Abstract—Industry 4.0 is a vision of manufacturing in which smart, interconnected production systems optimize the complete value-added chain to reduce cost and time-to-market. At the core of Industry 4.0 is the smart factory of the future, whose successful deployment requires solving challenges from many domains. Model-based systems engineering (MBSE) is a key enabler for such complex systems of systems as can be seen by the increased number of related publications in key conferences and journals. This paper aims to characterize the state of the art of MBSE for the smart factory through a systematic mapping study on this topic. Adopting a detailed search strategy, 1466 papers were initially identified. Of these, 222 papers were selected and categorized using a particular classification scheme. Hence, we present the concerns addressed by the modeling community for Industry 4.0, how these are investigated, where these are published, and by whom. The resulting research landscape can help to understand, guide, and compare research in this field. In particular, this paper identifies the Industry 4.0 challenges addressed by the modeling community, but also the challenges that seem to be less investigated.

I. INTRODUCTION

Industry 4.0 is the current trend of integrating automation systems with processes and stakeholders of the complete value-added chain as well as part of the high-tech strategy of the German Federal Ministry for Education and Research [1]. This “fourth industrial revolution” raises the new challenges for future smart factories driven by four disruptions: (1) data volumes, computational power, and connectivity; (2) the emergence of analytics and business-intelligence capabilities; (3) new forms of human-machine interaction; (4) and improvements in transferring digital instructions to the physical world, such as advanced robotics and 3D printing. The interplay of these four disruptions led to recognizing four particular Industry 4.0 design principles [2]:

- Interoperability: connect production systems, devices, sensors, and people.
- Information transparency: query data and connect digital planning with the runtime data collected from sensors.
- Technical assistance: provide the right abstraction to understand the complexity of Industry 4.0 systems and processes.
- Decentralized decision making: enable autonomous systems.

All of these aim to enable more efficient production down to the individualized mass production of “lot-size 1” [3] or, to put it differently:

This work was supported by the French project Clarity: www.clarity-se.org

“The big money (of Industry 4.0) is on two things; zero unscheduled downtime and resource efficiency”

– Bill Ruh, vice president of GE Software [4]

Model-based systems engineering (MBSE) is a key enabler for building complex systems of systems as can be seen by the increased number of related publications in key conferences and journals [5], [6], [7]. For successfully engineering Industry 4.0 systems of systems, the MBSE community plays a crucial role by enabling the aforementioned design principles.

As a research area matures, there often is a sharp increase in the number of reports and results. Hence, it becomes important to summarize and provide an overview of those results. There are different methods for structuring a scientific landscape, such as systematic literature reviews [8] or systematic mapping studies [9]. Systematic literature reviews are a “form of secondary study that use a well-defined methodology to identify, analyze, and interpret all available evidence related to a specific research question in a way that is unbiased and (to a degree) repeatable” [8]. Systematic literature reviews are common to software engineering, model-based engineering, software product lines, or domain-specific language engineering, etc. Mapping studies are less common. A systematic mapping study structures a body of research through its reports by categorizing these, which often yields a visual summary, the map, of its results. Such a map supports understanding what has been addressed by the community for a particular domain.

Through this mapping study, we aim to characterize the state of the art of MBSE for Industry 4.0 in a broad sense, i.e., we include 3D modeling, knowledge representation, business process modeling, and other modeling techniques into our study. Adopting a detailed search strategy using multiple digital libraries, 1466 papers were initially identified. Of these, 222 papers were selected and categorized using a particular classification scheme along the contribution types, research types, Industry 4.0 concerns, and modeling contributions. We present the concerns addressed by MBSE for Industry 4.0, how these concerns are investigated, when and where the results are published, and by whom. The resulting research landscape can help to understand, guide, and compare research in this field. In particular, this paper identifies the Industry 4.0 challenges addressed by the modeling community as well as the challenges that seem to be less investigated. As a main contribution, we obtain a classification scheme and structure the research on modeling for Industry 4.0.
In the remainder, first, Sec. II highlights selected mappings in software engineering and presents related literature reviews and surveys. Afterwards, Sec. III describes the research method used for the mapping study presented in this paper, before Sec. IV discusses main findings regarding this study. Then, Sec. V presents the threats to validity. Finally, Sec. VI summarizes the main contribution of this paper and discusses research directions that could be investigated by the modeling community.

