and society in general.

Systems with these characteristics are generally called safety-critical systems (SCS) [113,124,148]. Over the last 40 years, a considerable amount of research effort has been invested into improving the engineering of SCS. One of the most significant challenges for companies that develop SCS is to create and establish a complete, correct, unambiguous, testable, and yet understandable requirements specification and/or understanding shared among stakeholders [15,64,65,92,106,126]. This is crucial for the IT mainstream [20], but it is even more important for companies developing SCS, when considering product/system safety certification process compliance [100,180]. Compliance and ultimate certification and associated processes are required for companies developing SCS [91,98].

The literature on SCS has reported on many cases where systems have failed due to a lack in requirements specifications, or misunderstandings traced to problems in requirements engineering, contributing to accidents that cause damage to the environment, injury to people and even the loss of lives [8] and [13],177]. Of course, when accidents happen, they have a strong negative impact for the companies responsible for the associated SCS. As discussed by Leveson in [16], the causes of accidents involving complex technological systems usually are multifactorial. The hierarchical model of accidents causes proposed by Lewycky [39] points out that it is the constraints, or lack of them, on technical and physical conditions, social and human interactions [60,82,93,125,146], the management system and organizational culture, as well as governmental or socioeconomic policies [16] and [24] that are the root causes of accidents. Such factors are closely related and have influenced the way of approaching the safety requirements of SCS [61,74].

With the increasing complexity of SCS [12,191], the rules and standards for safety certification and associated processes defined by governments and international agencies are becoming more difficult and expansive [25]. The requirements specifications, and related processes for requirements engineering, play a very important role during the safety certification process [26,157,159,160], both in relation to process-based compliance and safety-assurance standards [15] and [19],145,150,174]. In addition, with the system functionalities increasingly moving from hardware to software [70], the safety certification process becomes even more complex.

Considering the importance of the requirements engineering process for improving the safety of SCS [78,165,171,178,185], we conducted a systematic literature review (SLR) to investigate what approaches have been proposed to elicit, model, specify or validate safety requirements in the context of SCS, as well as to what extent such approaches have been validated in industrial settings. Furthermore, we investigate the relationship between safety analysis and requirements engineering practices in order to analyse how integrated these areas are and what communication issues emerged from them. In this paper we analyse and discuss the SLR results considering four perspectives: (i) the requirements engineering approaches to treat safety requirements; (ii) how the safety requirements approaches have been validated by their proponents; (iii) how the usefulness and usability of the approaches have been measured by their proponents; and (iv) to what extent the safety requirements approaches support communication among the actors throughout the SCS lifecycle. To the best of our knowledge, this is the first SLR on the topic of requirements engineering for SCS.

This paper is organized as follows. Section 2 presents background and related work. The research methodology adopted to conduct the SLR is presented in Section 3. The results and the analysis related to our research questions are presented in Section 4. Our conclusions are presented in Section 5.

2. Background and related work

2.1. Definitions

In order to set the scope and make clear the adopted terms used in the SLR, and to ensure consistency throughout this paper, we present the following definitions, organized in alphabetical order:

Accident. An undesirable (negative) event involving damage, loss, suffering or death [7,120].

Approach. In the context of this SLR, we are interested in the following types of approaches: technique, model, framework, method, process, methodology or tool to elicit, model, specify or validate safety requirements for safety-critical systems.

Functional safety requirement. The requirement to prevent or mitigate the effects of failures identified in safety analysis [6].

Hazard. A system state that might, under certain environmental conditions, lead to a mishap. Hence, a hazard is a potentially dangerous situation that may lead to an accident [1] and [7].

Safety. Firesmith defines safety as “the degree to which accidental harm is properly addressed (e.g. prevented, identified, reacted to, and adapted to)” [2]. According to Leveson “safety must be defined in terms of hazards or states of the system that when combined with certain environmental conditions could lead to a mishap.” [7,133].

Safety requirement. A requirement that describes the constraints or actions to support and improve a system's safety. Firesmith defines the safety requirement as “any quality requirement that specifies a minimum, mandatory amount of safety in terms of a system-specific quality criterion and a minimum level of an associated metric.” [2].

Safety-critical. According to Medikonda and Panchumarthy “those software or system operations that, if not performed, performed out of sequence, or performed incorrectly could result in improper control functions, or lack of control functions required for proper system operation, that could directly or indirectly cause or allow a hazardous condition to exist” [1].

Usability. How easy an approach is to be used by practitioners.

Usefulness. The fact of being useful and bringing value for practitioners.

Validated approach. It is one that was tested, piloted or performed in some way into the industry setting.

2.2. Related work

Nair et al. [62] conducted a SLR on evidence for safety. The study considered 171 peer-reviewed publications with the intention of answering four questions: “What constitutes the evidence for safety?”, “What techniques are used for structuring safety evidence?”, “What techniques are used for assessing safety evidence?” and “What challenges and needs have been the target of the investigation in relation to safety evidence?” The authors argue that they intentionally conducted a SLR with a broad scope, not restricting themselves to a particular standard or domain. The stated reason for such a decision is that the breadth of scope enabled them to provide a more general and thorough analysis of the state of the art on evidence for safety. Additionally, they classified the various notions of evidence gleaned from the literature into a hierarchical taxonomy, which includes 49 evidence types.

Mellado et al. [3] conducted a systematic review of the literature concerning security requirements engineering in order to summarize evidence regarding security requirements approaches and to provide a framework to appropriately support new research activities. The research question that they tried to answer was “Which approaches have been carried out to develop secure Information Systems by means of Security Requirements Engineering?” They found 22 studies that completely fit their previously defined inclusion criteria.

Rodríguez-Dapena [19] discusses software safety certification as a multi-domain challenge. This work highlights the problem of new systems that are built from subsystems, which come from different application domains, because there is no certification scheme for inter-domain systems yet. The author comments on the influence of the issues related to safety requirements into multi-domain safety certification, and suggests that more work is required to ensure that certification is based on a complete and consistent set of requirements.

Iankoulova et al. [4] conducted a systematic review on cloud computing security requirements with the intention of providing a comprehensive and structured overview of cloud computing security requirements and solutions. Yahya et al. [5] carried out a review of tool supports for security requirement engineering. Both studies focused on security requirements, which are requirements that describe how to protect assets and system services against malicious actions and security threats [133]. Although safety and security are closely related, they have been developed as different research topics, according to Leveson [133] “one related to threats to life and property and the other with threats to privacy or national security.”

In 2000 Lutz published a report [13] on the current state of software engineering with regard to safety and proposed directions for necessary work. The report provides an overview of the state in the following areas of software engineering for safety: hazard analysis; safety requirements specification and analysis; designing for safety; testing; certification and standards. The report points out six directions of necessary work, which are the following: integration of informal and formal methods; constraints on safe product families and safe reuse; testing and evaluation; runtime monitoring;

education; and collaboration with related fields such as security and survivability, software architecture, theoretical computer science and human factors engineering.

In 2007 Heimdahl published a report [8] addressing issues that pose challenges for the future in terms of safety and software intensive systems. The issues discussed in the report are: misunderstandings about safety concepts and poor usage of known system safety techniques; a failure to demonstrate whether safety requirements specifications have been met; problems when introducing modelling and automated tools to increase productivity in safety-critical systems development [129]; and the increase of safety-critical systems relying on data.

In 2014 Hatcliff et al. published a report [15] considering the future of software engineering, challenges and research directions related to the development and certification of safety-critical software dependent systems. In this report the authors identified seven primary challenges: the need to establish foundational principles for software safety assurance; the nature of criteria in safety certification; the issues related to safety requirements specifications; the need to develop engineering and assurance approaches to support compositional certification and reuse of certified components; the use of tools in the certification process; the increasing automation in hazard analysis [73,184]; and the need of building competence to engineer software for safety critical systems. The reports produced by Lutz [13], Heimdahl [8], and Hatcliff et al. [15] are complementary and offer a good view about the issues in safety-critical systems and how the area has evolved over the last two decades.

Although the above works cover several aspects related to safety and security for safety-critical systems, none of these works perform an extensive identification and mapping of elicitation, modelling, specification and validation of safety requirements approaches for safety-critical systems.

3. Research methodology

In this section we present the design and the execution of the SLR. The research methodology to conduct the SLR was based on the guidelines proposed by Kitchenham [9] and Biolchini et al. [10]. The need for this SLR (Step 1, Fig. 1) was presented in the Introduction of this paper. Fig. 1 shows the research process we have adopted and its steps are described in the next sections. In order to determine if similar work had already been conducted we searched the Compendex, Inspec, and Google Scholar digital libraries (performed on 26 September 2014). The following string was used to search within titles, abstracts and keywords:

(Safety AND/OR Critical) AND Requirement* AND (Review OR Taxonomy OR SLR OR Research overview)

None of the retrieved studies were directly related to the objectives expressed in the research questions (Step 2–Defining Research Questions). In Table 1 we describe the research questions and their aims. The research questions elaborate on what should be extracted from the selected studies and points to directions for further discussions.

The Technology Acceptance Model (TAM) proposed by Davis [46a] and reviewed by Legris et al. [45a] motivated us in relation to RQ1.2. We used TAM as a theory to support in the verification of how effective the safety requirements approaches have been. TAM suggests that perceived ease of use and perceived usefulness are the two most important factors in explaining technology acceptance [45a].

Enabling communication and coordination among the stakeholders is a central purpose of Requirements Engineering [36] and [43]. According to Ramesh et al. [47a] to reach intensive communication among the stakeholders is the most important factor for successful agile requirements engineering. Taking into account the importance of such aspect in requirements engineering processes, the motivation of RQ2 is to verify how the safety requirements approaches have supported the communication among the stakeholders throughout SCS lifecycle.

3.1. Search strategy

For the identification of papers we adopted a search strategy presented in Fig. 2, which is based on the search strategy conducted by Unterkalmsteiner et al. [11]. The keywords defined to compose our search string were extracted from the research questions. In order to evaluate the quality of the proposed search string, we carried out a trial search on Google Scholar (Step 5). We manually selected ten relevant publications in the area of safety-critical systems and compared them with the results from the trial search. The adopted search string captured nine out of ten reference publications. The search string used in the SLR is specified below:

"Safety Requirements" AND ("Safety-Critical Systems" OR "Safety-Critical Software") AND ("Hazard" OR "Accident")

The expression ("hazard" or "accident") was included as part of the search string to guarantee that the selected studies would bring discussions related to accident and hazard analysis connected to safety requirements. We made this decision taking into account that safety requirements are typically "derived" from hazards during the requirements elaboration. Therefore, seeing hazards alongside safety requirements leads to better scoping for the SLR.

In order to manage and organize the papers selected in the SLR we adopted Mendeley (http://123.233.119.36:80/rwt/119/http/P75YPLUNMWYGI3LMMW6T6Z5QNF/).

3.2. Review protocol

We developed a review protocol (Step 3), in which the main elements are as follows: the selected resources chosen were ACM digital libraries, IEEE Xplore, Springer-Verlag, and Science Direct; the search method was based on research through web search engines available for use in digital libraries; the population was composed of peer-reviewed publications reporting approaches to elicit, model, specify or validate safety requirements for safety-critical systems; the aim of the intervention was to collect empirical evidence in relation to (i) what approaches exist; (ii) how they were validated; (iii) how the usability and usefulness of these approaches were measured; and (iv) how these approaches support communication among actors in safety-critical systems development.

As studies selection criteria we determined that papers should be scientific articles from journals, transactions, magazines, conferences and workshops. Only articles written in English would be considered. The inclusion criteria was defined as: any study that presents, compares or discusses approaches for eliciting, modelling, specifying and validating safety requirements for safety-critical systems. Studies relating safety requirements in the context of safety analysis, hazard analysis, or safety-critical standards must also be included. The exclusion criteria was defined as: studies that do not bring any discussion about approaches for eliciting, modelling, specifying or validating safety requirements for safety-critical systems.

As part of the protocol review evaluation (Step 4), we conducted a trial with the studies selection criteria before initiating the selection of the primary studies. In the trial, one participant in the SLR used the proposed selection criteria and found ten articles that should be included in the primary studies set, and ten articles that should be excluded. These same twenty articles were evaluated by the other participant, using the same selection criteria. This resulted in 86% agreement between the two participants using the studies selection criteria.

3.3. Procedure for studies selection

Fig. 3 shows the procedure for the selection of primary studies, which was based on the inclusion criteria presented in Section 3.2. Additional selections were carried out when the title and abstract of the study under analysis did not offer sufficient information to decide if it met the inclusion criteria or not. In such a case, the additional selection process included the perusal of the introduction and conclusion of the study.

We used the search string presented in Section 3.1 to capture the primary studies from the selected resources, based on the review protocol presented in Section 3.2. We captured a total of 611 studies, which were distributed into: the ACM digital library, IEEE Xplore, Springer Link, ScienceDirect (see Table 2 for the number of the studies in each phase of the studies selection). We chose these four digital libraries taking into account that they cover the main conferences and journals related to SCS. We screened them according to the procedure for primary studies selection presented in Fig. 3. After applying our selection criteria into the 611 studies (phase 1), we selected 193 studies for data extraction (phase 2).

3.4. Data extraction

Initially 193 studies were accepted for data extraction (Step 7). However, during the data extraction process, 28 studies were excluded (phase 3) because some of them were out of the scope of this SLR and others were a variation of the same study, not characterizing a new contribution. Besides, 14 studies were excluded because they were categorized as a “report,” i.e. they discussed challenges and tendencies in safety-critical systems, but did not propose that any new approaches be evaluated (phase 4). Considering these exclusions, 151 studies were finally accepted for data extraction (for SLR references please see Appendix A).

We prepared an initial form with the properties presented in Table 3, which also shows the mapping between the properties and the research questions. For properties P1, P2, P5, P6 and P7 a list of expected values was defined previously. For the properties P3, P4 and P8 the values should be extracted from the studies. Each property will be explained in subsequent subsections.

3.4.1. Research method (P1)

The studies were categorized according to the applied research method. We partially adopted the categorization used by Unterkalmsteiner et al. [11], and Ivarsson and Gorschek [14]. The categories and criteria to classify the studies were as follows:

Case study: the study empirically evaluates an approach or a theoretical concept in an industrial setting, having a clearly defined goal [11], [27] and [32];

Industry report: the aim of the study is to report industrial experience without declaring research questions or theoretical concepts [11]. Such a category complies with the “Project monitoring” method described in [27];

Experiment: the study carries out an experiment clearly defining its design, although not necessarily in an industrial setting;

Conceptual analysis: the aim of the study is to discuss a theoretical concept or a new approach, but without validating it, based on a case study or an experiment;

Lessons learned: the aim of the study is to report on the lessons learned from practitioners for an industrial setting;

Report: the study includes a discussion to point out challenges and future tendencies based on an analysis of past and state-of-the-art technology;

Example: the study presents a new approach and conduct a preliminary discussion about it based only on the authors assertions or it provides some examples. Such a category complies with the “Assertion” method described in [27];

Survey: the study collects data based on a questionnaire or interviews [28] and [32];

SLR: the study conducts a systematic literature review.

3.4.2. Context (P2)

The studies were categorized into industry and non-industry cases. The industry cases included studies in which research was conducted in collaboration with, or promoted within, industry. The non-industry cases comprised studies conducted in an academic setting.

3.4.3. Safety requirements approach (P3)

The categories of the approaches were extracted from the studies. Two major categories were identified: (i) a process-independent approach, which can be a technique, model, language, tool or method, which can be used in any process or framework; and (ii) a process-dependent approach, which can be a framework, methodology or a process model usually supporting several steps of the safety-critical systems development [77,107,118,138,140].

The process-independent versus process-dependent dichotomy emerged during the SLR, it was not previously defined by us. However, we decided to maintain this dichotomy in order to facilitate the comparison among the approaches. The process-dependent approaches have a broader scope than the process-independent ones, for this reason it would not make sense comparing and analysing approaches with different system lifecycle coverage into the same group.

3.4.4. Activity/process (P4)

We analysed the studies looking for approaches with the potential to improve safety requirements in the following activities or processes: safety analysis, requirements elicitation, requirements specification, requirements modelling, requirements validation, hazard analysis and safety case specifications [112].

3.4.5. Validation (P5)

To help us during the validation of the studies, we adopted a model for evaluating rigor and the industrial relevance of technology evaluations proposed by Ivarsson and Gorschek [14]. Since we are interested in identifying to what extent the proposed approaches were validated by the authors (RQ1.1), the procedure of how evaluate industrial relevance proposed in [14] seems to be adequate to support us in the SLR.

The model captures rigor and the relevance of technology evaluations in software engineering research, enabling a classification in order to characterize research performed in an applied field. In this model, the industrial relevance is evaluated based on four aspects: (1) the subjects that participated in the studies; (2) the context in which the studies were performed; (3) the scale used in the studies evaluation; and (4) the research method adopted in the studies. The first three aspects were concerned with the realism of the environment where the investigated studies were performed, and the fourth aspect was concerned with how the research methods influenced the produced results. Based on these aspects the levels of relevance can range from 0 to 4.

The rigor is evaluated based on three aspects: (1) the degree to which the context is described; (2) the degree to which the study design is described; and (3) the degree to which the validity is discussed. Based on these three aspects the levels of rigor can range from 0 to 3.

3.4.6. Usefulness and usability (P6 and P7)

We use the concepts of usefulness and usability to verify the effectiveness of the proposed approaches reported in the selected studies. For both concepts we investigated "how" and "why" the authors measured them in the studies. In order to measure the usefulness and usability of the approaches, we proposed the following criteria:

Weak: no evidence or indication of how/why usefulness/usability were measured are presented;

Medium: evidence or indications of how/why usefulness/usability were measured are briefly presented;

Strong: evidence of how/why usefulness/usability were measured and described in detail to a degree where the reader can understand and compare it to other works.

3.4.7. Communication (P8)

Communication among actors plays an important role during the safety-critical system development process. In RQ2, we are interested in verifying to what extent the proposed approaches have addressed such a property. To evaluate how the approaches support the communication of the safety requirements among the actors involved in critical-safety projects we considered the following aspects:

Artefact: any sub-product or delivered result derived from the use of an approach.

