







Analiza danych i optymalizacja w Przemysle 4.0

Jerzy Baranowski , Edyta Kucharska , Waldemar Bauer ,
Katarzyna Grobler-Dębska , Nataliai Kashpruk ,
Marta Kraszewska , Rafał Mularczyk ,
Cezary Piskor-Ignatowicz , Adrian Dudek , Daniel Dworak ,
Jan Kapusta , Kazimierz Kawa 

AGH Akademia Górniczo-Hutnicza, Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej, Kraków

Streszczenie: W pracy w skrótovej formie omówiono zagadnienia Przemysłu 4.0 i to, w jakim zakresie badane są one w Laboratorium Informatyki w Sterowaniu i Zarządzaniu w Katedrze Automatyki i Robotyki. Omawiane są kwestie zarządzania przedsiębiorstwem, optymalizacji i diagnostyki procesów. W podsumowaniu wskazano na powiązanie tej tematyki z kształceniem.

Słowa kluczowe: Przemysł 4.0, analiza danych, optymalizacja

DATA ANALYSIS AND OPTIMIZATION IN INDUSTRY 4.0

Abstract: Industry 4.0 issues and the extent to which they are developed in the Laboratory of Computer Science in Control and Management at the Department of Automatic Control and Robotics are discussed in a short form. We discuss business management, process optimization and diagnostics. The work ends with a summary indicating the connection of this subject with education.

Keywords: Industry 4.0, data analysis, optimization

https://doi.org/10.7494/978-83-67427-00-5_3

1. Wprowadzenie

Przemysł 4.0 oznacza przełomowe zmiany spowodowane przez rozwój internetu rzeczy oraz systemów cyberfizycznych w produkcji i logistyce. Termin ten obejmuje innowacje dotyczące inteligentnych produktów i procesów, głębszej integracji wzdłuż łańcucha dostaw oraz technologii cyfrowej w produkcji. Celem jest szybsza reakcja na coraz bardziej zmienne wymagania klientów. Prowadzi to do zwiększenia konkurencyjności na dynamicznych rynkach. W artykule przedstawiono zagadnienia związane z Przemysłem 4.0 realizowane w Laboratorium Informatyki w Sterowaniu i Zarządzaniu¹ Katedry Automatyki i Robotyki Akademii Górniczo-Hutniczej. W szczególności skupiono się na zagadnieniach wspomagania zarządzania, optymalizacji i diagnostyce procesów.

2. Przemysł 4.0

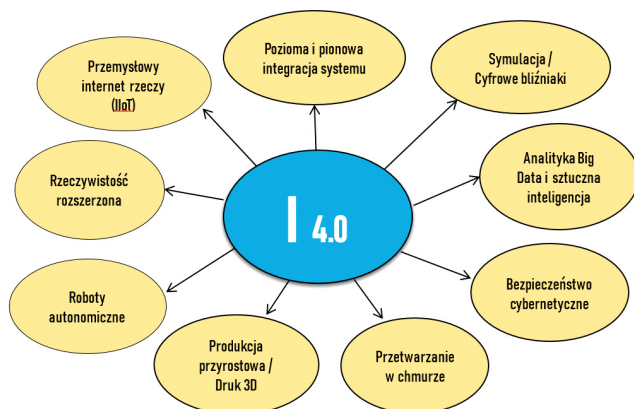
Industy 4.0 to połączenie automatyzacji i wymiany danych w różnych technologiach. Podstawę stanowią systemy cyberfizyczne wykorzystujące nowoczesne systemy sterowania z wbudowanym oprogramowaniem połączonym przez IoT (Palwe i Sirsikar 2021).