II. RELATED WORK

Several mapping studies have been published in software engineering [9], such as on the classification of techniques for test-set generation and selection [10], on software development effort and cost estimation [11], on the use of experimental studies [12], on object-oriented design [13] and the use of patterns [14], on the usage of UML diagrams [15], on the empirical evaluation of software requirements specifications techniques [16], on software product lines [17], [18], and on domain-specific languages [19]. These mapping studies vary in the form of analysis and in the number of included publications (between 35 and 400). We found only a single mapping study on model-driven engineering [20]. In this study, Mehmood and Jawawi survey existing researches on aspect-oriented modeling (AOM) and code generation. To the best of our knowledge, there is no mapping study on modeling and systems engineering or modeling and Industry 4.0, yet. However, several literature reviews and surveys focus on the Industry 4.0 domain.

Originally initiated in Germany in 2011, Industry 4.0 has attracted much attention in recent literature. In their perspective on Industry 4.0, Vogel-Heuser and Hess identify a set of challenges for the domain [21]. In particular, they identify challenges for software engineering that are well-known to the modeling community:

- Transition to modular and maintainable interfaces as a fundamental basis for adaptable and evolvable systems.
- Tracking of changes in hundreds of heterogeneous and distributed machines or plants on different operation sites operated over decades.
- Management while ensuring consistency of software variants and versions, including self-adaptation and reconfiguration at runtime.
- Adaptation of big data algorithms and technologies.

Following this paper, Mosterman and Zander discuss the needs and challenges for designing and operating cyber-physical systems (CPS) along with corresponding technologies to address the challenges and their potential impact [22]. In the same trend, Turowski et al. identify the current challenges on Industry 4.0 faced by companies through a survey [23]. The survey aims at understanding the stakeholder’s expectations, requirements and the potential challenges Industry 4.0 poses in real case studies. Complementary to these works, Trappey et al. [24] provide a consolidated review of the latest CPS literature. In this survey, they provide a complete review of international standards and an analysis of patent portfolios related to the CPS architecture model. Hermann et al. identify design principles of Industry 4.0 based on a quantitative text analysis and a qualitative literature review [25]. This paper illustrates how the identified design principles support practitioners in identifying Industry 4.0 scenarios. Recently, Lu proposes an extensive literature review on Industry 4.0 [26]. In this review, 88 papers related to Industry 4.0 are grouped into five research categories. In addition, this paper outlines the critical issue of the interoperability of Industry 4.0, and proposes a conceptual framework of interoperability regarding Industry 4.0. It also discusses challenges and trends in Industry 4.0, but does not focus on its modeling concerns.

III. RESEARCH METHOD

A systematic mapping study identifies and classifies primary studies of the field under investigation. As such, it aims to provide a systematic overview on the topics of research contributed to this area and the forms of contribution. We conducted this study following established guidelines [8], [9] and included useful practices and suggestions from similar studies [19], [27], [28], [29]. Ultimately, we employed a five-phase process for conducting this study: (1) Define research questions; (2) Search for primary studies; (3) Identify inclusion and exclusion criteria and screen primary studies based on these criteria; (4) classify primary studies; (5) Extract and aggregate data. In the first phase, we characterize the scope of this study. In the second phase, we build up the initial corpus of potentially relevant publications. In the third and fourth phases, we reduced this corpus to include only relevant publications and classified according to research qualities derived from the research questions. In the fifth phase, we extracted data to enable answering our research questions. The remainder of this section describes the activities and decisions of these phases.

A. Research Questions

We aim to identify relevant publications about MBSE in Industry 4.0, which concerns MBSE research addresses in this context, how research addressing these concerns is conducted, and which tools and languages are used to contribute to these concerns. Moreover, we investigate who is contributing to MBSE in Industry 4.0, where the contributions are published, and when they occurred. This manifests in the following research questions:

RQ1 What are the expected benefits from applying MBSE to contribute to Industry 4.0? This question aims to identify the benefits expected by applying MBSE at Industry 4.0.