Actor: any player involved in the safety-critical system development and/or certification process, such as safety engineer, requirements engineer, system engineer, auditor, etc.

Activity/process: any activity or process embedded in the steps of the safety-critical system development and/or certification process, such as safety analysis, hazard analysis, risk assessment, requirements elicitation, requirements specification, requirements validation, etc.

3.5. Study quality assessment

According to Kitchenham et al. [9], the study quality assessment (Step 8) should help researchers that are conducting the SLR, offering guidance in the interpretation of findings obtained from the selected studies. The study quality assessment we have conducted was an evaluation of how well the studies were reported. We answered the questions presented in Table 4 for each paper considered during the data extraction, while the quality assessment questions were adapted from [11].

3.6. Threats to validity

3.6.1. Publication bias

A general problem related to publication bias is the trend of researchers to publish positive research outcomes rather than negative ones [9]. We considered this threat as low since the research questions in the SLR do not address the performance or efficiency of the approaches to handle safety requirements. The same can be considered to the threat of sponsoring.

We did not limit the sources of information to specific publishers, journals, or conferences. On the contrary, the coverage of the review reached a variety of scientific media vehicles, including the main conferences and journals in the area sponsored by different publishers. In order to accumulate reliable information, we decided not to include technical reports, works in progress, unpublished, or non-peer-reviewed publications.

3.6.2. Identification of primary studies

The initial set of studies were obtained based on the search string presented in Section 3.1. The metric we used to decide about the quality of the search string was the recall of the search result, which is expressed as the ratio of the retrieved relevant studies and all existing relevant studies [33]. Of course, it is not possible to know all existing relevant studies. Taking this into account, the recall of the search string was estimated based on a pilot search as described in Section 3.1 (see Fig. 2). The search string was conducted on Google Scholar and the recall was checked against known relevant studies into an iterative refinement process. The refinement of the search string continued until we reached a recall with at least 90% of the known relevant studies. However, even when using such a strategy to refine the search string, the threats of missing relevant studies still remains. For instance, we only found one study in the Oil & Gas domain, it might be the case that the researchers in this domain do not use the same terminology and publish relevant papers under completely different categorizations and in other forums.

3.6.3. Data extraction consistency

Data extraction was carried out based on the form previously designed for us. We prepared the data extraction form considering the extraction properties discussed in Section 3.4 (Table 3), which are related to the RQs we intend to answer in the SLR. In order to check the consistency of the data extraction form, we conducted a pilot data extraction taking into account 7% of the total studies that we had selected (12 out of 165 studies).

With the extracted data from these studies, we simulated an analysis to check if the types of data that were extracted could help us answer the proposed RQs. After the refinement process of adjusting the form, several categories and parameterized attributes were created to help the researcher conduct the data extraction as a systematic process.

The data extraction form was implemented as a spreadsheet composed of 31 columns (data extraction attributes), which was structured as follows: 3 columns related to article identification, 2 columns related to context (P1 and P2), 4 columns related to the safety requirements approaches and activity/process (P3 and P4), 7 to validation (P5), 10 to usefulness and usability (P6 and P7), and 5 to communication process (P8).

We believe that the adopted strategy to build the data extraction form helped to attenuate threats to the data extraction consistency. However, the threats to missing some important data, and the subjective judgment that may have been used by the researcher during data extraction, still exist.

4. Results and analysis

A total of 165 studies discuss safety approaches, which are method/model/technique or solutions to elicit, specify, model or validate safety requirements, including safety and hazard analysis related to identifying and analysing safety requirements in the context of SCS.

The inspected publications were classified according to the applied research methods, as can be seen in Fig. 4, where case studies (33.33%) and experiments (30.3%) together clearly constitute the most common methods applied. These are followed by conceptual analysis (15.15%), reports (8.48%), examples (4.84%), industry reports (3.63%), lessons learned (2.42%), surveys (1.21%), and SLRs (0.6%).

Applying rigor and relevance (Section 3.4.5) gives the overview seen in Fig. 5 (applied on 151 studies, with 14 excluded and classified as reports). As we can see, there is nearly a symmetry in the studies dispersion, with 38.41% of the studies obtaining the highest score of relevance (4 points) and 31.12% of the studies obtaining the lowest score of relevance (0 points). 25.16% of the studies reveal rigor ⩾2 (ranging from 0 to 3), and 47.68% of the studies show rigor <1.5. Only 5.96% of the studies reached high scores (relevance ⩾3 and rigor ⩾2.5).

When we apply rigor and relevance for each research method category, the studies from the case study category obtained the best results, with 94.54% of them achieving a relevance ⩾3 and 16.36% boasting rigor ⩾2.5. This was in stark contrast to studies from the other research methods category where only 2% of the studies from the experiment category reached a relevance ⩾3, and likewise for rigor. These results are not surprising in that case studies are relevant since they are studied in industry. However, it is surprising that the rigor is high, which contradicts the results presented by Ivarsson et al. [34], where rigor was much lower in "solution" based studies. This can indicate that the application domain here is more mature, or has higher demands in relation to rigor than can be seen in the Requirements Engineering field in general. In the case of experimentation, a low level of relevance is not surprising as experiments often have to be conducted in a controlled environment, scaled down [27]. What is surprising is that rigor also lags. Normally the procedures and demands on experimental design as well as operations are very high [35].

Fig. 6 shows the distribution of the studies by activities/processes related to safety requirements handling. Safety analysis (43.71%) and requirements specifications (39.07%) are discussed in the majority of the studies, followed by requirements elicitation (31.79%), requirements validation (24.50%) and hazard analysis (13.91%) with a considerable number of the studies concerned with these activities. Safety cases and requirements modelling are discussed in 8.61% and 6.62% of the studies, respectively.

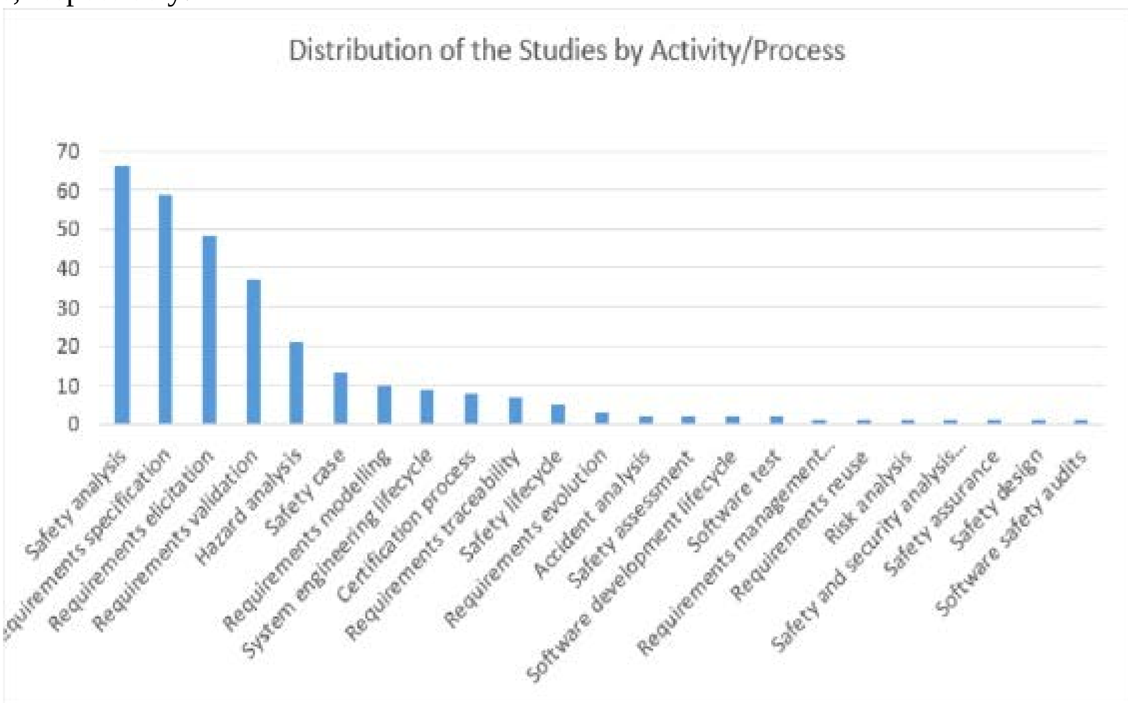


Fig. 6. Distribution of the studies by activities/processes.

Just a few studies (1.3%) have reported approaches supporting the relationship between safety requirements and system testing [94]. This finding seems to be in accordance with what was pointed out by Nair et al. in [48a], where the authors observed the lack of published work on system testing for safety critical systems in the context of safety certification. Additionally, Nair et al. [49a] made a comparison of importance given in practice and importance observed in the technical literature for safety evidences. They found that "testing results", which is considered a type of safety evidence in the context of safety certification, has high importance in practice, but low in technical literature.

Lutz [13,86] has discussed the importance of testing in the context of safety-critical systems and has highlighted the need for approaches to improve the area. Safety requirements elicited during hazard analysis should be used during testing to validate if the as-built system satisfies them [13]. This is particularly important considering that all SCS must demonstrate that they are in compliance with process-based safety standards or to show how safe they are based on the safety assurance process (safety case oriented approaches) [8] and [15]. Showing evidence that the safety requirements were taken into account during the test cases designing is highly recommended to get success into the safety certification process [15]. The lack of approaches supporting these crucial steps of the system development may suggest that the communication among safety-critical team actors are not properly performed. This is further discussed in Section 4.4.2.

4.1. Approaches for safety requirements (RQ1)

4.1.1. Results

The purpose of this research question was to identify which approaches have been proposed to elicit, model, specify or validate safety requirements in the context of safety-critical systems.

Fig. 7 shows the distribution of the process-independent approaches, i.e., approaches classified as techniques, methods, models and languages that can be used in any process, methodology or framework for safety-critical systems development. As can be seen, FTA is the most adopted approach, appearing in 15.23% of the studies, followed by FMEA (9.93%), HAZOP (7.95%), FMECA (6.62%), PHA (5.96%); these approaches are traditional in the context of safety engineering [16]. The newer approaches of GSN and RSML/RSML−e were independently adopted in 4.63% of the studies.

73 different approaches were found in isolated studies, i.e., each one was found in only one study. This may indicate that there are a couple of well researched and developed approaches, and many smaller ones that were newly presented in individual papers, with many that later undergo no further development or validation in either an academic setting by the researchers themselves, or in an industrial setting. This is not uncommon in research; many approaches are proposed but few “survive.” What is surprising is that we cannot find any evidence of validation and evaluation.

Fig. 8 shows the distribution of the approaches among the activities/processes related to safety requirements. As we can see, FTA is the most frequent approach for safety analysis, requirements specification, requirements elicitation and requirements validation. The use of FTA in such activities is followed by FMEA, HAZOP, FMECA and PHA, respectively. From the selected studies in this SLR, GSN is the only approach used for documenting safety cases.

4.1.2. Analysis and discussion

We found that 70.37% of the studies published in the 1990s suggest that at least one of the following approaches should be adopted by developers to improve safety analysis and/or safety requirements specification: FTA (10), FMEA (3), HAZOP (2), FMECA (2) and PHA (2). When we analysed the studies developed during the 2000s, we found that 40.0% of the studies suggested using at least one of these same approaches, with the following distribution: FTA (17), FMEA (11), HAZOP (7), FMECA (7) and PHA (6). Although in absolute numbers such traditional approaches are greater in the 2000s than in the 1990s, in relative numbers there is a decreasing tendency, from 70.37% in the 1990s to 40% in the 2000s.

The decreasing tendency about these approaches along the last 25 years may suggest that the safety-critical system community is trying to find out other alternatives to improve safety requirements [29]. On the other hand, 40.0% of all studies means that a considerable amount of effort still lies in such traditional approaches and suggest that the safety-critical systems developers still are adopting such approaches on a large scale. The additional 73 different approaches found (each one in a different study), as mentioned previously, seem to support the argument that developers prefer the traditional approaches [15,57,89,156] and that the new approaches are not yet mature or not convincing enough.

Regarding these isolated approaches, most of them (21.81%) are formal methods, which presented no validation in the industry context. Considering all 73 isolated approaches, 56.36% were discussed in papers that adopted experiments, conceptual analysis or examples as research methods. Most of the isolated approaches were not proposed in partnership projects with industry, as well as being poorly suited for practitioners' real problems.

A poor integration between the safety engineering and requirements engineering areas is pointed out by several researchers as an open challenge in SCS [8], [15] and [16],134]. New approaches to improve safety cases is also pointed out as an important topic for further investigation

[8,58,120,122,139,147,152,181,188]. However, very few of the isolated approaches addressed the above-mentioned open topics. On the other hand, the majority of them (65.45%) approached requirements specification and safety analysis issues in a non-integrated way. These reasons may explain the premature death of these isolated approaches.

Leveson [29] describes several reasons why the safety-critical community needs to move towards new approaches for safety engineering, among them the most important are: the fast pace of technological change, reduced ability to learn from experience, the changing nature of accidents, new types of hazards, increasing complexity and coupling, difficulty in selecting priorities and making trade-offs, increasing the control of systems with automation, and changing regulatory and public views of safety. Despite there being no new approach clearly taking over from the traditional approaches, our findings show that some newer approaches are becoming more prevalent than others, such as RSML/RSML−e for requirements specification, and GSN to describe safety arguments. RSML was developed by Nancy Leveson's group as a formal language for specifying the behaviour of process control systems. The main characteristic of this language, that distinguishes it from others, is its improved readability and understandability as well as support for end users, system engineers, managers and auditors from regulatory agencies [161]. Additionally, the requirements specified with this language can be further analysed by a model checker and theorem prover commercially available [135], partially automatizing the analysis.

In a recent work Hatcliff et al. [15] stress the importance of requirements engineering for safety-critical systems. In their work it is noted that system engineering, safety engineering and software engineering are still not particularly well-integrated disciplines. They believe that the isolation starts with "competence-building infrastructure" and continues through domain vocabularies, paradigms and tools [97]. The recognition of such a gap by the safety-critical system community is not new and it was already mentioned by other researchers, e.g. Heimdhal [8], Hansen [67], and Leveson [16,134]. Looking at Fig. 6 the selected studies show that safety engineering and requirements engineering processes are largely discussed by researchers concerned with safety-critical systems. The selected studies in this SLR strongly cover safety and hazard analysis, discussed in 66 and 21 studies respectively, as well as requirements specification, elicitation and validation discussed in 59, 48 and 37 studies, respectively. The capturing and definition of safety requirements, both during the safety analysis and the requirements engineering processes, is a common topic discussed in these studies, which may indicate that the gap between safety engineering and requirements engineering is getting smaller [76,104,142,155,162,172,169,192]. However, this finding does not mean that these disciplines are integrated in a structured way, but rather that further research should be conducted to investigate how the integration between them is progressing and/or what gaps need to be addressed.

Additionally, as we can see in Fig. 8, the approaches named FTA, FMEA, HAZOP, FMECA and PHA are the most commonly used during safety analysis, requirements elicitation, specification and validation, as well as in hazard analysis. The concentration of these approaches among certain activities/processes seems to indicate a tendency towards integration between requirements engineering and safety engineering approaches. Particularly, the strong use of FTA for both safety analysis and requirements engineering activities (i.e., requirements elicitation, specification and validation), might suggest an approximation of such activities/processes during the early stages of the system engineering process. This might imply that both safety engineering and requirements engineering have similar needs and may contribute to each other. The common use of techniques may improve the communication process among the actors [166], by reusing the produced artefacts, and consequently decreasing the costs related to the specification and evolution of the safety requirements [36,83,186,193].

Another interesting finding that we can observe from Fig. 8 is that GSN was the only one approach mentioned in relation to the documentation of safety cases. According to Kelly [181] a safety case "should communicate a clear, comprehensible and defensible argument that a system is acceptably safe to operate in a particular context." In an industrial case study developed by Dardar et al. [56] the safety case was composed of four conceptual elements: (1) goals, which represent safety requirements that must be satisfied in order to ensure that the system is safe; (2) evidence, which represents the proof that the goals are achieved; (3) arguments, which represent the relationship between goals (safety requirements) and their evidence; and (4) context, which describes the domain where the system safety is to be argued. GSN is a widely studied and an easy graphic notation that is used to represent safety arguments in safety cases. GSN is composed of the following conceptual elements [181]: goal, justification, context, assumption, solution and strategy; and two types of relationships: "in context of" and "solved by." The fact that GSN was the only one mentioned in the inspected studies seems to indicate that such notation is becoming a prevalent approach in a certification process oriented by goal-based safety standards. However, there are other approaches that could be considered to represent safety arguments in safety cases, such as SafeML [154] and SafeUML [166]. Both approaches are aimed at addressing communication among actors in the development and certification processes in SCS. SafeML is a SysML profile designed for modelling the safety-related concerns of a system, and SafeUML is a UML profile designed to enable precise modelling and monitoring implementation of safety requirements during the development process.

Further work comparing such approaches may bring interesting insights into new possibilities to represent safety arguments. For instance, Heimdahl points out that despite the fact that GSN might help in arguing safety cases, further investigations are necessary with regard to how to effectively argue safety cases [38] as well as in order to discern the acceptable types of evidence when building safety cases [8]. Could SafeML or SafeUML help safety engineers and requirements engineers in producing better arguments than GSN? Further research is necessary to answer such questions.

Since the 1990s, researchers in academia have had high expectations in relation to formal methods as an approach to improve safety analysis and safety requirements specifications. Several approaches were developed during the last 25 years, as presented in [8], [26] and [47a],51,84,116,158,161,172]. Despite the efforts of the academic researchers and the great expectations of formal methods, they still remain little used in an industrial setting [161]. In this SLR we observed that RSML/RSML−e (found in 7 studies), VDM/VDM++ (3 studies) and Petri nets (2 studies) were the most discussed approaches based on the formal methods to handle safety requirements.