Podstawowymi dziesięcioma technologiami, czyli filarami Przemysłu 4.0 są (Reis i Geert 2017, Para i in. 2018) (rys. 1):

- **Analityka Big Data** – wykorzystanie statystyk, wykresów i analizy danych do pomiaru wydajności i dostrajania systemów między innymi w celu optymalizacji produktywności.
- **Roboty autonomiczne** – roboty, które mogą funkcjonować samodzielnie przy minimalnym nadzorze. Zastosowanie ich umożliwiło powstanie „ciemnych fabryk” (znanych również jako produkcja Lights-Out), w których maszyny wykonują całą produkcję przy minimalnym wykorzystaniu kadry i kontynuacji pracy w godzinach nocnych.
- **Symulacja** – korzystanie z pomocy wizualnych i oprogramowania do symulacji złożonych procesów i projektów stało się powszechne w każdej firmie inżynierskiej lub produkcyjnej.
- **Pozioma i pionowa integracja systemu** – spójność procesów w firmie uzyskuje się dzięki wdrożeniu sieci integracji danych, co także znacznie ułatwia współpracę między przedsiębiorstwami.

¹ <https://isz.agh.edu.pl>

- **Przemysłowy internet rzeczy (IIoT)** – powszechne stosowanie komputerów wbudowanych pozwala na decentralizację, ponieważ każde urządzenie może komunikować się zdalnie i dostarczać dane w mgnieniu oka. Proces decentralizacji znacznie zmniejsza czas odpowiedzi w czasie rzeczywistym.
- **Bezpieczeństwo cybernetyczne** – zwiększona łączność wiąże się ze zwiększonym ryzykiem naruszenia bezpieczeństwa, a to powoduje konieczność zastosowania cyberbezpieczeństwa i odpowiednich środków w celu ochrony informacji.
- **Chmura** – udostępnianie danych i przejrzystość mają kluczowe znaczenie dla Przemysłu 4.0. Przyjęcie technologii chmury pozwala na bezproblemowe udostępnianie danych i lepszą łączność.
- **Produkcja dodatkowa** – w ramach procesu projektowania drukowanie i modelowanie 3D stały się niezbędne do budowania prototypów i lekkich projektów, co pozwala firmom odzwierciedlić symulacje w rzeczywistości.
- **Rzeczywistość rozszerzona (AR)** – to ostateczna granica paradygmatu Przemysłu 4.0. Obecne zastosowania obejmują wybór produktów w magazynach i wizualizację możliwych opcji projektowych. W przyszłości z pewnością nastąpi rozwój tej technologii w celu dostarczenia pracownikom różnorodnych potrzebnych danych i informacji w czasie rzeczywistym.

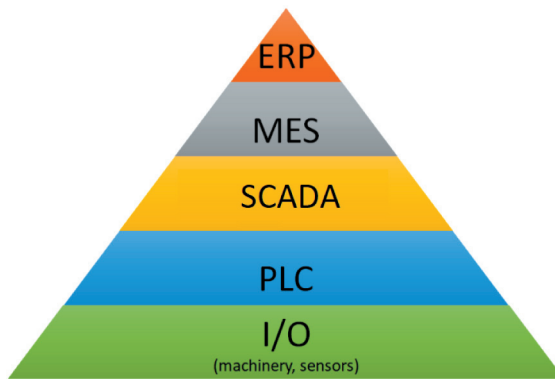


Rys. 1. Podstawowe technologie Przemysłu 4.0

3. Systemy wspomaganie zarządzania

Wraz z rozwojem Przemysłu 4.0 firmy dążą do umożliwienia przeprowadzenia całego procesu produkcyjnego z minimalnym udziałem ludzi. W tym celu duży nacisk kładzie się na rozwój automatyzacji komunikacji z wykorzystaniem systemów cyberfizycznych (rys. 2). Ważną rolę obecnie odgrywa wspomaganie przez technologie

Przemysłu 4.0 planowania zasobów przedsiębiorstwa w systemach ERP (*Enterprise Resources Planning*). Systemy ERP są połączone urządzeniami wejściowymi i wyjściowymi, takimi jak skanery, urządzenia drukujące i komunikacyjne, a także z innymi systemami sprzętowo-programowymi w przedsiębiorstwie. System ERP jest istotną częścią systemu cyberfizycznego (Majstorovic i in. 2020) zawierającą dane główne i transakcyjne organizacji. Aktualnym trendem jest w systemach ERP wspomaganie planowania i realizacji procesów biznesowych przedsiębiorstwa dzięki zmniejszeniu manualnych czynności użytkowników. Rozwój systemów ERP wraz z technologiami Przemysłu 4.0 w efekcie zautomatyzowanego planowania zasobów zmierza również w kierunku optymalizacji procesów produkcyjnych, łańcucha dostaw i operacji logistycznych, usprawnienia śledzenia zapasów oraz zarządzania pracownikami. Bardzo często są również integrowane z bardziej złożonymi systemami planowania i realizacji produkcji takimi jak APS i MES (Kucharska i in. 2015).