RQ2 Which Industry 4.0 concerns are addressed through applying MBSE? With this question, we investigate which concerns of Industry 4.0 are addressed by the MBSE researchers and to which extent.

RQ3 Which MBSE tools and languages are used in Industry 4.0 and which concerns do they address? This question relates the findings of RQ2 to the solutions contributed to the research field.

1Available from companion website: http://gemoc.org/modeling4Industry4.0/
RQ4 What are the most frequently applied research methods in the context of applying MBSE to Industry 4.0? With this question, we aim to understand how research in MBSE for Industry 4.0 is performed and how this relates to the concerns of RQ2 and the tools of RQ3.

RQ5 Who researches MBSE in Industry 4.0? Industry 4.0 was coined as “Industrie 4.0” in Germany. This question investigates who has adopted this notion and contributes to MBSE research in this field.

RQ6 Where have the contributions been published? Similar to RQ5, we like to uncover which venues are relevant to MBSE for Industry 4.0.

RQ7 When did the contributions on MBSE to Industry 4.0 occur? With this question, we aim to comprehend when MBSE actually started contributing to smart manufacturing and whether Industry 4.0 might be an already declining.

To answer these questions, we conducted the initial literature search presented in the next section.

B. Search Strategy and Data Sources

Finding possibly relevant publications to answer the research questions requires creating an appropriate search clause and selecting relevant libraries to apply this clause to. As Industry 4.0 aims at interconnecting all participants of the value-added chain to optimize various aspects of manufacturing, at its core, it focuses on manufacturing, production processes, and ultimately the “factory of the future” [30], [31] or the “smart factory” [32], [33]. Consequently, we included these terms into our search clause. Similarly, MBSE uses various terminology (such as “model-based”, “model-driven engineering”, “model-based software development”) with varying meaning depending on the authors. Instead of selecting a possibly biased set of near-synonyms, we decided to include the objects of MBSE into the search clause instead. Hence, we consider papers mentioning metamodels, DSLs, modeling languages, or UML as relevant contributions to MBSE in our context. This ultimately lead to the logical search clause depicted in Fig. 1.

Essentially, this is a conjunction of two disjunctions: The first part of the conjunction specifies terms related to Industry 4.0. The second part specified terms representing the objects of MBSE research. As we conducted a full-text search with this clause, we omitted including synonyms for “DSL” or “modeling language”. Papers contributing to MBSE should at least use these terms in either related work or the referenced literature. Although we cannot exclude missing a small number of relevant publications that do not provide such discussions, searching this way yields better results than just searching titles and abstracts. We also did not enforce any inferior year-limit and included papers published until April 2017. We applied the search clause to ACM Digital Library, IEEE Explore, SpringerLink, and Google Scholar, which produced the results depicted in Table I. Where such complex logical conditions were not supported, we searched for parts of the query and joined the results manually.

![Fig. 1. Logical search clause defined to identify relevant literature.](http://gemoc.org/modeling4Industry4.0/)

<table>
<thead>
<tr>
<th>Digital Library</th>
<th>URL</th>
<th>Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM Digital Library</td>
<td><a href="https://dl.acm.org">https://dl.acm.org</a></td>
<td>104</td>
</tr>
<tr>
<td>Google Scholar</td>
<td><a href="https://scholar.google.com">https://scholar.google.com</a></td>
<td>971</td>
</tr>
<tr>
<td>IEEE Explore</td>
<td><a href="https://ieeexplore.ieee.org">https://ieeexplore.ieee.org</a></td>
<td>151</td>
</tr>
<tr>
<td>SpringerLink</td>
<td><a href="https://link.springer.com">https://link.springer.com</a></td>
<td>240</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1466</td>
</tr>
</tbody>
</table>

C. Screening Papers for Inclusion/Exclusion

Inclusion of a study into the classification phase of a systematic mapping study usually is decided on its title, abstract, and keywords. To reduce the corpus and enable reproduction of the study, we established the following criteria:

**Inclusion criteria:**
1) Peer-reviewed studies published in journals, conferences, and workshops.
2) Studies are accessible electronically.
3) From title, abstract, and keywords we can deduce that the paper focuses on contributing MBSE to Industry 4.0.