We applied rigor and relevance to the studies that reported formal methods. 23.8% of them achieved rigor ⩾2.5, and 28.6% of them boasted relevance ⩾3. One third of the studies were conducted into the industrial context, including within the avionics, defence, railway and nuclear plant domains. From this total, 19% of the studies reported approaches validated on an industrial scale and 81% into down-scaled industrial. We believe that the rigor and relevance of the studies reporting formal methods got reasonable scores, revealing that the studies were carried out close to the real problems. However, a common factor we found among the studies validated into an industrial scale was that almost all of them had validated the approaches that combined traditional techniques, e.g. FTA and FMEA, with formal methods. That was not the case when we looked at the studies validated into down-scaled industrial, where just a few of them had that combination of techniques. This may indicate that practitioners are interested in experimenting with new approaches based on formal methods, since they are combined with traditional and well known techniques adopted within an industrial setting [15,155].

As observed in Fig. 8, RSML/RSML−e were more used in the process of safety analysis, requirement specification and validation. Petri nets and VDM/VDM++ were more used in the process of requirements specification. However, compared to the other approaches such as FTA, FMEA and HAZOP, it is clear that formal methods are not the preference of the practitioners. For safety requirements specifications, we have found that practitioners' preference still rest on informal approaches such as natural language or approaches based on the approaches such as FTA, FMEA, HAZOP, FMECA and PHA. The main reason for this seems to be the mathematical skills usually necessary to handle formal methods. Despite the insistence of those supporters who believe that the necessary mathematical skills to use formal methods are easy, they do recognize that proper training is crucial to the adoption of such methods [30,68]. In addition, time-to-market pressure, in terms of delivering products, which is extended by formalism [30] and [31], combined with necessary additional training, and also the practitioners' perception that new types of mistakes can be introduced when adopting formal methods [80, 121], may be considered realistic barriers for their adoption.

The development of exotic additions to formalism, tools, and support structures (like model checking, etc.) are the focus of the research community in formal methods. At the same time, the research community just assumes that the "base" is used. If this is not true, the base is not used, then we have to ask ourselves two main questions: Will it help to further develop exotics and additions as a means to enable their use and spread? Does the base need to be changed, and do the basic assumptions need to be questioned?

4.2. Validation of safety requirements approaches (RQ1.1)

4.2.1. Results

The purpose of this research question was to identify which approaches the practitioners used in an industrial setting and to what extent they were validated through their use in real development environments and projects. Fig. 9 shows how the selected studies in the SLR were validated by the authors while taking into account the subjects, context, scale and research method adopted in the studies.

Fig. 9 shows that in 57.61% of the studies, only researchers participated. In 27.16% of the studies, researchers and practitioners collaborated with each other, and in 15.23% of the studies only the practitioners participated. Considering the context in which the studies were carried out, 60.93% were conducted in industry cases and 39.07% in non-industry cases. In 42.38% of the studies the scale was industrial, in 54.30% it was down-scaled industrial, and only in 3.32% the scale was classified as toy examples. In 36.42% of the studies, the adopted research method was case study, in 33.11% and 16.56% it was experimental and conceptual analysis respectively, and in 13.91% the research methods of the studies were distributed into examples (5.30%), industry reports (3.97%), lessons learned (2.66%), surveys (1.32%) and SLRs (0.66%).

Fig. 10 shows the distribution of the studies by application domain. 28.48% of the studies were classified as domain-independent, the remainder of the studies were developed in the following application domains: avionics (17.88%), railway (11.26%), spacecraft (10.78%), defence (6.62%), medical devices (5.96%), automotive (5.29%), nuclear plants (5.29%), and others (12.58%) including process control [63,75,101,195], health care [149], road traffic control [141], robotics [66,69,95], chemical plant [171,51], fire systems [144], oil & gas [102], and rebreather equipment [58].

The high maturity and volume of business in the avionics and railway domains might explain the concentration of the studies in these areas. The spacecraft and defence domains are also mature and involve a high volume of business, however the development in such domains usually demands secrecy and confidentiality, which might make it difficult to publish results. We found only one study in the oil & gas domain [102]. We find this surprising, as this domain type requires strict adherence to e.g. environmental restrictions. Similarly we would have expected more papers on areas such as water and environmental protection.

4.2.2. Analysis and discussion

Despite the prevalence of the industrial context reported in the selected studies (60.93%), the other aspects of industrial relevance (subjects, scale and research methods) indicate that the most of the studies were not developed in close industry participation. Most of the studies (57.61%) was developed without, or with only minimal participation of, industry practitioners. Despite safety-critical systems being a strongly applied field, only 36.42% of the studies reported that the proposed approaches for safety analysis and safety requirements specification were validated based on the case studies. The scale in which the proposed approaches were validated also suggests that the proposed approaches to improve safety requirements should be tested in a more industrial setting, in that only 42.38% of the studies were validated on an industrial scale.

When we analyse how many studies were conducted by practitioners on an industrial scale, we found that 23.84% of them were carried out by researchers and practitioners in collaboration, and 9.93% of them by practitioners only. These indicators reinforce the need for closer ties with industry. As discussed by Ivarsson et al. in [34], the technology transfer from research results to the industry setting is probably one of the main goals of an applied field, such as requirements engineering. A trustworthy validation of new technologies is closely linked to the context and the subjects that participate in the validation studies, i.e., experienced practitioners validating new technologies in an industrial context, usually increase the reliability of a new technology [102,108,167]. Considering how important high quality requirements specification are for SCS, and their strong influence on the safety certification process [8] and [15], it would be helpful if managers and decision makers could base their choices about new approaches based on those validated by an increasing amount of practitioners in an industrial context.

In 43 studies (28.48%), the authors mentioned or gave some indication that the proposed approaches could be used in different safety-critical domains. Such studies were classified as domain-independent. Considering specific application domains, avionics was the domain where more approaches for safety requirements and hazard analysis were validated (17.88%). Avionics is a domain strongly ruled by regional and international regulators with well- established safety standards around the world, such as DO-178B [45a,72,145,194], ARP4754A [46a,137] and ARP4761 [15]

Fig. 11 shows the results after applying the model for evaluating industrial relevance in the selected studies. 51.85% of the studies in the avionics domain achieved a high degree of evaluation, with a relevance score $\geqslant 3$ (ranging from 0 to 4), while only 20.93% of the domain-independent studies received such a score. A possible explanation for the dominance of domain-independent studies in relation to the dependent ones is the difficulty to conduct case studies involving real projects from industry (i.e., in a specific domain), as indicated by the low relevance score for independent-domain presented in Fig. 11. Another possible explanation could be that researchers focus on generalizability – thus not wanting to create domain specific "solutions". However, according to our results, we can observe that studies conducted in close collaboration with industry, e.g., in the avionics domain, can reach a higher score of relevance – so the question is if research will be more useful if domain adapted.

The high risks involved to users in the avionics segment pushed the area to find the best practices to improve their safety-critical systems [105,109,170,174,175]. This factor seems to suggest why the studies in the avionics domain were developed more closely with industry than the domain-independent studies. Several studies in this SLR reported a partnership between the avionics industry and academia [45a,72,100,105,161,162,175].

Miller et al. [161] reported a case study conducted by the Advanced Technology Centre of Rockwell Collins, the Critical Systems Research Group at the University of Minnesota, and the NASA Langley Research Centre to determine how well a formal analysis could fit in an industrial example. Abdul-baki et al. [45a] reported on specification-based testing, analysis tools, and associated processes that were used to independently validate, verify, and provide for certifying, safety-critical software developed for the Traffic Alert and Collision Avoidance System. The reported case study was developed at the Rannoch Corporation and the practitioners adopted formal methods during the system specification. Modugno et al. [162] explored the feasibility of such a comprehensive safety analysis on the specification of a guidance system for high-speed civil transportation being developed at NASA Ames. The goal of the reported case study was to determine the feasibility of performing such an analyses and to evaluate the techniques and their contribution to the safety analysis processes. In these three studies, the authors discuss the need for better requirements engineering in the avionics industry. The common focus was to test approaches to improve software requirements specification and validation.

Hatcliff et al. [15] argue that, as in mainstream IT, one of the biggest challenges in safety-critical systems engineering is to establish a complete, correct, unambiguous, and understandable requirements specification. Our findings confirm such assertions, since most of studies have reported approaches to improve safety requirements specifications. Despite this stringent way of specifying requirements, it seems to run counter to the current agile and lean software development approaches, and it is important to remember that the trade-off between agility, cost reduction and safety is a complex problem in SCS development, with ethical implications [16,121]. The main factors that explain why a more stringent requirements specification approach is necessary in the SCS domain are the following: requirements of regulatory agency recommending accurate requirements specifications [15,174,181]; understanding in the SCS community that deficiency in requirements is the biggest source of unanticipated cost and delivery delay [8], [15] and [26]; the need for certifiably safe systems before deployment, and taking into account that SCS can trigger accidents with high severity level [29].

Additionally, another interesting finding related to how the investigated studies have validated the proposed approaches, in order to improve the safety requirements, was the use of safety standards to support such approaches. In 51.66% of the studies we found that the authors had based the proposed approaches on recommendations from safety standards, or at least recognized the importance of such recommendations for their proposals. The most cited safety standards were IEC 61,508 (reported in 28 studies), which is a generic safety standard for electrical/electronic programmable safety-related system [22,127,176]; DO-178B (in 20 studies), which provides safety guidance for the avionics domain; EN 5012x (in 11 studies), which is a set of safety standards for the railway domain; and MIL-STD 882x (in 9 studies), which is a set of military standards for the defence domain published by DoD-USA. Considering that any commercial safety-critical system controlled by software must be submitted to a safety certification process, which is based on one or more safety standards, the amount of studies that are "safety standard-oriented" may suggest a promising tendency in academia towards proposing approaches in compliance with safety standards.

The research community in SCS recognizes the importance of safety standards and strive to develop new approaches taking into account such standards. However, only few works have reported case studies involving actors covering the whole SCS lifecycle, i.e., from the requirements definition up to the safety certification process. This issue is further discussed in Section 4.4.2.

4.3. Evidence of usefulness and usability measurements for safety requirements approaches (RQ1.2)

4.3.1. Results

The purpose of this research question was to evaluate to what extent usability and usefulness have been explored for the approaches found in the selected studies. In order to achieve this evaluation we proposed some criteria to classify how and why usability and usefulness were measured in the studies. The criteria were based on the evidence presented in the studies according to three levels of evidence: weak, medium and strong (defined in Section 3.4.6).

Fig. 12 shows that the studies have presented reasonable evidence of how and why the usefulness of the proposed approaches were measured. We found that 25% of the studies revealed strong evidence on how the usefulness of the approaches was measured, and 61.03% showed medium evidence. Regarding the why usefulness was measured, we found that 53.68% of the studies boasted strong evidence, and 36.03% showed medium evidence.

On the other hand, the studies had poor evidence of how and why the usability of the proposed approaches was measured. In 94.85% of the studies we found weak evidence related to how the usability was measured, and in 95.59% of the studies the evidence of why the usability was measured was classified as weak.

4.3.2. Analysis and discussion

The results presented in Fig. 12 show that the selected studies bring much more evidence about the usefulness than the usability when it is considered how and why they were measured. In relation to the aspect of usefulness, we found that the studies bring better evidence of "why" than "how" the studies were measured. This finding shows that the selected studies generally emphasize the problems of a poor safety requirements specification and their consequences rather than possible solutions to solve those problems. It seems to be clear to the authors why they proposed the approaches to improve safety requirements, however the validation of the usefulness of such approaches still demand more evidence.

When we move to the evidence about the usability of the proposed approaches, the results are disturbing. 3.68% of the studies revealed medium evidence of how the usability was measured, and only 1.47% had strong evidence. This is even more disturbing when we analyse the studies that proposed tools to support techniques, models and methods. From the 29 studies that proposed tools to improve the processes of safety analysis and/or requirements specification and validation, only 2.2%

of them had medium or strong evidence of how the usability of the tools was measured. Two studies had medium evidence and only one study boasted strong evidence. The lack of evidence supporting the benefits of the new approaches in software engineering has already been highlighted by Kitchenham et al. [40] and [41]. Our results seem to indicate that the same problem happens in the SCS domain. The lack of evidence about the usability of the new approaches, added to the relatively low participation of the practitioners in the validation process of such approaches (as discussed in Section 4.2.2), seems to indicate that the technology transfer from academic research to industrial setting in the SCS domain have faced difficulties. Additionally, the practitioners' resistance to change from traditional approaches to new ones, as discussed in Section 4.2.1, also might have its roots in the lack of evidence on the new approaches benefits.

Daramola et al. [55] reported the use of a framework for reuse-oriented HAZOP supported by KROSA (Knowledge Reuse-Oriented Safety Analysis). To validate the usefulness and usability of the tool a one-day workshop on how to use the tool was conducted involving all the participants in the case study, after which they had one week to try and interact with the tool. The participants also had a detailed user-manual as a further guide for using the tool. The safety engineers that participated in this study were unanimous in confirming that the tool would provide valuable support for the conduct of HAZOP, with the potential of alleviating the complexity of the HAZOP process by enabling the reuse of experience. Mavin et al. [90] reported the use of a tool called ART-SCENE, which is a scenario technique for discovering and documenting stakeholder requirements. It was proposed that ART-SCENE could be used to derive safety requirements from an engine control system scenario. A trial was carried out to assess the effectiveness of ART-SCENE against a standard safety method (HAZOP) for deriving requirements. The tool was used by systems engineer and safety engineer during a trial.

Panesar-Walawege et al. [187] described a tool-supported solution based on Model-Driven Engineering (MDE) to support the verification of compliance to safety standards. An empirical validation was performed based on an industrial case study that showed how the concepts of a sub-sea production control system can be aligned with the evidence requirements of the IEC 61,508 standard. A subsequent survey examined the perceptions of practitioners about the solution.

Unfortunately, such validation approaches to measure the usefulness and usability of the proposed approaches, as presented in these three mentioned studies, do not appear to be the common practice adopted by researchers in the selected studies. However, these validation approaches are good examples of how to validate new approaches and they show that the effective participation of practitioners is essential to get insightful ideas to rethink and refine the proposed approaches in order to fit the practitioners' needs.

The survey conducted by Panesar-Walawege et al. [187] aimed to investigate the likelihood of an approach proposed by them being adopted by practitioners in industry. Complexity was among the factors that they considered important for the adoption of technology, and it was defined as "the degree to which a technology is perceived to be difficult to understand or use" [42]. The results from this survey indicated that the practitioners found their approach easy to understand and use, and that they would be willing to adopt it in real situations. Such result indicate the importance attributed by practitioners to the issue of usability. Our results in this SLR, as well as the survey conducted by Panesar-Walawege et al., seem to indicate that practitioners are willing to adopt new approaches since such approaches show evidences that they bring value to practitioners and also that they are easy to use.

4.4. Communication and understanding among development project actors supported by safety requirements approaches (RQ2)

4.4.1. Results

The purpose of this research question was to investigate to what extent the approaches enable safety requirements communication and understanding among development project/team actors. In order to answer this question, we distributed the approaches found in the selected studies among the actors in which the approaches were designed to be used for.

Fig. 13 shows the distribution of the use of process-independent approaches among the project actors found in the selected studies. The majority of the approaches were designed to be used by safety engineers and requirements engineers. From the 151 selected studies, 114 studies reported approaches targeted at supporting safety engineers, and 112 studies targeted requirements engineers. The use of FTA by safety engineers was reported in 22 studies, and the use of the same technique by the requirements engineers was reported in 19 studies. The use of FMEA by safety engineers was reported in 14 studies, while its use by requirements engineers was reported in 12 studies.

Few studies mentioned, or suggested, the use of any approaches by other actors involved in the safety-critical system development and/or certification process. The proposed approaches that considered some kind of support for certification auditors were discussed in 8 studies; 16 studies reported approaches considering software engineers; 8 studies considered software designers; and 8, 20 and 15 studies reported approaches involving systems designers, systems engineers and testers, respectively.

Another group of approaches, which we call the process-dependent approaches, also were found in the SLR. From the 151 selected studies, 16 studies (10.6%) reported processes, methodologies or process models with large coverage throughout the SCS lifecycle. Almost all of these approaches were reported in isolated studies, i.e., each one was discussed in a different study. The only exception was the Systems-Theoretic Accident Model and Process (STAMP), proposed by Leveson [29], which is discussed in 3 studies [111,143,153].

4.4.2. Analysis and discussion

Almost all studies that reported process-independent approaches to support safety engineers and requirements engineers took into account both type of actors, i.e., the proposed approaches could help both of them in different but complementary activities, such as safety analysis and safety requirements specification. Again, this finding seems to suggest a tendency of integration between safety engineering and requirements engineering; the approximation between these disciplines was already detected in the findings discussed in Section 4.2.2.

On the other hand, few approaches have been reported involving the other actors in the safety-critical systems development and certification process. For instance, the certification auditors play an important role in the certification process. However, from the selected studies we only found three process-independent approaches being discussed in a context where certification auditors were included. Hawkins et al. [181] presented the concepts of assurance cases and prescriptive software assurance. They describe how an assurance case could be generated for the software of an aircraft wheel braking system. In the case study reported by the authors, GSN was used as an approach to represent the software assurance arguments, which would be assessed by certification auditors in a future process.

Dodd and Habli [174] reviewed existing practices in software safety certification. The authors explored how software safety audits are performed in the civil aerospace domain [37]. They proposed a statistical method based on GQM for supporting software safety audits. As stated by the authors, the results from the case study revealed that "the proposed method can help the certification authorities and the software and safety engineers gain confidence in the certification readiness of airborne software and predict the likely outcome of the audits." We are not evaluating whether such a claim is true, but instead highlighting an approach that was focused on the certification process during the carrying out of the study. Only a few studies (5.3%) in the SLR reported approaches to improve the interaction with auditors during the safety certification process.

Despite the fact that some studies reported approaches in which system engineers, software engineers, software designers and testers are supported, none of them investigated how well the proposed approaches could communicate safety requirements among these actors. FTA and HAZOP were the most reported approaches in the studies aimed at supporting a variety of different actors in the safety-critical system engineering process (see Fig. 12). Nevertheless, the focus of such studies was strongly directed at how well the approaches could support the activities performed by safety engineers and requirements engineers, while nothing was said about the communication with the other actors in the SCS development. This finding is in accordance with what was pointed out by

Fricker et al. in [36] and [43], where the authors claimed that the requirements engineering focused on requirements specification practices but still needed to find better solutions for effective requirements communication.