Rys. 2. Automatyzacja stosowana na różnych poziomach w fabryce: poziom 0 (I/O) – urządzenia i czujniki dostarczające elementarne dane z produkcji, poziom 1 (PLC) – sterowanie urządzeniami za pomocą sterowników PLC, poziom 2 (SCADA) – monitorowanie i nadzór systemów produkcyjnych, poziom 3 (MES) – planowanie procesów produkcyjnych od surowców do gotowego produktu, poziom 4 (ERP) – planowanie wszystkich zasobów przedsiębiorstwa, łącznie z finansami, zasobami ludzkimi i inwestycjami

Przemysł 4.0 wpłynął również na rozwój systemów zarządzania relacjami z klientem (CRM). Metody analizy danych z wykorzystaniem algorytmów uczenia maszynowego umożliwiają przetwarzanie i analizowanie ustrukturyzowanych i nieustrukturyzowanych danych o klientach. Ponadto pozwalają wykrywać ukryte wzorce i pozyskiwać cenne informacje o klientach. W ten sposób można wykryć preferencje zakupowe klientów. Zastosowanie takich rozwiązań znacznie poprawia jakość obsługi klienta.

Nowoczesne systemy ERP przenoszą się do chmury. Jest to związane z bezpieczeństwem danych i ułatwieniem przetwarzania ogromnej liczby danych pozyskanych również dzięki technologiom Przemysłu 4.0 (Tongsuksai 2020). Przetwarzanie w chmurze umożliwia efektywniejsze przetwarzanie danych, obniżenie kosztów IT, szybszą mi-

grację i integrację z innymi narzędziami biznesowymi. W tym momencie bardzo istotnym elementem jest również opracowanie odpowiednich metodologii wdrożeniowych adekwatnych do nowoczesnych projektów w zmienionym środowisku (Domagała i in. 2021). Nowoczesne systemy ERP stały się niezbędną częścią nowoczesnych fabryk.

Ponadto jednym z ważnych aspektów automatyzacji jest coraz częstsze zastosowanie przetwarzania języka naturalnego (NPL – *Natural Language Processing*) w Przemysle 4.0. W szczególności rozwija się zagadnienie dotyczące automatycznej klasyfikacji dokumentacji związanej z rozwojem oprogramowania, finansami oraz procesami biznesowymi, takimi jak planowanie projektów czy zarządzanie sprzętem i materiałami (Shui-Shun 2009, Gupta i in. 2020).

5. Optymalizacja

Czwarta rewolucja przemysłowa ma na celu zwiększenie wydajności systemów produkcyjnych i dlatego wymagane jest zastosowanie różnych metod optymalizacji do rozwiązania coraz trudniejszych problemów na liniach produkcyjnych i w procesach logistycznych.