**Exclusion criteria:**
1) Studies not available in English.
2) Studies not subjected to systematic peer-review, such as books, slides, websites.
3) Teasers and short papers of less than two pages, such as calls for papers, editorials, or curriculums.
4) Studies where Industry 4.0 is mentioned as future application, related work, or broad context only, e.g., papers on Internet of Things (IoT) or Cyber-Physical Systems (CPS) mentioning Industry 4.0 as a possible use case only.

We applied the criteria to titles, keywords, and abstracts. Where this did not suffice to determine inclusion, we temporarily included the publications for the classification phase. There, the final inclusion or exclusion could be decided based on the publication’s full text. Thus, this phase only eliminates publications clearly not within our study’s scope and publications failing on formal requirements (such as being available in English). In detail, we eliminated 402 publications outside this study’s scope, 166 non-English publications, 78 books 43 bachelor, master, and Ph.D. theses, 38 technical papers.

reports, 30 teasers, 28 full proceedings (Google Scholar produces complete conference proceedings as results), 10 patents, 10 project deliverables, and 3 websites. Whenever a publication was given in another language than English, we excluded it for this reason alone. Overall, this left 658 papers for classification.

D. Classifying Studies

In the classification phase, the remaining papers were investigated to assign qualities of the dimensions derived from the research questions to them. To achieve this, we followed [9] in considering at least introduction and conclusion. However, for many papers this was insufficient and we considered the complete paper for proper classification. This also is the last phase in which publications can be eliminated. Hence, after further elimination of 436 unrelated papers, a total of 222 publications remained. These were classified along the facets described in the following.

**Contribution Type Facet:** The first facet is inspired by [9] to classify publications according to the type of research they contribute (RQ4). We adapted this to our study by employing the five contribution types presented in [Tbl. II] The contribution types are disjoint and each paper was classified to provide exactly one contribution type.

<table>
<thead>
<tr>
<th>Contribution Type Facet</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyses</td>
<td>Papers contributing investigations without constructive contributions.</td>
</tr>
<tr>
<td>Concepts</td>
<td>Papers suggesting ways of thinking things, such as new metamodels or taxonomies (this was titled “models” in [9], which is misleading in the context of your study)</td>
</tr>
<tr>
<td>Methods</td>
<td>Papers suggesting ways of doing things, such as novel approaches to use UML for integrating materials flows and the participating cyber-physical production systems [34].</td>
</tr>
<tr>
<td>Metrics</td>
<td>Papers suggesting ways of measuring things.</td>
</tr>
<tr>
<td>Tools</td>
<td>Papers presenting software tools.</td>
</tr>
</tbody>
</table>

**Research Type Facet:** Also inspired by [9], we classified the publications according to the research type they contribute. This enables addressing RQ4 on the most frequently applied research methods contributed to MBSE for Industry 4.0. Again, we adjusted the taxonomy [9] to better match our study. In particular, we eliminated the category of philosophical papers as these did not occur. The resulting, disjoint, research types are depicted in [Tbl. III] Each paper was classified to belong to exactly one research type.

**Industry 4.0 Concern Facet:** We also classified the publications along the concerns addressed by the various publications. This addresses RQ2 and aims to uncover which concerns are investigated how often. During classification, keywording (cf. [9]) the abstracts, introductions, conclusions, and in complex cases the complete paper, produced the following Industry 4.0 concerns. These concerns are not disjoint and included papers can contribute to multiple concerns.

**Modeling Technique Facet:** To answer RQ3 on the MBSE tools and languages used in Industry 4.0, we also classified the publications along this dimension. Overall, we identified 81 modeling techniques and many papers addressed more than one modeling technique. To prevent dissipating the results we aggregated modeling techniques into groups (such as 3D modeling, architecture description languages, or business process modeling techniques) and isolated modeling techniques more specific to Industry 4.0 (such as AutomationML). This led to the 15 groups presented in [Tbl. V] Moreover, we also investigated whether the classified publications report on real-world industrial applications. Out of the 222 publications included in classification, only 10 reported such applications in automotive [52], avionics manufacturing [47], packaging [53], production of white goods [54], oil production [55], and production of windows and doors [56]. The next section presents our main findings along the classification dimensions.