In Section 4.2.2 we pointed out that the SCS research community has increasingly proposed approaches taking into account the dominant safety standards, which is a positive finding. However, the most of these approaches focuses on the activities performed by safety engineers and requirements engineers, paying little attention to the other actors that participate in the SCS lifecycle, such as testers and certification auditors. This makes it difficult to assess to what extent the proposed approaches may improve the requirements communication among the actors in the whole lifecycle. Two factors may explain such a preference. One is the central role assigned to safety engineers and requirements engineers during the requirements communication process. Two, the difficulty to carry out case studies involving the whole chain of actors in the SCS lifecycle. However, we believe that such a preference in relation to the new approaches for requirements communication is too restrictive, taking into account that the goal of the requirements is to communicate and enable a shared understanding among all stakeholders [36], [43] and [44], not only between requirements engineers and safety engineers.

Communication efficiency is an essential aspect to successfully develop SCS. However, very few selected studies in this SLR addressed communication efficiency while taking into account the whole system engineering lifecycle [168,190]. Bjanarson et al. [44] have identified four main factors that affect requirements communication, namely scale, temporal aspects, common view, and decision structures. According to the results from the empirical study conducted in [44], the first factor (scale) is related to the size of the organizations and the complexity of the products. In the SCS domain, the products are usually complex and developed by large organizations [50,52], then the scale factor is commonly present and affects the requirements communication in SCS development. Another relevant communication factor identified by Bjanarson et al. is the temporal aspect, which is related to the requirements communication discontinuity throughout the project's lifecycle, resulting in communication gaps between the requirements and the development teams. In the SCS domain, such a discontinuity issue can be even more problematic because there are two additional teams to be considered, i.e., the safety engineering and the certification process teams [19] and [25],194]. Despite the fact that these factors (scale and temporal aspect) clearly influence communication in SCS development, we did not find studies reporting approaches to address them.

Leveson [29] discussed the essential role of the communication process throughout the SCS lifecycle, as well as the challenges to integrate safety as part of any system engineering process. Leveson presented the concept of intent specification, which is based on systems theory and system engineering principles, as a tool to help practitioners integrate safety into system engineering, as well as to improve the communication process among the project teams. Among the selected studies in this SLR we found two studies that reported the use of intent specification. Yin et al. [111] presented a method for safety requirements generation based on System-Theoretic Process Analysis (STPA) [29] combined with intent specification. They applied the method to generate safety requirements in an Automatic Train Protection system. Leveson et al. [135] discussed requirements for effective software reuse in embedded systems. They reported the use of the SpecTRM specification/modelling tool, which is based on intent specifications, in a case study to investigate reuse at the software behavioural requirements level on a real spacecraft. However, both studies did not report on the influence or impact of intent specification for communicating requirements throughout the project lifecycle.

Regarding the type of artefacts produced by the approaches reported in the selected studies, we found the following categories: tables and natural language specification were reported in 48 studies (31.79%); diagrams and graphics notation were reported in 59 studies (39.07%); and formal specifications were reported in 17 studies (11.26%). The three most reported artefacts were Fault Tree diagram, FMEA/FMECA table and List of Hazards. The Fault tree diagram was the most reported artefact among the studies (in 21.32%), the FMEA/FMECA table was reported in 16.91% of the studies, and list of hazards was reported in 9.59% of the studies. This result shows the prevalence of informal and semi-formal approaches over formal ones in order to document and communicate requirements, design decisions and relevant information among the project teams and actors in the development and certification of safety-critical systems.

In terms of process-dependent approaches, we found three studies based on STAMP to improve the safety issues during SCS development. STAMP was the only approach reported in more than one study in this SLR. As mentioned above, Yin et al. [111] presented a method to safety requirement generation based on STPA, which is a hazard analysis process based on STAMP, but the authors did not present evidence that their method might improve the communication process among the actors throughout the SCS development lifecycle. Qureshi [143] pointed to the need to move on from the traditional event-based accident modelling approaches, such as FTA and FMEA, to new system-theoretic approaches of modelling and analysis of accidents in safety-critical systems. Among other approaches discussed in his work, Qureshi highlighted STAMP as a promising alternative for such a change. However, the study did not discuss the use of STAMP to improve the communication chain among the SCS project teams. Weber-Jahnke et al. [153] presented an adaptation of STAMP for safety engineering of an Electronic Medical Records (EMR) system. Despite the fact that the authors applied the proposed approach in a real case study, they did not discuss the influence of STAMP upon the communication among the involved actors.

Considering the potential benefits of using STAMP for the communication process among the SCS project teams and actors, as discussed by Leveson [29], it is a surprise that we found no study reporting research to improve aspects of communication based on STAMP.

5. Conclusion

In this paper we presented a systematic literature review that investigates studies reporting approaches proposed to elicit, specify, model and validate safety requirements for safety-critical systems. The most relevant findings from this review and their implications for further research are as follows.

Integration between requirements engineering and safety engineering areas. A considerable part of the studies in the SLR reported approaches to support requirements specification (39.07%) and safety analysis (43.70%), however the same approaches were not necessarily proposed to support both processes together. A common subject discussed in these studies is the capturing and definition of safety requirements, both during the safety analysis and the requirements engineering processes. Despite the high number of studies aimed at this common subject, only 7.28% of them clearly proposed approaches to support both the requirements specification and safety analysis in a cooperative way. This finding seems to indicate a trend of approximation between requirements engineering and safety analysis areas. However, further research is necessary to investigate to what extent the integration between these disciplines is accomplished by the proposed approaches.

Dominance of the traditional approaches. Traditional approaches such as FTA, FMEA, HAZOP, FMECA and PHA, which currently are domain-independent, seem to be dominant in the safety engineering area and they are being increasingly adopted in the requirements engineering process to help during elicitation, specification and validation of safety requirements. FTA and FMEA were originally created for the military and defence domain in the 1960s, since 1970s they became popular and spread out to other domains such as avionics and automotive. HAZOP was originally created for chemical industry domain in the 1960s, however, it has been used in several other domains. In this SLR we found the use of HAZOP in domains such as railway, medical devices, robotics and avionics. Formal methods still remains relatively little used by practitioners, despite significant effort from the academic community to introduce them into an industrial setting. Among the formal approaches, it is worth mentioning RSML/RSML−e among the selected studies. This is the most reported approach in the studies (7), and its solid readability and the available tools for model checking [114] and theorem proving, appear to be its strength. In relation to the approaches to support safety case description, GSN seems to be consolidating. These two approaches seem to be gaining more acceptance among practitioners, however they are still far from the popularity achieved by the traditional approaches.

Early mortality of new approaches. Seventy-three out of 151 studies reported on approaches that were only mentioned once in one study, revealing a trend to early mortality for new approaches. Most of these approaches (56.36%) seem to have been proposed by researchers without industry collaboration. Furthermore, they have not addressed relevant topics for the SCS community, such as integration between safety analysis and requirements-specification-based model development to support validation and verification [[21] and [49a], 54, 99, 136, 163], and certification process based on safety assurance [151]. These reasons may explain the premature death of these approaches. On the other hand, some researchers have pointed out the necessity of changing from traditional event-based accident modelling approaches to new ones [29, 103, 123, 143]. STAMP appears to be a promising approach in the direction of changing paradigms. It is based on the system-theoretic view of causality. However, STAMP was reported only in a few studies (3) in this SLR.

Need for industry validation. Despite the maturity of the SCS area, which seems to be greater than the IT mainstream (according to the rigor and relevance results presented in Section 4.2), the research in safety requirements for SCS still needs more validation within an industrial setting, i.e., the research community should develop more real cases with the participation of industry practitioners. The avionics domain reached the highest score when it came to industry relevance, with 51.85% of the studies getting a relevance score ⩾3, which is a good score. However, 79.07% of the studies classified as process-independent domain, in a total of 43 studies, got a relevance score <= 2. In relation to the participants, in more than 65% of the studies only researchers participated in the validation of the proposed approaches. Down-scaled industrial and toy problems were found in more than 57% of the studies that were tested or validated.

Lack of evidence for the usefulness and usability measurements. The reviewed studies revealed poor evidence with regard to how and why the usefulness and usability of the approaches were measured. The studies provide more evidence of why approaches were measured instead of how. They show much more evidence related to usefulness measurements than usability. The evidence related to the usability of the proposed approaches was quite poor, with 95% of studies not providing any evidence or indications related to how and why the usability was measured. The low level of participation among practitioners in the conducting of case studies and experiments to validate new approaches (discussed in Section 4.2.2) is linked to the lack of evidence of the usefulness and usability of the approaches. Despite the maturity of this area, our results seem to indicate that the SCS domain faces the same difficulties as other software engineering domains related to technology transfer from an academic to an industrial setting.

Communication process throughout the lifecycle. The majority of the reviewed studies (61%) reported approaches to support the requirements engineers and safety engineers during the elicitation, analysis and specification of the system hazards and safety requirements. However, these studies neither really focus on the communication issues between these actors nor how it could be improved by the proposed approaches. Thirteen out of 151 studies mentioned that the proposed approach supported in some way the development and certification teams throughout the SCS lifecycle, but even these studies provide little if any evidence about the impact of the proposed approaches over the communication process among the SCS project team actors. We found only three studies [83, 154, 166] that proposed approaches that explicitly addressed the improvement of the communication process among the actors in the SCS lifecycle. The lack of studies that investigate how to improve the communication issues in the SCS domain was a surprising finding, particularly considering the impact of such issues in the certification process, which is mandatory in the SCS industry.

Advice to SCS practitioners. Several new approaches have been proposed by the research community to help SCS practitioners in handling safety requirements. However, as we mentioned earlier, practitioners still seem to prefer traditional approaches. Practitioners should be open to try new approaches in order to improve the way they treat safety requirements, as well as providing a real evaluation of such approaches. Several innovative approaches identified in this SLR, such as STAMP, SafeML, SafeUML, GQM, GSN, RSML, among others, need more validation. These innovative approaches have the potential to help SCS practitioners in many important issues, such as in the integration between safety and requirements engineer teams, in improving the requirements communication throughout SCS lifecycle, as well as in the certification process. We recommend to the practitioners join forces with researchers to perform an extensive validation of such innovative approaches in large scale industrial settings. Additionally, we advise the project managers and the certification staffs to try some of the new approaches we have reported in this SLR. For example, STAMP has a large scope and can be instantiated to cover the whole SCS lifecycle, which has the potential to improve the requirements communication among the teams at the different levels of the organization. Besides, based on this new approach, a thorough safety program can be implanted in the companies interested in improving the safety of their products.

Education and Training. Even if this SLR is not directly related to training and education, it is possible to infer some points given the results. As discussed earlier, we found that practitioners prefer to use the traditional approaches, which may indicate that these traditional approaches are still being taught to the software engineering students (future practitioners), rather than new approaches. In order to get a seamless transition from the traditional to the new paradigms, it would be useful if both the traditional and the new approaches were taught for requirements engineering students. For example, introduce to the students the promising new approaches, such as STAMP, GSN and RSML, and compare them with the traditional ones. Requirements Engineering can assume different nuances based on what audience the teacher has. For computer science students who would become software developers, which are requirement consumers, Requirements Engineering teaching should focus on the importance of traceability among requirements, design artefacts and code, as well as its impact to software inspection and testing and consequently to the improvement of the software safety. For information technology students who usually become consultants in companies, e.g. business analysts and requirements specialists, which are requirement suppliers, Requirements Engineering teaching should emphasize the adoption of approaches to improve the structure of the requirements documentation and the requirements readability.

It should be observed that as the results indicate the level of actual validation and evidence of usefulness and usability, there is nothing to show that the older frameworks are better than the new ones.

Research agenda. Motivated by the results of this SLR we propose a research agenda for the SCS community, where a number of questions need to be considered, some examples are listed below:

(i)To what extent does the combination of traditional and new approaches improve the requirements communication throughout the SCS lifecycle?

(ii)What are the real difficulties and barriers related to industry practitioners changing from traditional event-based accident modelling approaches, to new ones, such as those based on the system-theoretic view of causality?

(iii)What are the common problems and conflicts between safety and security requirements approaches - in order to propose a common groundwork for developing new integrated approaches?

(iv)To what extent are the current safety standards supporting practitioners to improve the safety requirements in multi-domain systems?

(v)To what extent may the lean and agile requirements engineering approaches improve the integration among safety, requirements, test and certification teams?

(vi)To what extent can model-driven approaches help during requirements communication process throughout SCS lifecycle?

The results of this systematic review should provide insights and encourage further research into the design of studies to improve the requirements engineering for SCS, particularly in connection with the communication of the safety requirements among the project team actors, and the adoption of other models for hazard and accident models beyond the traditional event-based approaches. Additionally, we hope our results encourage researchers to incorporate a more empirical and industry-oriented base for the invention and creation of approaches, but especially in the evaluation of approaches.

Acknowledgements

This work was partially funded by the Federal University of São Paulo (UNIFESP), and the Brazilian research agency CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico - under grant agreement no. 249059/2013-04.

References

1. B.S. Medikonda , S.R. Panchumarthy , A framework for software safety in safe- ty-critical systems, SIGSOFT Softw. Eng. Notes 34 (2) (2009) 1–9 March .
2. D. Firesmith , Engineering safety requirements, safety constraints, and safety–critical requirements, J. Object Technol. 3 (3) (2004) 27–42 march/april .
3. D. Mellado , C. Blanco , L.E. Sánchez , E. Fernández-Medina , A systematic review of security requirements engineering, Comput. Stand. Interfaces 32 (4) (2010) 153–165 .
4. I. Iankoulova , M. Daneva , Cloud computing security requirements: a system- atic review, in: Sixth International Conference on Research Challenges in In- formation Science (RCIS), 2011, pp. 1–7 .
5. S. Yahya , M. Kamalrudin , S. Sideck , A review on tool supports for security requirements engineering, in: IEEE Conference on Open Systems, ICOS, 2013, pp. 190–194 .
6. K. Allenby , Deriving Safety Requirements Using Scenarios, in: Proceedings Fifth IEEE International Symposium on Requirements Engineering, 2001, pp. 228–235 .
7. C. Johnson, Forensic software engineering: are software failures symp- tomatic of systemic problems? Saf. Sci. 40 (9) (2002) 835–847, doi: 10.1016/ S0925-7535(01)0 0 086-8 .
8. M.P.E. Heimdahl , Safety and software intensive systems: challenges old and new, in: FoSE 2007: Future of Software Engineering, 2007, pp. 137–152 .
9. Kitchenham, B. and Charters, S. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering, Technical Report EBSE-2007-01, Software Eng. Group, Keele Univ. and Dept. of Computer Science, Univ. of Durham, UK.
10. Biolchini, J., Mian, P.G., Candida, A., Natali, C., and Travassos, G.H. (2005). Systematic Review in Software Engineering. Technical Report RT-ES 679/05, COOPE/UFRJ, (pp. 1–30).
11. M. Unterkalmsteiner , S. Member , T. Gorschek , A.K.M.M. Islam , C.K. Cheng , R.B. Permadi , R. Feldt , Evaluation and measurement of software process im- provement —a systematic literature review, IEEE Trans. Softw. Eng. 38 (2) (2012) 398–424 .
12. C. Heitmeyer, Managing complexity in software development with formally based tools, Electron. Notes Theor. Comp. Sci. 108 (2004) 11–19, doi: 10.1016/ j.entcs.2004.11 .
13. R. Lutz , Software engineering for safety: a roadmap, in: FoSE 20 0 0: Future of Software Engineering, 20 0 0, pp. 215–224 .
14. M. Ivarsson, T. Gorschek, A method for evaluating rigor and industrial rele- vance of technology evaluations, Empirical Softw. Eng. 16 (3) (2010) 365–395, doi: 10.1007/s10664- 010- 9146- 4 .
15. J. Hatcliff, A. Wassyng , T. Kelly , C. Comar , P. Jones , Certifiably safe soft- ware-dependent systems: challenges and directions, in: Proceedings of the on Future of Software Engineering - FOSE, 2014, pp. 182–200 .
16. N.G. Leveson , Safeware: System Safety and Computers, Addison-Wesley, Boston, 1995 .
17. N.G. Leveson, Safety as a system property, Commun. ACM 38 (11) (1995) 146, doi: 10.1145/219717.219816 .
18. C.M. Holloway , C.W. Johnson , Reducing our ignorance: finding answers to cer- tain epistemic questions for software systems, in: Proceedings of the 6th IET International Conference on System Safety, 2011, pp. 1–6 .
19. P. Rodríguez-Dapena , Software safety certification: a multidomain problem, IEEE Softw. Mag. 16 (4) (1999) 31–38 .
20. A. Van Lamsweerde, Requirements engineering in the year 00: a research per- spective, in: Proceedings of the International Conference on Software Engi- neering (ICSE), 20 0 0, pp. 5–19, doi: 10.1109/ICSE.20 0 0.870392 .
21. B.S. Andersen, G. Romanski, Verification of safety-critical software, Commun. ACM 54 (10) (2011) 52–57, doi: 10.1145/20 01269.20 01286 .
22. R. Bell , Introduction to IEC 61508, in: Proceedings of the ACS Workshop on Tools and Standards, 2005, pp. 1–15 .
23. T. Cant , B. Mahony , B. Atchison , Revision of Australian defence standard def (Aust.) 5679, in: Proceedings of the 10th Australian Workshop on Safety-Re- lated Programmable Systems, vol. 56, 2004, pp. 85–94 .
24. D. Kopec , S. Tamang , Failures in complex systems: case studies, causes, and possible remedies, SIGCSE Bull. 39 (2) (2007) 180–184 .
25. A. Kornecki, J. Zalewski, Certification of software for real-time safety-critical systems: state of the art, Innov. Syst. Softw. Eng. 5 (2) (2009) 149–161, doi: 10. 10 07/s11334-0 09-0 088-1 .
26. J. Hatcliff, M. Heimdahl, M. Lawford, T. Maibaum, A. Wassyng, F. Wurden, A software certification consortium and its top 9 hurdles, Electron. Notes Theor. Comput. Sci. 238 (4) (2009) 11–17, doi: 10.1016/j.entcs.2009.09.002 .
27. M.V. Zelkowitzayb , D. Wallace , Experimental validation in software engineer- ing, Inf. Softw. Technol. 39 (11) (1997) 735–743 .
28. S.L. Pfleeger , B.A. Kitchenham , Principles of survey research part 1: turning lemons into lemonade, ACM SIGSOFT Softw. Eng. Notes 26 (6) (2001) 16–18 .
29. N.G. Leveson , Engineering a Safer World: Systems Thinking Applied to Safety, The MIT Press, Cambridge, 2011 .
30. A. Hall , Seven myths of formal methods, IEEE Softw. Mag. 7 (5) (1990) 11–19 .
31. J. Abrial , Formal methods in industry: achievements, problems, future, in: Proceedings of the IEEE International Conference on Software Engineering, 2006, pp. 761–767 .
32. P. Runeson , M. Höst ,Guidelines for conducting and reporting case study re- search in software engineering, Empirical Softw. Eng. 14 (2009) 131–164 .
33. T. Saracevic , Evaluation of evaluation in information retrieval, in: Proceedings of the 18th International Conf. on Research and Development in Information Retrieval, 1995, pp. 138–146 .
34. M. Ivarsson , T. Gorschek , Technology transfer decision support in require- ments engineering research: a systematic review of REj, Requirements Eng. J. 14 (3) (2009) 155–175 .
35. C. Wohlin , Experimentation in Software Engineering: An Introduction, Kluwer Academic, Heidelberg, 20 0 0 .
36. S. Fricker , T. Gorschek , C. Byman , A. Schmidle , Handshaking with implemen- tation proposals: negotiating requirements understanding, IEEE Softw. 27 (2) (2010) 72–80 .
37. K.J. Hayhurst , C.M. Holloway , Challenges in software aspects of aerospace sys- tems, in: Proceedings of Annual NASA Goddard Software Engineering Work- shop, 2001, pp. 7–13 .
38. W.S. Greenwell , J.C. Knight , C.M. Holloway , J.J. Pease , A taxonomy of falla- cies in system safety arguments, in: Proceedings of the International System Safety Conference, 2006, pp. 1–10 .
39. P. Lewycky , Notes toward an understanding of accident causes, Hazard Pre- vention 23 (2) (1987) 6–8 .
40. B. Kitchenham , S. Linkman , D. Law , DESMET: a methodology for evaluat- ing software engineering methods and tools, Comput. Control Eng. J. (1997) 120–126 .
41. B. Kitchenham , Evaluating software engineering methods and tools part 1: the evaluation context and evaluation methods, in: ACM SIGSOFT Software Engineering Notes, vol. 21, 1996, pp. 11–15 .
42. E.M. Rogers , Diffusion of Innovations, fifth ed., Free Press, New York, 2003 .