Przemysł 4.0 jest również ściśle związany z koncepcją ZDM (*Zero-Defect Manufacturing*) (Psarommatitis i in. 2022). ZDM jest strategią kompleksową. W perspektywie krótkoterminowej obejmuje opracowanie i wdrożenie systemu kontroli procesu w czasie rzeczywistym do wyeliminowania produkcji wadliwych komponentów z powodu różnic w materiałach, komponentach i właściwościach procesu. Perspektywa długoterminowa dotyczy minimalizowania wszystkich awarii dzięki ciągłej optymalizacji procesu produkcyjnego i systemu wytwarzania. Koncepcja „zero błędów” może być praktycznie wykorzystana w każdym środowisku produkcyjnym w celu poprawy jakości i zmniejszenia kosztów. Dlatego cały system powinien zawierać następujące główne fazy: faza 1 – pozyskiwanie danych (automatyczne przechwytywanie, czyszczenie i formatowanie odpowiednich danych przy użyciu systemu inteligentnych czujników), faza 2 – przetwarzanie sygnału (automatyczne przetwarzanie sygnału, filtrowanie i ekstrakcja cech), faza 3 – diagnostyka (eksploracja danych i odkrywanie wiedzy na potrzeby diagnostyki), faza 4 – ocena prognostyczna (eksploracja danych i odkrywanie wiedzy do celów prognostycznych) oraz faza 5 – harmonogram serwisu i produkcji (narzędzie optymalizacyjne do dynamicznego harmonogramowania w warunkach niepewności).

Problemy harmonogramowania, które spełniają założenia strategii ZDM, uwzględniają podejmowanie decyzji w warunkach niepewności. Tego rodzaju harmonogram powinien być predykcyjno-reaktywny. Do takiej optymalizacji można wykorzystać **metodę przełączania modeli algebraiczno-logicznych**, która zawiera elementy umożliwiające zamodelowanie rozwiązania różnych problemów z zaburzeniami w postaci uszkodzeń produktów lub awarii maszyn produkcyjnych (Grobler-Dębska i in. 2021).

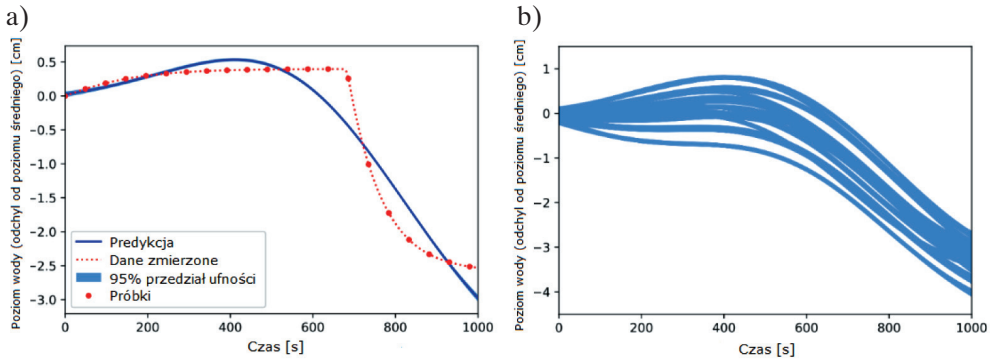
6. Diagnostyka procesów przemysłowych

Skuteczne i niezawodne monitorowanie oraz diagnostyka instalacji sterowania procesami mają ogromne znaczenie dla efektywnego zarządzania przedsiębiorstwami. Algorytmy do wykrywania i izolacji anomalii umożliwiają zwiększenie żywotności systemu, zmniejszenie przerw w jego pracy oraz mogą przyczynić się do znacznych oszczędności długoterminowych. Główna trudność w ich rozwoju polega na tym, że takie instalacje mają zazwyczaj dość wysoki poziom złożoności. Są zwykle nieliniowe i mają na nie wpływ zaburzenia stochastyczne oraz zmiany parametrów. Trudność lub brak możliwości zastosowania podstawowych modeli na szerszą skalę powoduje konieczność opracowania metod związanych z modelami statystycznymi lub uczeniem maszynowym.

Zazwyczaj uczenie maszynowe i wykorzystanie modeli *data-driven* (oparte na danych) umożliwiają zbudowanie skomplikowanych modeli typu *black-box* (czarna skrzynka), które są wysoce nieprzejryste oraz trudne do zinterpretowania. Wady te są mniej zauważalne w podejściach statystycznych. To powoduje większe zainteresowanie nimi w tej dziedzinie. Modele z obu tych grup nie radzą sobie w typowych sytuacjach, gdy rzeczywiste dane dotyczące usterek systemu są niezwykle rzadkie, a nawet jeśli występują, to są często niekompletne. Dlatego ważne jest badanie metod radzenia sobie z kwestiami reprezentatywnych lub brakujących danych (Baranowski i in. 2021).