### IV. Main Findings

This section presents our findings on the expectations researchers in MBSE for Industry 4.0 have regarding the impact of their papers as well as on the contribution types, research
in terms of contribution types (Tbl. II) and research methods (Tbl. III). Investigating this, we found that most publications on MBSE for Industry 4.0 contribute methods to challenges in digital representation (78 publications), integration (58), and processes (32). Moreover, many publications contribute concepts to digital representation (25) and integration (15). Out of the 347 concerns addressed by the included publications, these five combinations of contribution types and concerns are addressed by 60% of the papers in this corpus. Overall, the majority of papers contributes methods (61%) or concepts (17%), whereas analyses, metrics, or tools are contributed less often. With contributions claiming to reduce costs and time (cf. Sec. IV-A), the lack of papers contributing metrics to track these claims is surprising. The results concerning contribution types (as inquired by RQ4) are depicted on the left part of Fig. 2 and these findings are reflected by the types contributions on its right part. Most contributions are solution proposals (i.e., application of existing techniques to solve particular problems) that focus either on digital representation or integration challenges. It is also surprising, that only few papers address modeling for the smart product, which is supposed to control its production processes in many visions of Industry 4.0. Overall, solution proposals make 79% of the research, whereas only few experience reports, validation research, evaluation reports, and vision papers are contributed. That most solution papers also are method papers might reflect the very constructive research typical to MBSE. However, the large number of method papers over paper contributing new concepts, validating new techniques, or proposing visions implies that research mainly approaches Industry 4.0 with established methods and techniques. This is supported by our findings on the modeling techniques contributed to Industry 4.0 presented in the next section.

| 3D Modeling | Techniques for representing geometric properties, e.g., for factory planning or augmented reality systems, including AutoCAD [41] and CATIA 3D [59]. |
| ADL | Techniques employing architecture description languages [42]. |
| AML | Techniques employing the AutomationML [43] plant engineering data exchange format. |
| BPM | Techniques for business process modeling in the context of Industry 4.0, for instance [44], [55]. |
| CMSD | Approaches based on Core Manufacturing Simulation Data (CMSD), such as [56]. |
| DSL | Domain-specific languages. |
| ER | Entity-relationship modeling (cf. [47]). |
| Formal Modeling | Automata-based and mathematical modeling approaches, including Petri Nets [48] or Priced Timed Automata [50]. |
| GPL | Techniques employing general programming languages (GPLs), for instance to model the services provided by a robotic manufacturing system [59]. |
| KR | Knowledge representation techniques, using, for instance, OWL [51]. |
| Meta | Various metamodeling techniques. |
| Simulink | Approaches using MATLAB/Simulink. |
| SysML | Techniques employing SysML. |
| UML | UML and UML profiles. |
| XML | XML-based modeling techniques. |

TABLE V
MODELING TECHNIQUE FACETS.

A. Expectations on the Impact of MBSE

With RQ1 we address the reasons why researchers contribute MBSE to the challenges of Industry 4.0. To this end, we investigated the expected impact of modeling approaches to Industry 4.0 challenges and extracted whenever these were made explicit. Out of the 222 publications included after classification, only 25 (11%) explicitly stated the authors’ expectations on the impact of their contribution. The expectations include reducing the cost for production system integration [57], saving energy on production system reconfiguration [49], and remaining internationally competitive in high-wage countries [54]. We classified the expectations into expectations on reducing time (development time, time-to-market), reducing costs (of development, integration, (re-)configuration), and improving either sustainability or international competitiveness.

Overall, the included publications 27 expectations. Out of these, most publications expected MBSE to either reduce time (12x mentioned) or costs (10x), whereas only few publications expect MBSE to improve sustainability (4x), increase international competitiveness (2x), or enhance the quality of products (1x). However, as the number of papers making the expectations of contributing MBSE to Industry 4.0 is rather small, this cannot be generalized.