43. S. Fricker , T. Gorschek , P. Myllyperkiö, Hand-shaking between software projects and stakeholders using implementation proposals, in: Proc. 13th Int'l Working Conf. Requirements Eng.: Foundation for Software Quality (REFSQ 07), 2007, pp. 144–159 .
44. E. Bjarnason , K. Wnuk , B. Regnell , Requirements are slipping through the gaps - a case study on causes & effects of communication gaps in large-scale soft- ware development, in: 19th International Requirements Engineering Confer- ence, 2011, pp. 37–46 .
45. P. Legris , J. Ingham , P. Collerette , Why do people use information technology? A critical review of the technology acceptance model, Inf. Manag. 40 (2003) 191–204 .
46. F.D. Davis , User acceptance of information technology: system characteristics, user perceptions, and behavioral impacts, Int. J. Man Mach. Stud. 38 (1993) 475–487 .
47. B. Ramesh , L. Cao , R. Baskerville , Agile requirements engineering practices and challenges: an empirical study, Inf. Syst. J. 20 (5) (2010) 44 9–4 80 .
48. S. Nair , J.L. de la Vara , M. Sabetzadeh , L. Briand , An extended systematic liter- ature review on provision of evidence for safety certification, Inf. Softw. Tech- nol. 56 (2014) 689–717 .
49. S. Nair , J.L. de la Vara , M. Sabetzadeh , D. Falessi , Evidence management for compliance of critical systems with safety standards: A survey on the state of practice, Inf. Softw. Technol. 60 (2015) 1–15 .

3D Bioprinting for Engineering Complex Tissues

Christian Mandrycky [1,*], Zongjie Wang [2], Keekyoung Kimb [2,*], Deok-Ho Kim [1]

[1] Department of Bioengineering, University of Washington, Seattle, WA 98195, USA

[2] School of Engineering, The University of British Columbia, Kelowna, BC V1V 1V7, Canada

* Corresponding Author. E-mail: keekyoung.kim@ubc.ca

Abstract. Bioprinting is a 3D fabrication technology used to precisely dispense cell-laden biomaterials for the construction of complex 3D functional living tissues or artificial organs. While still in its early stages, bioprinting strategies have demonstrated their potential use in regenerative medicine to generate a variety of transplantable tissues, including skin, cartilage, and bone. However, current bioprinting approaches still have technical challenges in terms of high-resolution cell deposition, controlled cell distributions, vascularization, and innervation within complex 3D tissues. While no one-size-fits-all approach to bioprinting has emerged, it remains an on-demand, versatile fabrication technique that may address the growing organ shortage as well as provide a high-throughput method for cell patterning at the micrometer scale for broad biomedical engineering applications. In this review, we introduce the basic principles, materials, integration strategies and applications of bioprinting. We also discuss the recent developments, current challenges and future prospects of 3D bioprinting for engineering complex tissues. Combined with recent advances in human pluripotent stem cell technologies, 3D-bioprinted tissue models could serve as an enabling platform for high-throughput predictive drug screening and more effective regenerative therapies.

Keywords: Bioprinting; Bioink; Tissue engineering; 3D printing; Hydrogel; Drug screening; Regenerative medicine.

1. Introduction

In the United States, one name is added to the organ transplant waiting list every 15 min (Abouna, 2008). While this list grows rapidly, less than one-third of waiting patients can receive matched organs from donors (Ozbolat and Yu, 2013). This growing deficit, however, is unlikely to be met by a supply of transplantable organs that has stagnated over the last decade (Bajaj et al., 2014). One of the most promising techniques to alleviate this organ shortage crisis is tissue engineering, the use of a combination of cell, engineering, and material methods to generate artificial tissues and organs (Langer and Vacanti, 1993). In tissue engineering, three strategies are used to replace or induce targeted tissues: (1) the use of cells alone, (2) the use of biocompatible biomaterials, (3) the use of a combination of both cells and biomaterials (Khademhosseini et al., 2006). These cells and biomaterials are combined into scaffolds through a variety of processes, which can generally be classified as either top-down, or bottom-up. In top-down approaches cells are often seeded sparsely and homogenously in biomaterials shaped to resemble biological geometries. On the other hand, in bottom-up approaches modular units of cells and biomaterials are combined to form macro tissues. Top-down methods have been in wide use for years, however, these methods often cannot accurately control the distribution of cells, and fail to generate the appropriate extracellular matrix (ECM) (Khademhosseini et al., 2006). Without a proper ECM microenvironment, cells cannot function as tissues properly. This limitation is addressed in bottom-up approaches that build up tissues brick by brick via micro- and nano-technologies. As a result, cell distribution can be defined at the micrometer scale, which significantly improves the controllability of scaffold fabrication (Jiao et al., 2014). Motivated by developments in nanotechnology, techniques like self-assembly and soft-lithography have been applied to bottom-up tissue engineering (Kim et al., 2013, Kim et al., 2014a and Shapira et al., 2014). Among the micro-scale bottom-up techniques recently applied to tissue engineering, bioprinting, a form of additive manufacturing, has become one of the most promising and advanced fabrication methods (Table 1).

In bioprinting, small units of cells and biomaterials are dispensed with micrometer precision to form tissue-like structures (Fig. 1). Unlike conventional 3D printing techniques that have been used to print temporary cell-free scaffolds for use in surgery (Bracci et al., 2013), bioprinting requires a different technical approach that is compatible with depositing living cells. The advantages of bioprinting include accurate control of cell distribution, high-resolution cell deposition, scalability, and cost-effectiveness. For those reasons, the development and subsequent applications of bioprinting have greatly increased during the last five years. In this review, we discuss the basic principles of bioprinting, including bioprinter device design, workflow, biomaterial options, and current and potential applications.

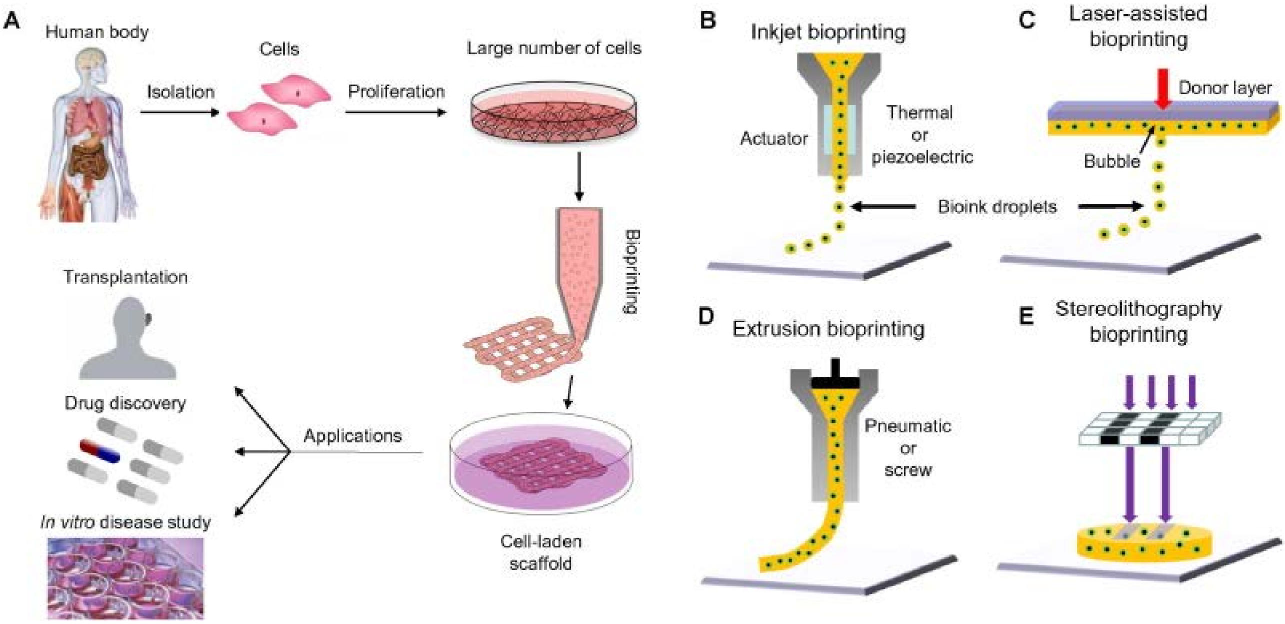


Fig. 1. Bioprinting process, techniques, and applications. (A) For human therapeutic applications, the typical workflow of bioprinting would involve the isolation and expansion of human cells prior to printing the desired cell-laden scaffold. These scaffolds could then ultimately be used as therapeutic devices themselves, as a testing platform for drug screening and discovery, or as an in vitro model system for disease. (B) Inkjet printers eject small droplets of cells and hydrogel sequentially to build up tissues. (C) Laser bioprinters use a laser to vaporize a region in the donor layer (top) forming a bubble that propels a suspended bioink to fall onto the substrate. (D) Extrusion bioprinters use pneumatics or manual force to continuously extrude a liquid cell–hydrogel solution. (E) Stereolithographic printers use a digital light projector to selectively crosslink bioinks plane-by-plane. In (C) and (E), colored arrows represent a laser pulse or projected light, respectively.

2. Bioprinting techniques

To date, no single bioprinting technique enables the production of all scales and complexities of synthetic tissues. The three major bioprinting techniques of inkjet, laser-assisted, and extrusion bioprinting each have specific strengths, weaknesses, and limitations. A concise comparison of these approaches is also provided in Table 2.

2.1. Inkjet printing

Inkjet bioprinting was the first bioprinting technology (Tuan et al., 2003) and is very similar to conventional 2D inkjet printing (Singh et al., 2010). A hydrogel pre-polymer solution with encapsulated cells (called a bioink) is stored in the ink cartridge. The cartridge is then connected to a printer head and acts as the bioink source during the electronically controlled printing process. During printing, the printer heads are deformed by a thermal or piezoelectric actuator and squeezed to generate droplets of a controllable size, as shown in Fig. 1B. The advantages of inkjet printing include: (1) low cost due to similar structure with commercial printers, (2) high printing speed conferred by the ability of the printer heads to support parallel work mode, and (3) relatively high cell viability (usually from 80% to 90%), as determined by many experimental results (Cui et al., 2012a, Cui et al., 2012b and Cui et al., 2013).

However, because current printer heads are based on microelectromechanical system (MEMS) devices, there is a relatively small deformation generated by either thermal or piezoelectric actuation at the nozzle opening. As a result, MEMS-based printer heads cannot squeeze out high viscosity materials (> 15 mPa/s) and do not work well with bioinks with high cell density (> 1 × 106 cells/mL). High cell density increases the average viscosity of bioinks, resulting in clogging of the head (Xu et al., 2005, Guillotin et al., 2010, Pepper et al., 2011 and Pepper et al., 2012). Recent research has highlighted another disadvantage of inkjet printing, named the settling effect (Pepper et al., 2011 and Pepper et al., 2012). When bioinks are initially loaded into the ink cartridge, they are well mixed. Over the entire printing process, however, cells begin to settle in the cartridge, increasing the viscosity of the bioink and often clogging the printer head.

The simplest way to build inkjet bioprinter is to modify a commercial printer. HP 26 printer heads (Hewlett-Packard, Palo Alto, USA) were combined with a controller to print bioinks (Mattimore et al., 2010). Similar print heads were further integrated with a modified HP G3110 scanner (Hewlett-Packard, Palo Alto, USA) to build a low-cost bioprinter (~$700) (Orloff et al., 2014). Such a low-cost system was achieved by using commercial print heads as the dispenser, a scanner as a 2-axis servo stage, and free control software. The resolution of the servo stage was approximately 500 μm, however, which is too coarse for micro-positioning. Additionally these print heads and cartridges are not capable of storing enough bioink to print large tissues, limiting the applications of this simple bioprinter.

Many efforts have been made to improve the stage resolution and enlarge the reservoir capacity. Screw-based servo stages with less than the 100 μm resolution in each direction were used to provide sub-micrometer positioning (Nishiyama et al., 2009 and Arai et al., 2011). External jugs and bottles were modified and connected to multiple print heads to increase the maximum bioink capacity. After adopting high accuracy stages and larger reservoirs, this inkjet bioprinter was able to achieve 10 μm positioning accuracy and 20 picoliter droplet volume.

2.2. Laser-assisted printing

Laser-assisted printing originated from laser direct-write (Bohandy et al., 1986) and laser-induced transfer technologies (Duocastella et al., 2007 and Kattamis et al., 2007). Fig. 1C shows a schematic of laser-assisted printing. The critical part of the laser-assisted printing system is a donor layer that responds to laser stimulation. The donor layer comprises a ‘ribbon’ structure containing an energy-absorbing layer (e.g., titanium or gold) on the top and a layer of bioink solution suspended on the bottom. During printing, a focused laser pulse is applied to stimulate a small area of the absorbing layer. This laser pulse vaporizes a portion of the donor layer, creating a high-pressure bubble at the interface of the bioink layer and propelling the suspended bioink. The falling bioink droplet is collected on the receiving substrate and subsequently crosslinked. Compared to inkjet printing, laser-assisted printing can avoid direct contact between the dispenser and the bioinks. This non-contact printing method does not cause mechanical stress to the cells, which results in high cell viability (usually higher than 95%). In addition, laser-assisted printing can also print highly viscous materials, and more types of bioinks can be used than in inkjet printing. These features of laser bioprinting are promising, but the side effects of laser exposure on the cell are not yet fully understood. Moreover, laser diodes with high-resolution and intensity are expensive compared to other nozzle-based printing methods, and control of the laser printing system is complex, limiting the technique's adoption.

Due to the high cost, there are few laser-assisted bioprinters, which are usually cumbersome and complex compared to other types of printers. A laser printing prototype was developed by combining optical laser sources with a lens (Nahmias et al., 2005). A more compact, high-throughput laser printing system was also built (Guillemot et al., 2010) and this system was further developed into a highly accurate version with 10 μm resolution (Guillotin et al., 2010). In addition to the high equipment cost, laser-assisted printing is still immature because of unexplored parameters affecting the droplet size and quality. Instead of building prototypes of laser-assisted bioprinters, more researchers have focused on investigating the relationships between laser parameters, such as wavelength, intensity, and pulse time, with the quality of printed patterns (Duan et al., 2013 and Duarte Campos et al., 2013).

2.3. Extrusion printing

Extrusion printing is a modification of inkjet printing. In order to print the viscous materials inkjet printers cannot deposit, extrusion printing uses either an air-force pump or a mechanical screw plunger to dispense bioinks, as shown in Fig. 1D. By applying a continuous force, extrusion printing can print uninterrupted cylindrical lines rather than a single bioink droplet. Almost all types of hydrogel pre-polymer solutions of varying viscosity as well as aggregates with high cell density can be printed with extrusion bioprinters. While extrusion bioprinters can print a wider range of materials, they also expose the encapsulated cells to larger mechanical stresses that are thought to reduce cell viability (Khalil and Sun, 2007 and Murphy and Atala, 2014).