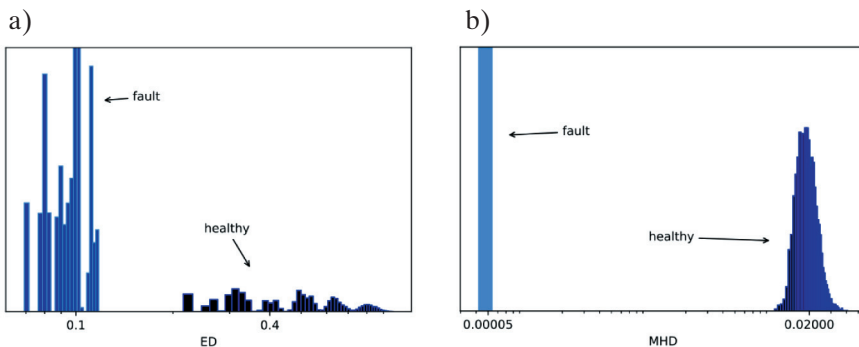
Dziedzina wykrywania oraz diagnozowania anomalii i stanów awarii jest zdominowana przez metody statystyczne, oparte na wariantach PCA (analizy głównych składowych) lub metodach jądrowych. Wykorzystywane metody uczenia maszynowego obejmują szeroki zakres klasyfikatorów kNN, przez SVD po sieci neuronowe (Cho i Jiang 2019). Podejścia probabilistyczne obejmują natomiast PCA lub metody jądrowe, takie jak CVA (Jiang i in. 2019). Przeprowadzone badania natomiast skupiały się na wykorzystaniu procesów gaussowskich wraz z analizą głębi danych w celu wypełnienia luki w diagnostyce stanów przejściowych procesów.

Proces gaussowski (PG) (rys. 3) jest procesem stochastycznym używanym do modelowania danych, które były obserwowane w czasie, przestrzeni lub obu naraz. Główną cechą charakteryzującą procesy gaussowskie jest to, że są one uogólnieniem normalnych rozkładów prawdopodobieństwa, gdzie każdy z nich opisuje zmienną losową (Rasmussen i Williams 2006). PG jednoznacznie opisuje odpowiednia funkcja kowariancji (reprezentująca wiedzę badacza o danym problemie) wraz z jej parametrami oraz średnia. W badaniach procesy były modelowane za pomocą procesów gaussowskich osobno w przypadku stanów normalnych oraz stanów awarii. Wykorzystując własności PG, mamy możliwość wygenerowania dziesiątek tysięcy nowych próbek (przez próbkę rozumiemy cały przebieg czasowy). Takie podejście umożliwia nam wykorzystanie głębi danych na wygenerowanych informacjach w celu utworzenia algorytmu wykrywania stanów awarii.



Rys. 3. Wynik dopasowania PG (dostosowywania parametrów) do posiadanych danych (a) oraz skumulowany wynik wygenerowanych danych za pomocą utworzonego modelu (b)

Głębkie danych określamy jako odległość pomiarów od środka punktu $x \in R^d$ względem rozkładu funkcji F . Najbardziej oddalone obserwacje mają mniejsze wartości niż dane znajdujące się blisko środka. Wszelkie wartości są wyznaczone za pomocą funkcji głębokości. Właśnie w dany sposób można określić, które z badanych próbek są anomaliami w konkretnym eksperymencie. Istnieje wiele metod głębi danych (m.in. Euklidesa, Mahalanobisa, Turkeya, Oji). W zależności od danych można wychwycić lepsze wyszukanie i rozdzielanie anomalii w eksperymencie. Podczas naszych badań skupiliśmy się na dwóch z nich, głębi Euklidesa oraz Mahalanobisa (rys. 4). Można zauważyć, że w obu przypadkach mamy podział na dwa histogramy danych. Oznacza to, że wykorzystane metody wychwytyją różnicę pomiędzy danymi zdrowymi a anomaliami (Gijbels i Nagy 2017).