B. Industry 4.0 Concerns Addressed with MBSE

With RQ2, we aim at uncovering which concerns of Industry 4.0 are addressed using MBSE and how they are addressed
C. Modeling Techniques Applied to Industry 4.0

Regarding RQ3, out of the 222 publications included into classification, 178 (80%) explicitly specified the (meta)modeling technique the authors applied to contribute to MBSE for Industry 4.0. Examining these publications produced 96 different techniques. Most notably among these are:

- Variants of UML, such as DiSpa [58], Mechatronic UML [59], UMM [60], and UML4IoT [61];
- The systems modeling language (SysML) [62] and its variants, such as SysML4Mechatronics [63] and SysML4Modelica [64];
- Knowledge representation techniques (mostly employing OWL [65], [66]);
- Various DSLs, such as the EXPRESS DSL for product data modeling [62], the virtual factory data model [68], the Industry 4.0 process modeling language [69], the graphical modeling language for value networks [70], or the graphical modeling framework for production processes [71].

Overall, out of the 222 classified papers, 47 employ UML and variants, 36 use DSLs specific to Industry 4.0 challenges, 26 employ knowledge representation techniques, and 19 papers use AutomationML [43]. Where a column of Fig. 3 sums up to a higher number (e.g., column UML), the column’s modeling technique was applied to multiple concerns. We also found eight papers applying explicit metamodeling techniques to Industry 4.0 challenges. These either used specific language workbenches, such as Xtext [72], [73], [74] or MetaEdit+ [75], [76] or more generic metamodeling frameworks, such as MOF [52], EMF/ECore [77]. Overall, 40% of the contributions address Industry 4.0 challenges with new DSLs, language profiles of UML or SysML, or metamodeling techniques. This hints at modeling challenges that cannot be properly addressed by established modeling techniques.

To answer RQ3, we also aim to understand which modeling techniques are applied to address the different Industry 4.0 concerns. The results, depicted in Fig. 4, show that UML is used mostly to solve challenges regarding digital representation (28 papers) and regarding integration (23 papers). This is consistent with identifying these as the most important challenges addressed by included publications. Hence, these are the two concerns most often addressed with knowledge representation techniques, DSLs, and architecture description languages as well. For failure handling and process modeling, knowledge representation techniques are more prominent than UML. Overall, the concerns digital representation and integration – addressed by either AutomationML, various DSLs, knowledge representation techniques, SysML, or UML – make up for 36% of the included publications. While the usage of UML is almost equally distributed between both concerns, DSLs and knowledge representation techniques lean towards digital representation challenges. This might imply that UML and its variants are more suitable to integration than the former.

The results also show that neither validation & verification, nor the human factors crucial to the success of Industry 4.0 or product modeling are investigated as much as integration and digital representation. Whereas the former might require solving digital representation and integration (to some degree) first, the lack of research on the latter two is elusive. Unless the smart factory of the future is fully automated, human interaction and control are necessary and should be considered appropriately.

D. Countries and Institutions Contributing to the Field

Investigating RQ5, we found that almost half of the papers were contributed by teams including German authors, followed by teams including authors from the USA, Italy, Greece, and France as depicted in Fig. 5. Overall 38 countries contributed to MBSE for Industry 4.0 in 278 contributions (papers with authors from multiple countries produce multiple country contributions). Out of these, the top 10 contributing countries produce 210 (i.e., 79%) of the 278 country contributions. Among these, 176 (66%) contributions are from European countries, which suggests that MBSE for Industry 4.0 still largely is a European research program. Whether this is due to the focus on manufacturing or on MBSE cannot be derived from the data.

Aside from the contributing authors’ countries, we also identified the institutions most actively engaging into MBSE for Industry 4.0 research. Due to Industry 4.0 originating in Germany and 51% of the included publications having German co-authors, it is not surprising that of the 10 most active institutions in this field, 7 are from Germany as depicted in Fig. 6. Out of the 197 overall contributing institutions, 124 (63%) are universities and 31 (16%) are research institutes. This indicates that despite the business-driven term, its research largely is academic. Further research is contributed by 42 companies and other organizations.