Most existing commercial bioprinters, including the Bioplotter (EnvisionTec, Gladbeck, Germany) and NovoGen 3D Bioprinting platform (Organovo, San Diego, USA), are based on extrusion technology. Extrusion bioprinting provides good compatibility with photo, chemical and thermal crosslinkable hydrogels of very different viscosities at a reasonable cost (Khalil and Sun, 2007 and Murphy and Atala, 2014). A typical extrusion printer, the multi-head tissue/organ building system from the Cho group, includes three-axis motion control with six dispensing heads, supporting up to six different bioinks (Lee et al., 2014). The substrate plate contains heating and cooling functions to control thermally sensitive hydrogels. Similar designs have been reported by two other groups (Chang et al., 2010 and Bertassoni et al., 2014a). The latest versions of extrusion printers include tissue–vessel parallel printing (Ozbolat et al., 2014) and parallel multi-bioink printing (Kolesky et al., 2014).

Dispensers in current extrusion systems have a few differences (Khalil and Sun, 2007). Pneumatic micro nozzles powered by compressed gases support a wider range of viscosity, but have difficulty precisely controlling the deposited mass. Screw-based nozzles can print without inlet air and are much cheaper, but they experience problems in high viscosity dispensing.

2.4. Other technical approaches

While these three printing methods are most commonly used by bioprinting researchers, the bioprinting paradigm itself has been challenged and novel printing methodologies remain under investigation. Rather than directly printing tissues, Miller et al. (2012) used a pneumatically controlled syringe to print molten sugar glass in the shape of a desired vascular network. Once printed, this artificial vascular network was embedded within a variety of hydrogels and could then be dissolved to form open channels within cross-linked tissues. While this approach sacrificed the ability to carefully control the deposition of cells within the bulk matrix, it enabled previously unachieved engineered vascular complexity in a synthetic tissue.

Stereolithography has also been modified for bioprinting purposes (Fig. 1E) (Gauvin et al., 2012 and Gou et al., 2014). Like laser-assisted printing, stereolithography bioprinters use light to selectively solidify a bioink in a layer-by-layer process that additively builds up objects (Fig. 2A). These printers use a digital light projector to cure bioinks plane-by-plane and have several advantages over traditional bioprinting methods. No matter how complex pattern in one layer is, the printing time is the same because the entire pattern is projected over the printing plane. As a result, the printer only needs a moveable stage in a vertical direction, which significantly simplifies the control of the printer. This reported stereolithography bioprinting system can achieve 100 μm resolution and printing times less than 1 h (Gauvin et al., 2012 and Gou et al., 2014) while maintaining very high cell viability (> 90%). Fig. 2B–E show the woodpile and hexagonal structures printed by the stereolithography system. The fluorescent images of the hexagonal structures encapsulated with human umbilical vein endothelial cells (HUVECs) are given in Fig. 2F–H. Recently, a commercial beam projector was adopted as the light source, providing an inexpensive solution (< $1,000) for stereolithography bioprinting (Wang et al., 2015c). This system was able to print hydrogel patterns with 50 μm resolution. A recent advance in stereolithographic 3D printing technology by the DeSimone group (Tumbleston et al., 2015) referred to as “continuous liquid interface production (CLIP)” dramatically improved both resolution and printing time for some materials. While this has not yet been applied to bioprinting, it may be an approach that enables the formation of more complex tissue architectures.

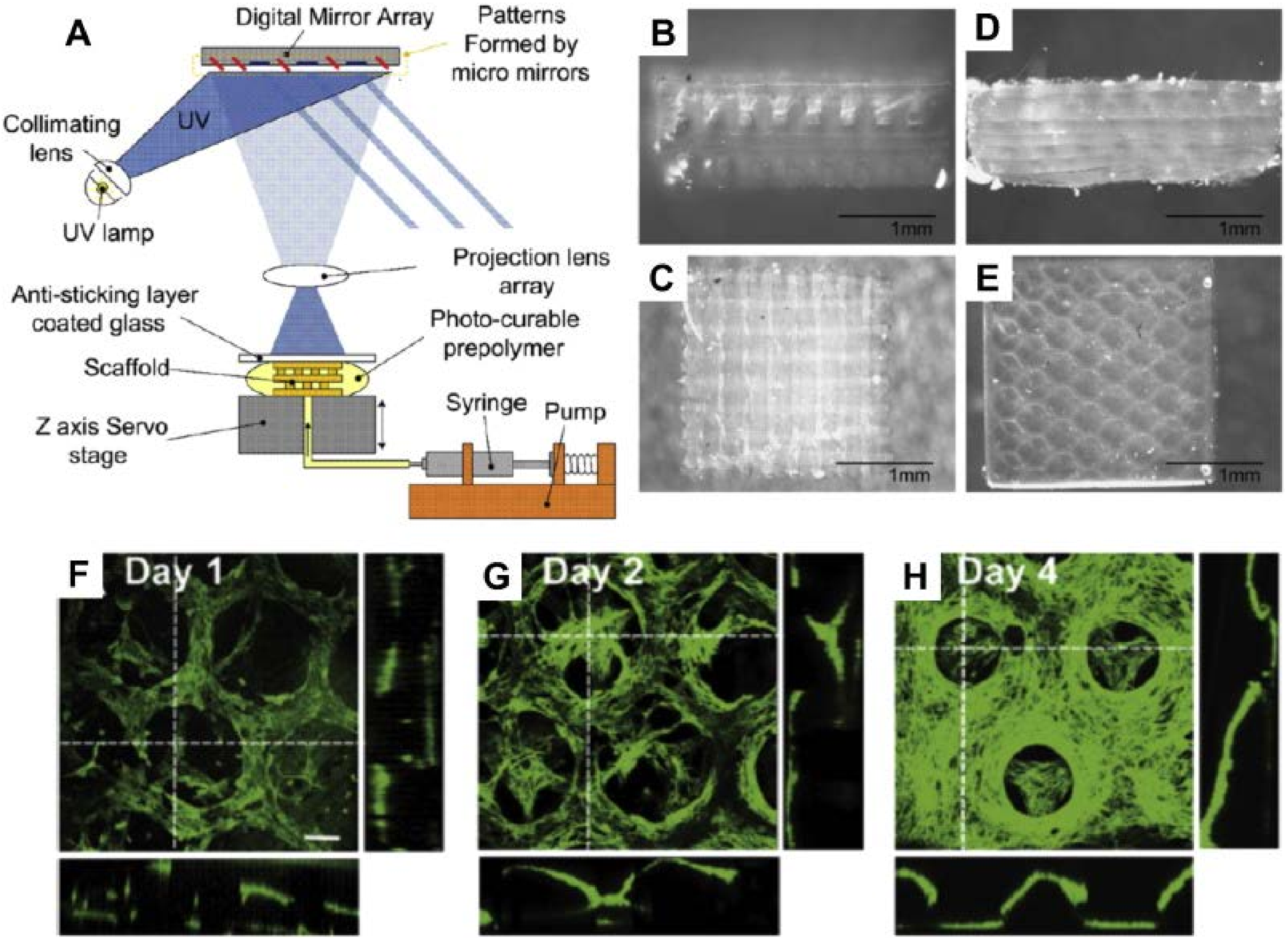


Fig. 2. Stereolithographic bioprinting. (A) Schematic illustration of a stereolithography system (Gauvin et al., 2012). (B)–(E) The side view of woodpile (B) and hexagonal (D), as well as the top view of woodpile (C) and hexagonal (E) structures generated by their stereolithography system. (F)–(H) 3D confocal images showing the proliferation of encapsulated HUVEC cells in day 1 (F), day 2 (G) and day 4 (H) (scale bar: 100 μm).

2.5. Bioprinting CAD, modeling, and the printing process

Bioprinters cannot print without instructions. To successfully create bioprinted tissues, it is necessary to generate the printing paths, select appropriate bioinks, control the bioprinter and perform quality control after printing (Murphy and Atala, 2014). The typical bioprinting process is as follows: (1) designers draw the printing geometry and manually verify its feasibility; (2) designers select appropriate cell types and hydrogels, and load the bioinks into the bioprinting system (3) through control language and protocols, such as RS 274 (G-Code, Massachusetts Institute of Technology, Cambridge, USA) and LabView (National Instruments, Austin, USA), the designed paths are sent to the bioprinting system; (4) the bioprinter builds structures by depositing bioinks under the control of a computer; (5) bioprinted tissues are checked manually via microscopy after bioprinting. After the bioprinting process, successfully printed constructs are transferred to an incubator for culturing. The bioprinting process is not currently highly automated and many manual operations at a variety of steps can result in slow processing speeds and increase the chance for mistakes and errors. To ensure printing quality and to improve the printing process, many researchers have investigated computer-aided design (CAD) and modeling technology for bioprinting. These CAD techniques can utilize computer automation systems to assist and accelerate the design process.

Bioprinting models, like models used in conventional rapid prototyping, are often converted to the STereoLithography (STL) file format as an intermediate between model and print path generation (Mironov et al., 2009 and Mondy et al., 2009). These files contain accurate surface information of complex 3D geometries, and can be designed via graphic user interfaces, or created from clinical images, including those from magnetic resonance imaging (MRI) and computed tomography (CT) (Keriquel et al., 2010 and Arai et al., 2011). In a process analogous to histologic sectioning, printing paths are created by "slicing" these STL model into layers and creating bioprinter toolpaths that trace out the perimeter and interior features of each slice. The thickness of these layers is often referred to as the resolution of a particular printer and is usually in the range of 100–500 μm depending on the machine and material used. These toolpaths are the instructions read and executed by the bioprinter for each layer and can include material selections. Layers are formed sequentially and stacked as the model is built up in an additive process forming a 3D object from a collection of 2D layers. All other things being equal, smaller resolutions are associated with higher quality and longer print times.

Clinical images can provide information regarding the in vivo tissue distribution of patients, and anatomically realistic tissue geometries can be determined via image processing. Clinical image-based STLs therefore have the potential to become the starting point for on-demand tissue production in the future. In addition, a smart program was coded for planning and optimizing bioprinting experimental design (Weiss et al., 2005). In summary, the introduction of Bio-CAD techniques has significantly improved the automation of bioprinting path generation.

Computer aided techniques, known as bio-computer-aided-manufacturing (Bio-CAM), also play an important role during and after bioprinting. Bio-CAM aims to predict the feasibility of the fabrication process by simulating relevant physical models on computers. To simulate bioprinting, both classical formula calculations and the finite element method (FEM) are applied. Currently, the most widely used physical model for bioprinting is laminar multi-phase flow. Although this model is oversimplified, ignoring complex issues generated by the inclusion of cells, simulations are still helpful for checking and optimizing the feasibility of specific designs. Many researchers are already attempting to model bioprinting results with the corresponding printing parameters. For extrusion printing, relationships between dispensing pressure, printing time, and nozzle diameter have been tested and modeled (Yu et al., 2013). Cell settling effects in inkjet printers, which are highly related to clogging and viscosity, change during printing and were modeled by both analytical and finite element methods (FEM) (Pepper et al., 2011 and Pepper et al., 2012). For laser printing, the effects of laser energy, substrate film thickness, and hydrogel viscosity on the viability of cells (Catros et al., 2011b), as well as droplet size (Duocastella et al., 2007, Mézel et al., 2010 and Gruene et al., 2011b), cell differentiation (Gruene et al., 2011a), and cell proliferation (Gruene et al., 2011a) have been investigated. Some researchers also focused on post-printing modeling of cellular dynamics (McCune et al., 2014), fusion (Yang et al., 2012, Yang et al., 2013, Sun and Wang, 2013 and Thomas et al., 2014), deformation (Sim et al., 2007) and stiffness (Tirella et al., 2011 and Mobed-Miremadi et al., 2012), as well as modeling of the typical types of printed tissues, including tumors (Zhao et al., 2014) and soft tissues (Zhang et al., 2013). Bio-CAM research not only provides a fast way to check design feasibility, but also gives designers a chance to better understand the physical and chemical principles governing printing. With the integration of Bio-CAD and Bio-CAM, an advanced design flow for bioprinting begins to take shape. Bio-CAD can accelerate the speed of the whole bioprinting process, and Bio-CAM can guarantee the quality of what is printed.

3. Materials for bioprinting

Bioinks typically consist of a hydrogel pre-polymer solution and cells. The desired properties of hydrogels are presented at the beginning of this section, and the characteristics of various types of crosslinkable hydrogels are summarized. Resources for the cells and materials used in bioprinting applications are briefly reviewed at the end.

3.1. Hydrogel bioink characteristics

Hydrogels play an essential role in bioprinting. They not only have direct contact with cells to provide structural support, but they also dominate the chemical and physical properties of bioinks (Williams, 2008). Ideally, hydrogels used for bioprinting should be characterized by the properties described below.

3.1.1. Printability and crosslinkability

Printability refers to the relationships between bioinks and substrates that results in printing an accurate, high-quality pattern (Murphy and Atala, 2014). In bioprinting, printability is usually associated with surface tension, which is measured by the contact angles between two media. Research has shown that the surface tension of supporting structures has significant and profound implications on cell attachment and development (Discher et al., 2005). To form 3D scaffolds, the printed hydrogel pre-polymer solution should not be too flat on the substrate. This means that the hydrogel pre-polymers are expected to maintain tension in the vertical direction and have a large contact angle with the substrate. Since glass slides and petri dishes are the most commonly employed substrates, ideal hydrogel pre-polymer solutions should be able to build highly vertical structures after printing on glass and plastic substrates. Unfortunately, most of the glass slide substrates have poor contact angles. This problem can be solved by coating the substrates with a thin layer of material, such as 3-(trimethoxysilyl)propyl methacrylate (TMSPMA) (Zhang et al., 2008), to enhance their hydrophobicity before printing (Bauer et al., 2012 and Nikkhah et al., 2012).

Printability is also influenced by how easily materials can be crosslinked. The three types of bioprinting technologies currently available are only capable of dispensing liquid materials and consequently hydrogels must be in liquid or paste-like form during printing. To accommodate different cell densities and printing technologies, the viscosity of the hydrogel pre-polymer solutions should be controllable over a wide range. At odds with this condition, bioinks must form a quasi-scaffold structure to support cell proliferation after printing. These conditions have effectively limited hydrogel pre-polymer solutions to either photo (Weiner et al., 2007 and Nichol et al., 2010), chemically (Li et al., 2005, Glowacki and Mizuno, 2008, Liu et al., 2009, Balakrishnan et al., 2012 and Araujo et al., 2014), or thermally (Gao et al., 2012 and Wu et al., 2012) crosslinkable polymers (Murphy et al., 2013 and Bajaj et al., 2014).

3.1.2. Mechanical properties

Hydrogels should maintain sufficient mechanical properties after polymerization to provide the cells with a stable environment for attachment, proliferation and differentiation (Limpanuphap and Derby, 2002 and Murphy et al., 2013). These mechanical properties include strain, shear stress, compressive modulus and mass swelling ratio. It is well understood that cell adhesion is significantly affected by the dynamic interactions between cells and hydrogels (Dou et al., 2012 and Benson et al., 2014). In fact, mechanical properties are considered to be highly essential for soft tissues, such as cartilage and skin, because the functions of such tissues mainly rely on their mechanical properties (Hutmacher, 2000 and Kim et al., 2012).

3.1.3. Biocompatibility and controllability of by-products and degradation

Biocompatibility refers to the ability of a material to perform with an appropriate host response in a specific situation (Hobkirk, 1988). In general, for in vitro applications, biocompatibility requires that the material itself is not harmful to cell proliferation and has the ability to provide proper binding with cells (Williams, 2008). For in vivo applications, biocompatibility adds the requirement that the material can be degraded by or integrated with the ECMs of cells without generating harmful by-products or having negative interactions with cells (Williams, 2008). It is desirable for implanted tissue to eventually fuse with other in vivo tissues. Therefore, hydrogel scaffolds need to be degraded or integrated with the in vivo ECM environment and hydrogels with a natural and controllable degradation rate which is similar to the ECM growth rate is highly desired (Murphy and Atala, 2014).

3.2. Bioinks

From the perspective of hydrogel design, there are basically two types of hydrogels: those based on natural polymers and those based on synthetic polymers (Zorlutuna et al., 2013). Natural hydrogels include polymers existing in ECM components, such as gelatin, collagen, laminin and fibronectin, as well as other natural polymers such as alginate, chitosan and silk fibroin. Interactions between natural hydrogels and cells have been well investigated (Zorlutuna et al., 2013). Synthetic polymers, unlike natural polymers, are made through chemical synthesis and are typically more controllable in terms of their chemical and mechanical properties (Zhu and Marchant, 2011). Their interactions with and effects on cells, however, have not yet been studied systematically (Zorlutuna et al., 2013). Natural polymers are widely used in bioprinting research, but some researchers have used a combination of natural and synthetic polymers (Schuurman et al., 2011 and Xu et al., 2013).

Decellularized extracellular matrices have been an increasingly promising material in tissue engineering as decellularization protocols have steadily improved. Recently, Pati et al. (2014) showed that dECMs from three tissues could be solubilized into bioinks and bioprinted. While many bioinks are compositionally simple, dECM bioinks contain the diverse array of ECM components characteristic of different tissues and as a result, more closely resemble the native tissue. Although the mechanical properties of dECM bioinks do not mirror the original tissue, they represent a promising addition to the bioinks available in bioprinting.

On the synthetic side, there is significant interest in developing conductive biomaterials (Balint et al., 2014). Recently Jakus et al. (2015) developed a printable high-content graphene: polyactide-co-glycolide bioink with high conductivity (800 S/m). Scaffolds printed with this material were able to support the growth of human mesenchymal stem cells (hMSCs) and have interesting possible applications in both biomedical devices and biologic scaffolds where enhanced conductivity is desirable. For example, conductive tracks through scaffolds could be pre-patterned in printed tissues simply by changing the bioink. This would complement a previously demonstrated method for installing these tracks by using α-hemolysin containing droplets (Villar et al., 2013). For a more complete discussion of existing biomaterials for bioprinting as well as the interaction and trade-off between desired hydrogel properties, we refer the reader to recent reviews (Bajaj et al., 2014 and Skardal and Atala, 2014).