Rys. 4. Histogramy wartości głębokości danych metodą Euklidesa (a) oraz Mahalanobisa (b)

Przedstawiliśmy wartości w przypadku danych zdrowych (*healthy*) i danych z anomalią (*fault*). Zauważalna jest różnica wartości pomiędzy nimi. W celu pokazania małych wartości oś poziomą przedstawiono w skali logarytmicznej.

Niniejsze badania pokazują wykorzystanie procesów gaussowskich oraz głębi danych do wykrywania anomalii w procesach. Jednym z głównych problemów są wartości uzyskane z metod głębi danych. W przypadku niektórych z metod nie da się ich zinterpretować. Dalsze prace będą opierać się na analizie i badaniu innych metod głębi danych. Dodatkowo będą wykorzystywane obliczenia i metody probabilistyczne w celu wyznaczenia i wyłapania niepewności.

8. Podsumowanie

W niniejszej pracy przedstawiliśmy ułamek zagadnień związanych z Przemysłem 4.0, które są przedmiotem badań w Katedrze Automatyki i Robotyki. Warto jednak zaznaczyć, że tematyka ta jest rozwijana bardzo dynamicznie. Przykładem tego jest uruchomienie wspólnie z Wydziałem Odlewnictwa AGH nowego kierunku studiów pierwszego stopnia technologii Przemysłu 4.0 od cyklu rekrutacyjnego 2022/2023. W ramach tego kierunku studenci poznają zagadnienia łączące Przemysł 4.0, automatykę i inżynierię materiałową. W latach 2022–2023 odbyły się lub odbędą również szkoły letnie oraz zimowe dla obcokrajowców w ramach projektu Vacation Learning at AGH for Sustainable Industry 4.0 education².

Warto również podkreślić, że zagadnienia związane z Przemysłem 4.0 badane są w Laboratorium Informatyki w Sterowaniu i Zarządzaniu w ramach licznych projektów finansowanych z NCN, NCBR, NAWA i środków europejskich.

Literatura

- Baranowski J., Dudek A., Mularczyk R., 2021, *Transient Anomaly Detection Using Gaussian Process Depth Analysis*, [w:] *MMAR 2021: 2021 25th International Conference on Methods and Models in Automation & Robotics: August 23–26, 2021, Międzyzdroje, Poland*, IEEE, Piscataway, s. 221–226. <https://doi.org/10.1109/MMAR49549.2021.9528470>.
- Cho S., Jiang J., 2019, *A fault detection and isolation technique using nonlinear support vectors dichotomizing multi-class parity space residuals*, *Journal of Process Control*, vol. 82, s. 31–43. <https://doi.org/10.1016/j.jprocont.2019.07.006>.
- Domagała A., Grobler-Dębska K., Wąs J., Kucharska E., 2021, *Post-Implementation ERP Software Development: Upgrade or Reimplementation*, *Applied Sciences*, vol. 11(11), 4937. <https://doi.org/10.3390/app11114937>.
- Garcia R., Luger R., Renault E., 2003, *Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables*, *Journal of Econometrics*, vol. 116(1–2), s. 49–83. [https://doi.org/10.1016/S0304-4076\(03\)00103-9](https://doi.org/10.1016/S0304-4076(03)00103-9).