E. Popular Venues for Publications on MBSE for Industry 4.0

Investigating RQ6, we determined that most papers are published at conferences (151, 68%), followed by journals (63,
Fig. 4. Individual Industry 4.0 concerns and modeling techniques addressing them.

28%), and workshops (8, 4%). We also identified the most popular journals, conferences and workshops of this particular field of research, to answer RQ6 on the most popular venues for MBSE research in the context of Industry 4.0. Fig. 7 presents the 10 most popular journals, where 41% of the related journal papers are published. Where journals produced the same number of publications, we selected in alphabetical order. The low numbers of publications in these most popular journals, however, do not support conclusions over their importance. As the Industry 4.0 will mature, future studies may draw such conclusions based on a larger corpus of relevant publications.

The 10 most popular conferences to MBSE for Industry 4.0, depicted in Fig. 7, publish, with 40%, a similar share of the related conference publications. However, the bigger number of conference publications supports the conclusion that the conference on Emerging Technologies And Factory Automation (ETFA) is the most important conference for publications on MBSE for Industry 4.0. The other conferences published between 1 and 2 papers on the topic and distinguishing their importance, hence, is not feasible. Overall, the publication activities regarding conferences and journals focus on venues on automation engineering. The 10 most popular journals and conferences publish in total 39% of the included papers, which hints at a healthy distribution of publications over multiple venues. This is reflected by the eight workshop papers included in the classification, each of which was published at a different workshop. Hence, no results on popularity can be drawn.

F. Publication Activities over Time

On RQ7, we found that MBSE in Industry 4.0 was already addressed as early as 1991 [78], although neither the term “Industry 4.0” was coined, nor MBSE was a common research topic. Over half (123) of related publications were published starting in 2015 and 182 (81%) of the publications are from 2011 (the year the term “Industry 4.0” was coined) or later (cf. Fig. 9).

V. THREATS TO VALIDITY

The presented study is subject to threats to construct validity (research design), internal validity (data extraction), and conclusion validity (reliability). Threats to external validity (generalizability) are irrelevant as the results of this study cannot be generalized to other problems domains (Industry 4.0) or other solution domains (MBSE).

Regarding threats to research design, the findings presented in Sec. III and Sec. IV are valid only for our sample of papers. Hence, it is crucial to ensure inclusion of as many relevant papers as possible. To achieve this, we included the Google Scholar digital library. Although we are aware that a great
number of subsequent exclusions for formal reasons (e.g., non-peer-reviewed materials) are due to including Google Scholar, its inclusion was useful to capture venues not published at the other libraries. Overall, considering Google Scholar led to including 108 papers that would have otherwise been omitted.

Furthermore, we did not restrict our search to papers mentioning “Industry 4.0” explicitly, but also included the related terms of the search clause’s first part. Similarly, the search clause’s second part included terms closely related to model-based software engineering, without narrowing it to the exact terms. Instead, we used terms one can expect from relevant contributions to be included in the full text. This enabled to capture related publications without focusing on the very specific, partly ambiguous, modeling terminology. Another threat to research design validity arises from the definition of the inclusion/exclusion criteria. During the screening, only title, abstract, and keywords were considered. To prevent excluding relevant publications based on the lack of investigation, we included papers we were uncertain of temporarily. In the subsequent classification phase, the complete papers were read and inclusion/exclusion were decided ultimately.

Threats to internal validity arise from the data extraction performed for screening and classification. The lack of established terminology, absent information, or unfavorably presented papers may lead to wrong decisions regarding inclusion and exclusion. To mitigate this, inclusion and classification of problematic papers was discussed among the authors. These discussions served as quality assessment also and, e.g., lead to excluding very short papers (cf. excluded teasers). We did, however, not discard papers based on their comprehensibility or venue alone. We also assigned each paper to the most suitable research type facet to yield a clear partitioning of the data set according to the categories of Tbl. III. Where debatable, we discussed this also among all authors. To prevent the threat of classification fatigue, the classification was performed in blocks of at most one hour broken up by at least 15 minute breaks. The resulting classification scheme comprises four facets that correspond to our research questions.