3.3. Cells

To form a highly mimetic tissue or organ on a macro scale, bioprinted cells must proliferate. Two main factors are considered when selecting cells for bioprinting: how closely the bioprinted cells can mimic the physiological state of cells in vivo, and to what degree the bioprinted cells can maintain or develop their in vivo functions under optimized microenvironments (Murphy and Atala, 2014). Artificial tissues are seeded by either printing functional primary cells with supporting cells (Keriquel et al., 2010, Cui et al., 2012a, Duan et al., 2013, Michael et al., 2013, Xu et al., 2013, Zhang et al., 2013 and Dolati et al., 2014) or printing progenitors or stem cells for further differentiation (Gruene et al., 2011a, Xu et al., 2011, Duarte Campos et al., 2013, Hong et al., 2013, Owens et al., 2013 and Visser et al., 2013). Direct printing of primary cells can rapidly increase the complexity of bioprinting. Since multiple types of cells embedded within the same or different hydrogels need to be printed in parallel, many bioinks need to be prepared for each print. Real-time alignment and printing step control are complicated by using many bioinks as each switch between bioinks has the possibility to introduce error into the bioprinting process. Printing with stem cells will usually reduce the total number of bioinks used for a given print, but also adds its own set of complications. Additional bioink formulations with different growth factor and small molecule signals may be desirable to attempt to guide site-specific differentiation. Even without this kind of approach, there is added difficulty in post-printing culture as growth factors and other differentiation stimulators must be deposited precisely to ensure the control of differentiation, especially when vascularization is desired.

Reliable cell sourcing poses a perennial problem to bioprinting. For clinical applications, cells for bioprinting would ideally be isolated from the patients themselves to avoid negative immune responses (Ozbolat and Yu, 2013). Because not all types of cells can regenerate after damage (e.g. cardiac muscle cells), stem cells (e.g. adipose derived stem cells) with the ability to proliferate and differentiate into the desired cell types are the most promising cell source. Examples of some of the cell types and organ systems targeted by recent bioprinting publications are presented in Table 3.

4. Applications of bioprinting

In this section, the current applications of bioprinting are reviewed in terms of several popular tissue types and its role in drug screening.

4.1. Vessels

While the ability to create vascular features in bioprinted tissues is often limited, novel bioprinting techniques may resolve this problem. Dolati et al. (2014), for example, utilized a coaxial nozzle system to print vascular conduits more than a meter long (Fig. 3A). These carbon nanotube reinforced alginate conduits were perfusable and supported the growth of human coronary artery smooth muscle cells within the matrix. Using this technique, the authors were able to fabricate conduits with diameters in the sub-millimeter range, but did not show an ability to print closer to capillary diameters. Another possible solution is to add magnetically controlled nanoparticles to bioinks and use these to print vessels. With this technique the position of the vessels within tissues could then be controlled by applying a magnetic field (Mironov et al., 2008 and Talelli et al., 2009). However, further research is needed to determine the efficiency and the potential effects of magnetic particles on cells and ECM. To reduce the size of vascular channels and to incorporate them directly into printed tissues, others have employed sacrificial inks to some success. Kolesky et al. (2014) used a Pluronic F127 fugitive bioink to print channels as small as 45 μm and were able to subsequently endothelialize them with HUVECs. This approach, combined with printing fibroblasts encapsulated in a gelatin methacrylate bioink, yielded multicellular bioprinted constructs (Fig. 3B). Once the constructs were printed and crosslinked, the temperature was lowered to 4 °C to liquefy and remove the Pluronic F127, leaving behind open vascular channels ready to be seeded. Previously, Miller et al. (2012) encapsulated and dissolved printed carbohydrate glass in various bulk extracellular matrices to form seedable channels as small as 150 μm (Fig. 3C). Rather than dissolving away the sacrificial material, Bertassoni et al. (2014b) cast hydrogels around printed agarose fibers and then aspirated or manually removed the fiber. The resulting lumen were perfusable and HUVECs could form an endothelial monolayer. These sacrificial techniques are exciting advances that may simplify not only the prepatterning of vascular features in bioprinted tissues, but also the speed at which large tissues can be printed.

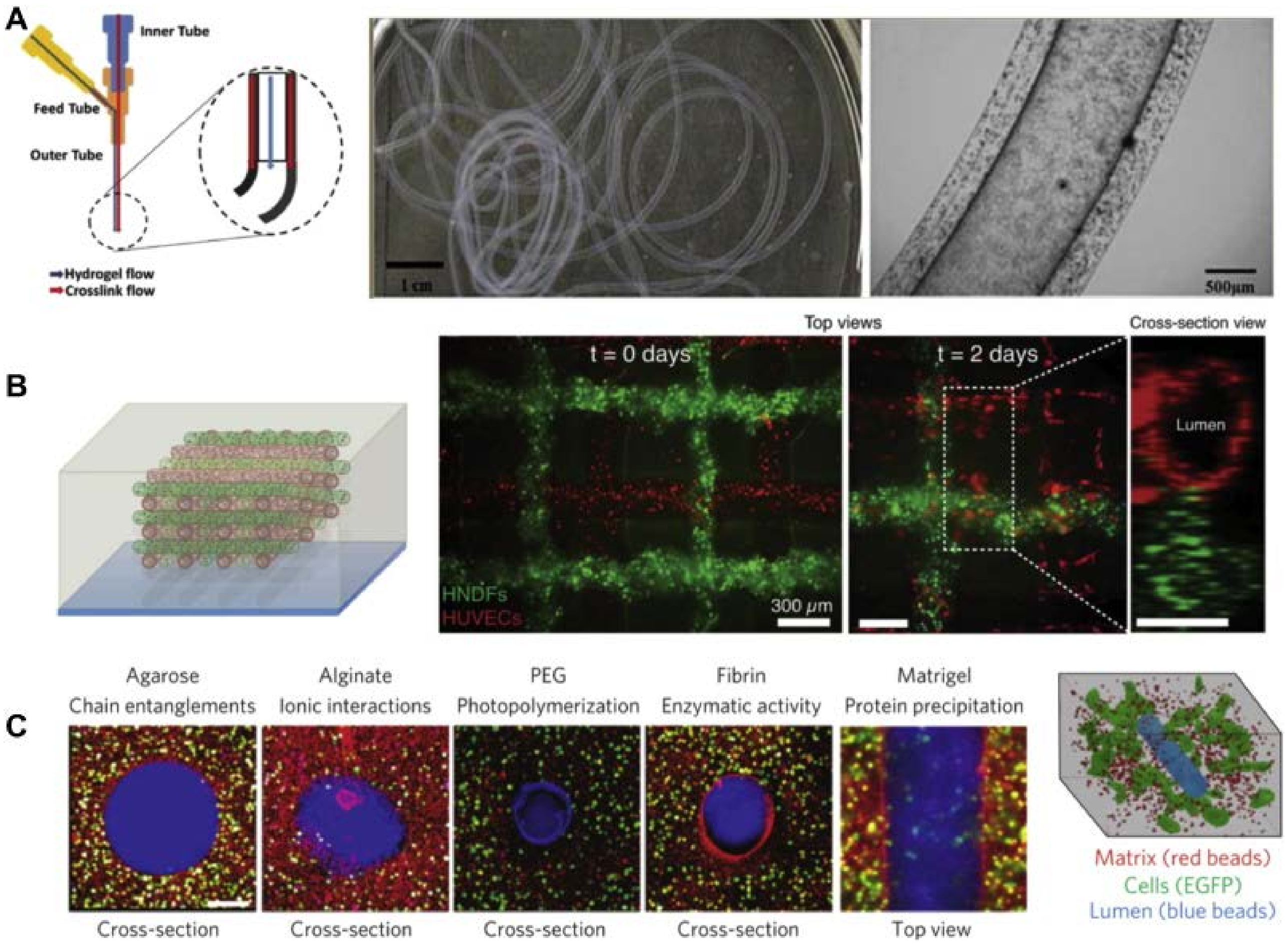


Fig. 3. Bioprinting strategies for vascularization. (A) Fabrication of long (> 1 m) vascular conduits using a coaxial nozzle system yielding internal lumen diameters below 1 mm (Dolati et al., 2014). (B) Pluronic F127 as a sacrificial bioink to form open lumens (red) while concurrently printing encapsulated cells around the vessels (green) (Kolesky

et al., 2014). (C) Carbohydrate glass to cast vascular features into a variety of hydrogels, forming perfusable vessels that support cell growth (Miller et al., 2012).
Figure adapted from Miller et al. (2012), Dolati et al. (2014), and Kolesky et al. (2014).

4.2. Bone and cartilage

The bone engineering space is interesting in that both conventional and bioprinting are poised to influence the field. Made to order metal 3D printed devices (Hsu and Ellington, 2015), 3D printed models for surgical planning (Pietrabissa et al., 2015 and Scawn et al., 2015), and 3D printed tools (Burleson et al., 2015) highlight some of the current and future biomedical applications of conventional 3D printing technologies. Bioprinting techniques have also been applied to bone tissue engineering. Yao et al. (2015) used anatomic data from CT scans of rabbits to print and test polycaprolactone-hydroxyapatite scaffolds which supported physiologically relevant loads. Wang et al. (2015a) printed poly(propylene fumarate) porous scaffolds, characterized the degradation process over a 224 day period, and showed the printed scaffolds were suitable for bone tissue engineering applications. Pati et al. (2015) enhanced the osteogenic potential of 3D-printed PCL/PLGA/ß-TCP scaffolds by using human nasal inferior turbinate tissue-derived mesenchymal stromal cells to deposit bone-like ECM. After a brief culture period, the scaffolds were decellularized and then investigated both in vitro and in vivo where they showed improved osteoinductive and osteoconductive properties.

Cartilaginous tissues have also been an area of interest in tissue engineering (Tatman et al., 2015). Kundu et al. (2013) printed alginate encapsulated chondrocytes with a supportive PCL structure and in vivo experiments suggested cartilage production. Lee et al. (2014) printed a PEG and PCL construct containing chondrocytes and showed that this material mixture could be used to print ear-shaped constructs (Fig. 4A). Similarly, Markstedt et al. (2015) developed a novel nanocellulose–alginate bioink with desirable printing properties. This ink supported the culture of human nasoseptal chondrocytes in printed tissues and could also be printed into complex shapes (Fig. 4B). Collectively, studies like these highlight the promise of bioprinting to produce unique 3D structures suitable for bone and cartilage tissue engineering.

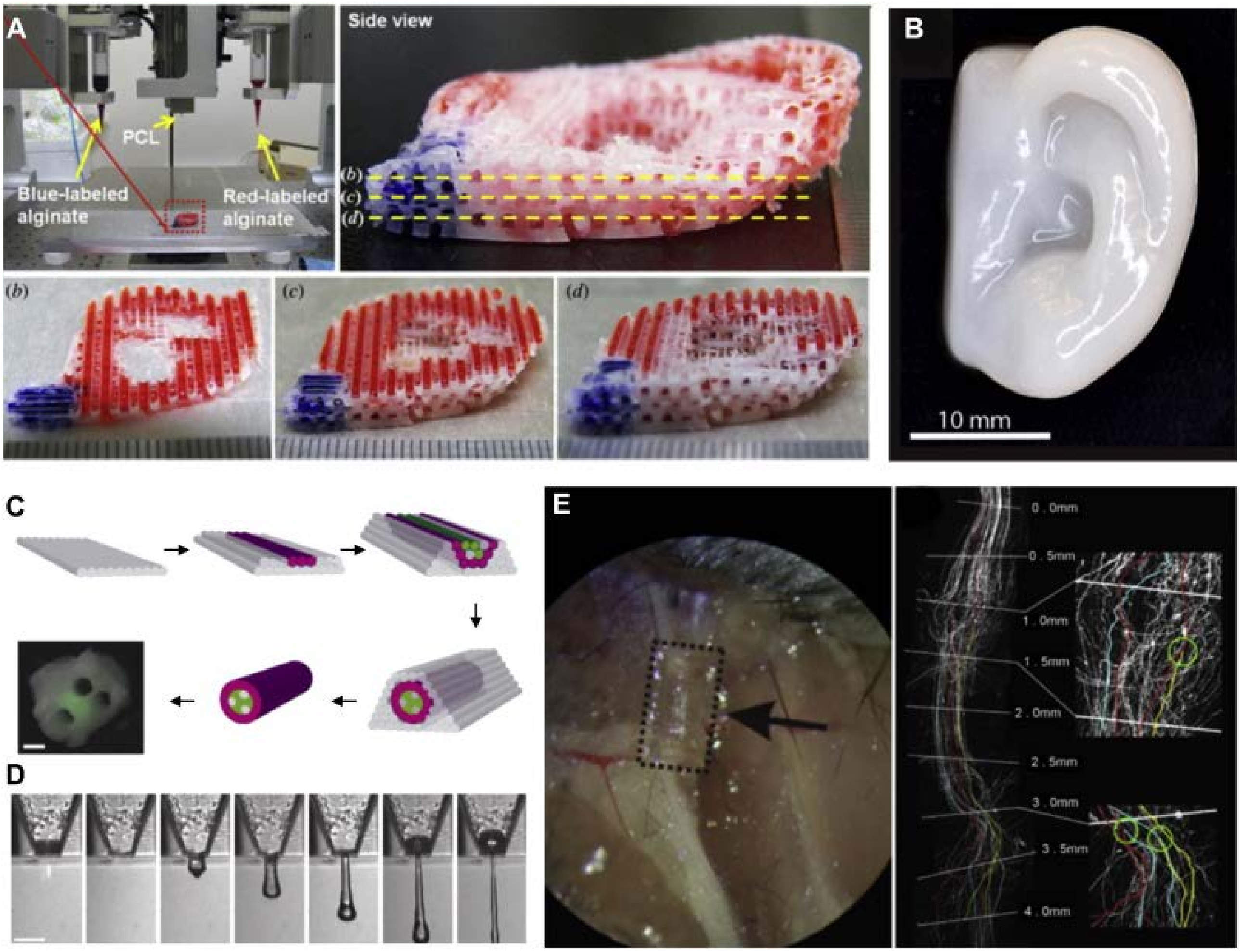


Fig. 4. Examples of bioprinted tissues and organs. (A) Printed ear-shaped PCL and alginate scaffolds with bioinks localized to certain tissue regions (Lee et al., 2014). (B) Cartilaginous ear scaffolds printed using a novel nanocellulose–alginate bioink supported human chondrocytes (Markstedt et al., 2015). (C) Fabrication of a synthetic nerve graft by printing cell-dense tubes of Schwann cells and BSMC (Owens et al., 2013). (D) Demonstration of the feasibility of printing mouse ganglion and glial cells (Lorber et al., 2014). (E) Printed PEG-based guidance conduits for nerve repair studies, showing their biocompatibility and efficacy (Pateman et al., 2015).
Figure adapted from Owens et al. (2013), Lee et al. (2014), Lorber et al. (2014), Markstedt et al. (2015), and Pateman et al. (2015).

4.3. Neuronal tissues

Bioprinting nervous tissue is another application that has been explored by researchers. Large synthetic tissues will need to integrate with the host nervous system, and bioprinting may be a means to generate new nervous tissue or to enhance the innervation of tissue engineered constructs. Owens et al. (2013) printed a synthetic nerve graft using cells alone. Isolated mouse bone marrow stem cells and Schwann cells were cast into 500 μm diameter tubes and then loaded into a bioprinter which extruded discrete tubes to form a dense nerve conduit of Schwann cell tubes surrounded by mouse bone marrow stem cell tubes for use in animal studies (Fig. 4C). These early stage proof-of-principle printed grafts performed similarly to control tissues and remain promising as the methodology is refined and improved. Lorber et al. (2014) also provided important validation on the feasibility of printing cells of the nervous system, showing rat retinal ganglion cells and glia can be used in inkjet printing systems (Fig. 4D). Pateman et al. (2015) used a microsterolithographic technique to print PEG-based nerve guidance conduits for nerve repair studies (Fig. 4E). Printed conduits had a finer resolution than those made through previously reported methods and performed comparably to autograft controls.

4.4. Construction of drug screening systems

Bioprinting is also promising in the design of drug screening systems. Compared to manual methods, bioprinting can deposit cells uniformly on the surface of micro devices. Such uniformity is highly desirable for testing and screening the interactions between cells and drugs (Huh and Kim, 2015 and Nam et al., 2015). Existing examples of bioprinted drug testing platforms include those for the liver (Snyder et al., 2011). Chang et al. (2010) developed an air-pressure based extrusion bioprinter to prototype a drug testing platform for the liver with alginate encapsulated immortalized hepatocytes. In this system, the authors were able to show differential drug metabolism. Snyder et al. (2011) expanded on this system by printing microfluidic channels in a co-culture system of liver and mammary cells to investigate tissue damage from radiation. Bioprinting has also been used to seed cell layers uniformly on each side of the interface of micro devices for the formation of organ-on-a-chip devices (Chang et al., 2010). Organ-on-a-chip systems mimic parts of typical organ functions to investigate the interactions between drugs and their potential effects on tissues (Wang et al., 2015b). Bioprinting may play an important role in organ-on-a-chip technology, given it is a practical solution for the formation of uniform and highly controllable tissue layers at low cost.

5. Present limitations and future prospects

5.1. Current limitations for bioprinting

5.1.1. Limitations of the current bioprinting approach

Although these three common bioprinting techniques have different printing principles and features, there are a few limitations to the typical bioprinting process as it stands today. All three techniques are based on a layer-by-layer printing method, which generally have difficulty printing complex hollow structures. In the simplest case of printing with a single material, each layer must be connected and mechanically supported as it is printed. When voids are introduced in one layer, subsequent layers that deposit material over the void may collapse causing a cascade of offset features and inaccurate geometries. One possible solution to this problem is to incorporate a sacrificial material, which is a method widely employed in the fabrication of suspended structures in MEMS (Taylor et al., 2013 and Bertassoni et al., 2014b). This sacrificial material provides the mechanical support each layer needs during fabrication and is then removed from the completed object in a post-processing step. This approach has been taken by several groups using several fugitive materials, including carbohydrate glass (Miller et al., 2012), Pluronic F-127 (Kolesky et al., 2014) and gelatin microparticles (Hinton et al., 2015). The introduction of extra materials, however, can increase the complexity of the printing process as the bioprinting platform must support rapid material exchanges or multiple nozzles loaded with different inks. Sacrificial materials must be printable under conditions compatible with non-sacrificial biomaterials and cells, and their method of removal and breakdown products must be cytocompatible. These difficulties have likely limited the development and adoption of new sacrificial materials.