² <http://vacation.agh.edu.pl>

- Ghysel E., Tauchen G., 2003, *Frontiers of Financial Econometrics and Financial Engineering*, Journal of Econometrics, vol. 116(1–2), s. 1–7. [https://doi.org/10.1016/S0304-4076\(03\)00101-5](https://doi.org/10.1016/S0304-4076(03)00101-5).
- Gijbels I., Nagy S., 2017, *On a general definition of depth for functional data*, Statistical Science, vol. 32, no. 4, s. 630–639. <https://doi.org/10.1214/17-STS625>.
- Grobler-Dębska K., Kucharska E., Baranowski J., 2021, *Formal scheduling method for zero-defect manufacturing*, International Journal of Advanced Manufacturing Technology, vol. 118(11–12), s. 4139–4159. <https://doi.org/10.1007/s00170-021-08104-0>.
- Gupta A., Dengre V., Kheruwala H.A., Shah M., 2020, *Comprehensive review of text-mining applications in finance*, Financial Innovation, vol. 6, 39. <https://doi.org/10.1186/s40854-020-00205-1>.
- Jiang Q., Yan X., Huang B., 2019, *Neighborhood variational Bayesian multivariate analysis for distributed process monitoring with missing data*, IEEE Transactions on Control Systems Technology, vol. 27(6), s. 2330–2339. <https://doi.org/10.1109/TCST.2018.2870570>.
- Kucharska E., Grobler-Dębska K., Gracel J., Jagodziński M., 2015, *Idea of Impact of ERP-APS-MES Systems Integration on the Effectiveness of Decision Making Process in Manufacturing Companies*, [w:] Kozielski S., Mrozek D., Kasprowski P., Małysiak-Mrozek B., Kostrzewa D. (eds.), *Beyond Databases, Architectures and Structures: BDAS 2015: Ustroń, Poland, May 26–29, 2015: proceedings*, Communications in Computer and Information Science, vol. 521, Springer, Cham, s. 551–564. https://doi.org/10.1007/978-3-319-18422-7_49.
- Majstorovic V., Stojadinovic S., Lalic B., Marjanovic U., 2020, *ERP in Industry 4.0 Context*, [w:] Lalic B., Majstorovic V., Marjanovic U., von Cieminski G., Romero D. (eds.), *Advances in Production Management Systems: The Path to Digital Transformation and Innovation of Production Management Systems. IFIP WG 5.7 International Conference, APMS 2020. Novi Sad, Serbia, August 30 – September 3, 2020. Proceedings, Part I*, Advances in Information and Communication Technology, vol. 591, Springer, Cham, s. 287–294. https://doi.org/10.1007/978-3-030-57993-7_33.
- Palwe S., Sirsikar S., 2021, *Industry 4.0 Technologies and Their Applications in Fighting COVID-19*, [w:] Agrawal R., Mittal M., Goyal L.M. (eds.), *Sustainability Measures for COVID-19 Pandemic*, Springer, Singapore, s. 237–251. https://doi.org/10.1007/978-981-16-3227-3_13.
- Para J., Del Ser J., Aguirre A., Nebro A.J., 2018, *Decision Making in Industry 4.0 Scenarios Supported by Imbalanced Data Classification*, [w:] Del Ser J., Osaba E., Bilbao M.N., Sanchez-Medina, Vecchio M., Yang X.-S. (eds.), *Intelligent Distributed Computing XII*, Studies in Computational Intelligence, vol. 798, Springer, Cham, s. 121–134. https://doi.org/10.1007/978-3-319-99626-4_11.

- Psarommatis F., Sousa J., Mendonça J.P., Kiritsis D., 2022, *Zero-defect manufacturing the approach for higher manufacturing sustainability in the era of industry 4.0: a position paper*, International Journal of Production Research, vol. 60, s. 73–91. <https://doi.org/10.1080/00207543.2021.1987551>.
- Rasmussen C., Williams C.K.I., 2006, *Gaussian Processes in Machine Learning*, MIT Press, Cambridge, Massachusetts.
- Reis M.S., Geert G., 2017, *Industrial Process Monitoring in the Big Data/Industry 4.0 Era: from Detection, to Diagnosis, to Prognosis*, Processes, vol. 5, no. 3, 35. <https://doi.org/10.3390/pr5030035>.
- Shui-Shun L., 2009, *A document classification and retrieval system for R&D in semiconductor industry – A hybrid approach*, Expert Systems with Applications, vol. 36(3), part 1, s. 4753–4764. <https://doi.org/10.1016/j.eswa.2008.06.024>.
- Tongsuksai S., Mathrani S., 2020, *Integrating Cloud ERP Systems with New Technologies Based on Industry 4.0: A Systematic Literature Review*, [w:] *IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, IEEE, Piscataway, s. 1–6. <https://doi.org/10.1109/CSDE50874.2020.9411570>.