Threats to conclusion validity arise drawing wrong conclusions and from the study’s replicability. Regarding the former, we have discussed various issues that could lead to wrong conclusions in the context of threats to internal and external validity. For replicability, we detailed the complete research method in Sec. III, which enables replicating every phase of this mapping study.

VI. CONCLUSION

Industry 4.0 is the application of systems engineering concepts, methods, and tools to the development of adaptive systems of systems that enable smart manufacturing with CPS in the IoT. Model-based systems engineering has shown to facilitate development of such systems, but its application to Industry 4.0 was not systematically investigated, yet. To comprehend the contribution of MBSE to Industry 4.0, we conducted a systematic mapping study, which revealed that digital representation of automation systems, i.e., their interfaces and data models, as well as their integration and (re-)configuration are the prime Industry 4.0 concerns addressed by MBSE. Also, the majority of papers contribute methods and concepts to solve particular challenges of Industry 4.0. Despite this is a relatively new term, we found only 9 vision papers addressing its various concerns. If this is considered an indicator for the domain’s maturity, the lack of experience papers and experience reports is notable. We do not consider this point regarding the mapping study validity. Indeed, systematic mapping studies in software engineering have been recommended mostly for research areas where there is a low number relevant, high-quality primary studies. There is a similar lack of research contributing metrics and analyses on the applicability of specific concepts, methods, and tools to the Industry 4.0 challenges. It is also surprising that there is relatively little research in validation & verification despite the huge costs that production system failures might entail. However, with Industry 4.0 being business-driven and aiming at reducing cost and time, such contributions might arise once the field has matured.

![Fig. 5. The top 10 countries with authors contributing to MBSE in Industry 4.0 are mostly from Europe and contribute 79% of the publications.](image1)

![Fig. 6. The 10 institutions engaging most actively in research on MBSE for Industry 4.0.](image2)
Regarding the most popular venues for publications on MBSE for Industry 4.0, we found that the publications are almost equally spread over the different journals, whereas the conference on Emerging Technologies And Factory Automation (ETFA) appears to be the most important conference for this field. It is, however, notable that no software-engineering-centric venues are among the top venues. We also found that Industry 4.0 is still largely a European research program with common contributions from the USA, China, and the Republic of Korea. Future studies will reveal whether Industry 4.0 will be addressed increasingly in other countries as well.

UML is the modeling technique applied most often to Industry 4.0, followed by DSLs and knowledge representation techniques. There is a conceptual gap between the software modeling techniques of UML and knowledge representation techniques whose bridging appears to be crucial to successful modeling the factory of the future. Domain-specific languages are the second most popular modeling technique. With the majority (82%) of DSL related contributions being published since 2010 and increasing in number since then, we expect to see more modeling techniques specifically tailored to Industry 4.0 in the future. It would be interesting to compare this trend to more mature domains that incorporated modeling, such as embedded systems in automotive or avionics, as well as to the corpora of MBSE-related publications on CPS and IoT.

Future research on MBSE for Industry 4.0 hence should address the integration of knowledge representation techniques with UML and AutomationML to ultimately achieve an integrated factory system model that enables fully automated reasoning about production optimization based on product information. Such a basis might lead to increasing research in validation & verification of integrated production systems as well as in including modeling the human factors of Industry 4.0 applications more often. Moreover, research is lacking metrics that enable checking whether Industry 4.0 fulfills the promises of reducing integration costs and energy consumption. Development of appropriate measures can help to guide future research in MBSE for Industry 4.0. Overall, this study has uncovered that research on modeling for Industry 4.0 currently focuses on digital representation and integration using established techniques, but also that there is an increasing trend towards customized modeling techniques to address the more specific challenges of Industry 4.0. Future studies will show whether established modeling techniques are adequate to address Industry 4.0 challenges. There also still is a lack of benchmarks for Industry 4.0. Consequently, there is no real comparison between approaches and only little work on their validation. Sharing benchmarks on top of existing case studies, such as MyJoghur [3] would be another challenge for the modeling community.