The lack of reliable methods to print pre-vascularized tissues is a hurdle that cannot be overlooked. This problem is not unique to bioprinting, but bioprinting is unique in its ability to create large tissues with high metabolic demands relatively quickly. Many of the small-scale tissues researchers currently print can survive through diffusion alone, but full-scale organs and large tissue constructs will require an embedded vasculature as well as mechanically robust conduits to connect to host arteries and veins. Small bioprinted tissues may take only minutes or hours to print, but the question of cell viability both within a pre-polymer bioink and within the polymerized early regions of large multi-day prints must be addressed. Self-assembly of vascular features is too slow a process to rely on when there is the threat of necrosis in partially assembled tissues still sitting on the printer. These sacrificial techniques represent the most promising approach in the current bioprinter's toolbox, but innovation could lead to better printed tissues.

In addition to the difficulty in fabricating hollow vascular features, bioink preparation can take several days to weeks due to cell culturing and biomaterial synthesis (Murphy and Atala, 2014). Once fully prepared, the working time of bioinks may also become an issue. This issue of time may be overcome by incorporating additional features into the bioprinter that support the maintenance of partially printed structures, the development of increasingly parallel bioprinters (e.g. multiple print heads working simultaneously) or other refinements to the printing process (e.g. CLIP; Tumbleston et al., 2015). Faster bioprinters with higher resolution would be poised to solve some of the problems faced by modern technology.

5.1.2. Cell and material limitations

Material selection remains a major concern and limitation for bioprinting. More biomimetic materials like dECM bioinks often lack the mechanical strength to be the sole material in printed tissues, requiring support from stronger but less bioactive inks, like PCL (Ousterout et al., 2013). Tunable bioinks with a wide range of material properties could be a solution to this problem and may be achievable through the creation of new composite mixtures to enhance crosslinking or incorporate other desirable features while maintaining the properties of the base bioink. Poly(ethylene glycol) (PEG) has received attention because of its tunable mechanics (Zustiak and Leach, 2010 and Kim et al., 2014b) and represents a suitable component for composite bioinks. The Khademhosseini group developed PEG:gelatin methacrylate (PEG:GelMA) and carbon nanotube-incorporated photocrosslinkable gelatin methacrylate (CNT:GelMA) composites with tunable mechanical and degradation properties that could have such applications (Shin et al., 2013). Similarly, the West group (Zhang et al., 2015) developed a low molecular weight–high molecular weight PEG composite which could mimic the anisotropy of heart valve leaflet moduli. These kinds of composites further expand the options available to researchers in bioprinting, and may lead to more complex and biomimetic structures.

Incorporating multiple materials also remains a challenge. For most bioprinters, materials to be printed are prepared in bulk before printing begins and switching materials involves changing to secondary pre-loaded reservoirs (e.g. a separate syringe or bioink cartridge). For example, the commercially available 3-D Bioplotter® (EnvisionTEC) is limited to three material cartridges for a single print job. While this approach enables multi-material printing, it makes creating smooth gradients of cells or growth factors impossible or arduous due to the need to prepare many independent solutions. To address this, the Lewis group recently developed an impeller based active mixing system for use in extrusion style printers (Ober et al., 2015). The inclusion of active mixing would reduce the number of solutions that need to be prepared and can enable more precise control over the concentration of deposited components. Although this does introduce some non-trivial complexity to the printing system, the benefits of on the fly mixing are significant. Such a system may also alleviate other concerns for long prints where cell suspensions, pre-polymer solutions, growth factors and other components can be stored in independently controlled reservoirs optimized for their contents.

5.2. Future prospects

In the future, bioprinting may be considered as much a nano-biofabrication technique as a tool for artificial organ generation. Due to its advantages on the micrometer scale, and highly controllable dispensing of live cells, bioprinting may fill a vital role in biofabrication. Bioprinting can be applied wherever the deposition or integration of live cells is desired. Bio-sensors (Xu et al., 2009a) and protein and DNA arrays of stem cells (Tasoglu and Demirci, 2013) have already been fabricated by bioprinting. These diverse applications illustrate the versatility and potential of bioprinting as a technology still in its infancy. Moreover, bioprinting remains a promising solution for addressing the growing international organ shortage. The ability to generate tissues for transplant on-demand with reduced immune response risk holds significant promise in the fabrication of artificial organs. Recent progress in hydrogel science, including the development of dynamic switchable hydrogels (Gillette et al., 2010) and oxygen producing hydrogels (Harrison et al., 2007), provide researchers with more and more methods to control cell microenvironments. In order to realize the potential of bioprinting and rapid prototyping, the printing speed, characteristics of hydrogels, preparation time for cells and hydrogels, vascularization of tissues, innervation of tissues, and the controllability of on-demand scaffold and cell maturation must be improved further. As the technology matures, bioprinting is poised to become a key technique in the fabrication of human-on-a-chip systems as well as on-demand anatomically realistic artificial organs.

6. Conclusions

Bioprinting is an advanced fabrication technique for the dispensing of cell-laden hydrogels, with a bright future accompanying numerous challenges and problems. Bioprinting has shown great potential in tissue engineering applications at its early research stage where many in vitro and even in vivo experiments have already hinted at the feasibility of bioprinted artificial organs. Due to advantages in micro scale, high-throughput, cell deposition, the applications of bioprinting are expanding rapidly. Bioprinting has become a strong fabrication tool to create complex micro- and macro-scale biomedical systems. Even with the progress that has been made, bioprinting remains an emerging and growing technology with incredible potential.

Acknowledgements

This work was supported by a National Institutes of Health R21 Grant (R21AR064395) and a Muscular Dystrophy Association Research Grant (MDA255907). This work was also supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (Application No. RGPIN-2014-04010)

References

1. Abouna, G.M., 2008. Organ shortage crisis: problems and possible solutions. Transplant. Proc. 40 (1), 34–38.
2. Agarwal, S., Greiner, A., Wendorff, J.H., 2013. Functional materials by electrospinning of polymers. Prog. Polym. Sci. 38 (6), 963–991.
3. Arai, K., Iwanaga, S., Toda, H., Genci, C., Nishiyama, Y., Nakamura, M., 2011. Threedimensional inkjet biofabrication based on designed images. Biofabrication 034113.
4. Araujo, J.V., Davidenko, N., Danner, M., Cameron, R.E., Best, S.M., 2014. Novel porous scaffolds of pH responsive chitosan/carrageenan-based polyelectrolyte complexes for tissue engineering. J. Biomed. Mater. Res. A 4415–4426.
5. Bajaj, P., Schweller, R.M., Khademhosseini, A.,West, J.L., Bashir, R., 2014. 3D biofabrication strategies for tissue engineering and regenerative medicine. Annu. Rev. Biomed. Eng. 16, 247–276.
6. Balakrishnan, B., Jayakrishnan, A., Kumar, S.S.P., Nandkumar, A.M., 2012. Anti-bacterial properties of an in situ forming hydrogel based on oxidized alginate and gelatin loaded with gentamycin. Trends Biomater. Artif. Organs 26 (3), 139–145.
7. Balint, R., Cassidy, N.J., Cartmell, S.H., 2014. Conductive polymers: towards a smart biomaterial for tissue engineering. Acta Biomater. 10 (6), 2341–2353.
8. Ballyns, J.J., Gleghorn, J.P., Niebrzydowski, V., Rawlinson, J.J., Potter, H.G.,Maher, S.A., et al., 2008. Image-guided tissue engineering of anatomically shaped implants via MRI and micro-CT using injection molding. Tissue Eng. A 14 (7), 1195–1202.
9. Bauer, M., Kim, K., Qiu, Y., Calpe, B., Khademhosseini, A., Liao, R., et al., 2012. Spot identification and quality control in cell-based microarrays. ACS Comb. Sci. 14 (8), 471–477.
10. Benson, K., Galla, H.J., Kehr, N.S., 2014. Cell adhesion behavior in 3D hydrogel scaffolds functionalized with D- or L-aminoacids. Macromol. Biosci. 14 (6), 793–798.
11. Bertassoni, L.E., Cardoso, J.C., Manoharan, V., Cristino, A.L., Bhise, N.S., Araujo, W.A., et al., 2014a. Direct-write bioprinting of cell-laden methacrylated gelatin hydrogels. Biofabrication 6 (2), 024105.
12. Bertassoni, L.E., Cecconi, M., Manoharan, V., Nikkhah, M., Hjortnaes, J., Cristino, A.L., et al., 2014b. Hydrogel bioprinted microchannel networks for vascularization of tissue engineering constructs. Lab Chip 14 (13), 2202–2211.
13. Billiet, T., Vandenhaute, M., Schelfhout, J., Van Vlierberghe, S., Dubruel, P., 2012. A review of trends and limitations in hydrogel-rapid prototyping for tissue engineering. Biomaterials 6020–6041.
14. Bohandy, J., Kim, B.F., Adrian, F.J., 1986. Metal deposition from a supported metal film using an excimer laser. J. Appl. Phys. 60 (4), 1538–1539.
15. Bracci, R., Maccaroni, E., Cascinu, S., 2013. Transient sunitinib resistance in gastrointestinal stromal tumors. N. Engl. J. Med. 368 (21), 2042–2043.
16. Burleson, S., Baker, J., Hsia, A.T., Xu, Z., 2015. Use of 3D printers to create a patient-specific 3D bolus for external beam therapy. J. Appl. Clin. Med. Phys. 16 (3), 5247.
17. Catros, S., Fricain, J.-C., Guillotin, B., Pippenger, B., Bareille, R., Remy, M., et al., 2011a. Laser-assisted bioprinting for creating on-demand patterns of human osteoprogenitor cells and nano-hydroxyapatite. Biofabrication 3 (2), 025001.
18. Catros, S., Guillotin, B., Bačáková, M., Fricain, J.C., Guillemot, F., 2011b. Effect of laser energy, substrate film thickness and bioink viscosity on viability of endothelial cells printed by laser-assisted bioprinting. Appl. Surf. Sci. 5142–5147.
19. Chang, C.C., Boland, E.D.,Williams, S.K., Hoying, J.B., 2011. Direct-write bioprinting threedimensional biohybrid systems for future regenerative therapies. J. Biomed. Mater. Res. B Appl. Biomater. 160–170.
20. Chang, R., Emami, K.,Wu, H., Sun,W., 2010. Biofabrication of a three-dimensional liver micro-organ as an in vitro drug metabolism model. Biofabrication 2 (4), 045004.
21. Cui, X., Breitenkamp, K., Finn, M.G., Lotz, M., D'Lima, D.D., 2012a. Direct human cartilage repair using three-dimensional bioprinting technology. Tissue Eng. A 1304–1312.
22. Cui, X., Breitenkamp, K., Lotz,M., D'Lima, D., 2012b. Synergistic action of fibroblast growth factor-2 and transforming growth factor-beta1 enhances bioprinted human neocartilage formation. Biotechnol. Bioeng. 109 (9), 2357–2368.
23. Cui, X., Gao, G., Qiu, Y., 2013. Accelerated myotube formation using bioprinting technology for biosensor applications. Biotechnol. Lett. 35 (3), 315–321.
24. Discher, D.E., Janmey, P.,Wang, Y.-L., 2005. Tissue cells feel and respond to the stiffness of their substrate. Science 310 (5751), 1139–1143.
25. Dolati, F., Yu, Y., Zhang, Y., De Jesus, A.M., Sander, E.A., Ozbolat, I.T., 2014. In vitro evaluation of carbon-nanotube-reinforced bioprintable vascular conduits. Nanotechnology 25 (14), 145101.
26. Dou, X.-Q., Yang, X.-M., Li, P., Zhang, Z.-G., Schönherr, H., Zhang, D., et al., 2012. Novel pH responsive hydrogels for controlled cell adhesion and triggered surface detachment. Soft Matter 9539.
27. Duan, B., Hockaday, L.A., Kang, K.H., Butcher, J.T., 2013. 3D bioprinting of heterogeneous aortic valve conduits with alginate/gelatin hydrogels. J. Biomed. Mater. Res. A 101 A (5), 1255–1264.
28. Duarte Campos, D.F., Blaeser, A., Weber, M., Jäkel, J., Neuss, S., Jahnen-Dechent, W., et al., 2013. Three-dimensional printing of stem cell-laden hydrogels submerged in a hydrophobic high-density fluid. Biofabrication 5 (1), 015003.
29. Duocastella, M., Colina, M., Fernández-Pradas, J.M., Serra, P., Morenza, J.L., 2007. Study of the laser-induced forward transfer of liquids for laser bioprinting. Appl. Surf. Sci. 253(19), 7855–7859.

Appl. Eng. Tech., 2016, 1(1): 60, C3.

Here is Your Paper's Title: Author Instructions

FULL First Author[1, a], FULL Second Author[2,b, *] and Others[3,c]

[1]Full address of first author, including country and zip code
[2]Full address of second author, including country and zip code
[3]List all distinct addresses in the same way
[a] email, [b] email, [c] email
* Corresponding Author: Prof. Name, PhD / M.D.. E-mail: abc@abc.com

Abstract. No more than 150 words. This document explains and demonstrates how to prepare your camera-ready manuscript for 1088.email Publications. The best is to read these instructions and follow the outline of this text. The text area for your manuscript must be 17 cm wide and 25 cm high (6.7 and 9.8 inches, resp.). Do not place any text outside this area. Use good quality, white paper of approximately 21 x 29 cm or 8 x 11 inches (please do not change the document setting from A4 to letter). Your manuscript will be reduced by approximately 20% by the publisher. Please keep this in mind when designing your figures and tables etc.

Keywords: keywords 1, keywords 2, keywords 3, keywords 4, keywords 5, keywords 6.

Introduction

All manuscripts must be in English. Please keep a second copy of your manuscript in your office (just in case anything gets lost in the mail). When receiving the manuscript, we assume that the corresponding authors grant us the copyright to use the manuscript for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

Use *italic* for emphasizing a word or phrase. Do not use boldface typing or capital letters except for section headings (cf. remarks on section headings, below). Use a laser printer, not a matrix dot printer.

Organization of the Text

Section Headings. The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph).

Page Numbers. Do *not* print page numbers: Please number each sheet toward the middle near the bottom (outside the typing area) with a soft pencil.

Tables. Tables (refer with: Table 1, Table 2, ...) should be presented as part of the text, but in such a way as to avoid confusion with the text. A descriptive title should be placed above each table. The caption should be self-contained and placed *below or beside* the table. Units in tables should be given in square brackets [meV]. If square brackets are not available, use curly {meV} or standard brackets (meV).

Special Signs. for example , α γ μ Ω () ≥ ± • Γ $\{11\bar{2}0\}$ should always be written in with the fonts Times New Roman or Arial

Figures. Figures (refer with: Fig. 1, Fig. 2, ...) also should be presented as part of the text, leaving enough space so that the caption will not be confused with the text. The caption should be self-contained and placed *below or beside* the figure. Generally, only original drawings or photographic reproductions are acceptable. Only very good photocopies are acceptable. Utmost care must be taken to *insert the figures in correct alignment with the text*. Half-tone pictures should be in the form of glossy prints. If possible, please include your figures as graphic images in the electronic version. For best quality the pictures should have a resolution of 300 dpi(dots per inch).

Color figures are welcome for the online version of the journal. Generally, these figures will be reduced to black and white for the print version. The author should indicate on the checklist if he wishes to have them printed in full color and make the necessary payments in advance.

Equations. Equations (refer with: Eq. 1, Eq. 2, ...) should be indented 5 mm (0.2"). There should be one line of space above the equation and one line of space below it before the text continues. The equations have to be numbered sequentially, and the number put in parentheses at the right-hand edge of the text. Equations should be punctuated as if they were an ordinary part of the text. Punctuation appears after the equation but before the equation number, e.g.

$$c^2 = a^2 + b^2. \quad (1)$$

Literature References

References are cited in the text just by square brackets [1]. (If square brackets are not available, slashes may be used instead, e.g. /2/.) Two or more references at a time may be put in one set of brackets [3,4]. The references are to be numbered in the order in which they are cited in the text and are to be listed at the end of the contribution under a heading *References*, see our example below.

Summary

On your CD, please indicate the format and word processor used. Please also provide your phone number, fax number and e-mail address for rapid communication with the publisher. Please *always* send your CD along with a hard copy that must match the CD's content *exactly*. If you follow the foregoing, your paper will conform to the requirements of the publisher and facilitate a problem-free publication process.

Acknowledgement

For example: This paper is supported by the National Foundation (123456, 34567 and 56789), the Research Program of 1088.Email (2016SRWLV007) and the Foundation of 1088.Email (1088EP2016-1234). The authors would like to acknowledge the constructive comments given by the anonymous reviewers.

References

[1] Dj.M. Maric, P.F. Meier and S.K. Estreicher: Mater. Sci. Forum Vol. 83-87 (1992), p. 119

[2] M.A. Green: *High Efficiency Silicon Solar Cells* (Trans Tech Publications, Switzerland 1987).

[3] Y. Mishing, in: *Diffusion Processes in Advanced Technological Materials*, edtied by D. Gupta Noyes Publications/William Andrew Publising, Norwich, NY (2004), in press.

[4] G. Henkelman, G.Johannesson and H. Jónsson, in: Theoretical Methods in Condencsed Phase Chemistry, edited by S.D. Schwartz, volume 5 of Progress in Theoretical Chemistry and Physics, chapter, 10, Kluwer Academic Publishers (2000).

[5] R.J. Ong, J.T. Dawley and P.G. Clem: submitted to Journal of Materials Research (2003)

[6] P.G. Clem, M. Rodriguez, J.A. Voigt and C.S. Ashley, U.S. Patent 6,231,666. (2001)

[7] Information on http://www.weld.labs.gov.cn

...

Application of Engineering and Technology
ISSN Pending
Volume 1 number 1
May 20, 2016
Published on *http://www.1088.email*