Research & Development

2026

Mechanical Engineering Letters 2026

Technical-Scientific Journal supported by the Institute of Technology, Hungarian University of Agriculture and Life Science (MATE), Gödöllő, published by GATE Nonprofit Ltd.

HU ISSN 2060-3789

Páter K. u. 1., Gödöllő, H-2103 Hungary

Volume 28 (2026)

Contents

Design of physics-based mathematical building model for the investigation of phase change materials integrated into heat emitting surfaces

Bence KAZÁR[1,2], Márta SZABÓ[1]

[1] *Department of Building Services and Energy Engineering, Institute of Technology, Hungarian University of Agriculture and Life Sciences*, MATE, Gödöllő

[2] *Doctoral Program of Mechanical Engineering, Hungarian University of Agriculture and Life Sciences,* MATE, Gödöllő

Abstract

The possibilities for applying phase change materials (PCMs) in buildings include their use in solid envelope structures, in thermal energy storage units of heating systems, and integrated into heat emission surfaces. These heat emission surfaces can range from simple electric or hot water radiators to surface heating systems integrated into building envelope components. To investigate the impact of PCMs on thermal comfort, energy efficiency, and control characteristics, it is essential to establish one or more mathematical models. In this paper, I examine the process of developing such a model.

Keywords

phase Change Materials (PCM), mathematical modeling, effective heat capacity method, heat emitting surfaces

1. Introduction

Possibilities for applying phase change materials (PCMs) in buildings include integration into solid building envelopes, thermal storage units of heating systems, and heat emission surfaces. These emission surfaces can be conventional electric or hot-water radiators, or radiant surface heating systems integrated into the building structures. To investigate the impact of PCMs on thermal comfort, energy efficiency, and control characteristics, it is essential to establish one or more mathematical models. This paper examines the process of developing such a model. The objective of the mathematical model is to simulate the temporal behavior of an arbitrary building both without PCMs and with PCMs installed near various heat emitters. A well-functioning model enables the analysis of indoor temperature fluctuations over time, the selection of the ideal control strategy, and the determination of heating energy consumption. The mathematical model I have developed is physics-based, meaning every parameter possesses physical significance.

2. Mathematical model

2.1. Derivation of the Governing Equations

The first step is the modeling of the building envelope structures. Since the objective is to investigate time-dependent processes, the governing equation is a differential equation. Given that temperatures vary across every layer of the structures at any given time, the process is described by a partial differential equation (PDE), specifically the Fourier heat conduction equation. The analytical solution of this PDE presents significant challenges. In many cases, the time-varying profile of the outdoor air temperature can only be approximated with great difficulty as a combination of elementary functions. Therefore, instead of an analytical approach, a numerical solution is sought. The explicit finite difference method (EFDM) provides a viable means for numerical solution, and the necessary calculations can be performed using Excel spreadsheets. In the following sections, the discrete formulations used for the solution are presented.
The notations used in the following sections are illustrated in Table 1.

Table 1. Nomenclature and symbols used

Physical Quantity	Symbol	Unit
Specific heat	c	$\frac{J}{kgK}$ or $\frac{J}{kg°C}$
Density	ρ	$\frac{kg}{m^3}$
Thermal conductivity	λ	$\frac{W}{mK}$ or $\frac{W}{m°C}$
Temperature	T	$°C$
Time	t	s
Thermal diffusivity	a	$\frac{m^2}{s}$
Heat capacity	C	$\frac{J}{K}$ or $\frac{J}{°C}$
Thickness	x	m

The first structure is the facade wall. I divide the wall into *n* layers of equal thickness, parallel to the plane of the wall. Only one-dimensional heat flow occurs through the wall. In the current model, the wall is homogeneous and isotropic. Its surface area, thickness, thermal conductivity, density, and specific heat capacity are known. Furthermore, the average values of the external and

internal convective heat transfer coefficients are also known (which, in reality, depend on several factors such as airflow and wind).

Wall thickness: x [m]
Layer thickness: $\Delta x=\frac{x}{n}$ [m]
Time step: Δt [s]
The time step must be chosen such that it satisfies the stability criterion. I will address this in a later section.

Determination of the thermal diffusivity:

$$a = \frac{\lambda}{\rho * c}\left[\frac{m^2}{s}\right] \tag{1}$$

I determine the temperature of every layer of the structure. (The node under investigation is always located at the center of the layer.) In the case of the two outermost layers, determining the mid-layer temperature is not sufficient; the surface temperatures must also be determined.

Determination of the temperature for the internal layers of the structure:

$$T_i^{k+1} = T_i^k + a * \frac{\Delta t}{(\Delta x)^2} * \left(T_{i-1}^k - 2 * T_1^k + T_{i+1}^k\right)[°C] \tag{2}$$

Where:
T_i^k is the temperature of the i-th layer at the k-th time step

Determination of the temperature at the center of the outermost (external) layer of the structure:

$$T_1^{k+1} = T_1^k + a * \frac{\Delta t}{\Delta x} * \left(\frac{T_{surface,1}^k - T_1^k}{\frac{\Delta x}{2}} + \frac{T_2^k - T_1^k}{\Delta x}\right)[°C] \tag{3}$$

In this correlation, the term on the left side of the parentheses refers to the heat transport process between the external surface and the center of the outermost layer, while the right side refers to the heat transport process between the center of the second layer and the center of the first layer. The centers of the two layers are at a distance of Δx from each other, whereas the distance between the external surface and the center of the first layer is only $\frac{\Delta x}{2}$.

Determination of the temperature at the center of the innermost (internal) layer of the structure:

$$T_n^{k+1} = T_n^k + a * \frac{\Delta t}{\Delta x} * \left(\frac{T_{surface,n}^k - T_n^k}{\frac{\Delta x}{2}} + \frac{T_{n-1}^k - T_n^k}{\Delta x}\right)[°C] \tag{4}$$

Principle of determining the external surface temperature:

Since the temperature assigned to the outermost layer also refers to its center, it is not identical to the surface temperature. I determine the external surface temperature by extrapolation. I assume that the heat flux density from the layer center towards the external surface (located at a distance of $\frac{\Delta x}{2}$) is equal to the heat flux density passing through the convective boundary layer. (In reality, during a time-varying process, there will be a minimal difference between the two heat flux densities due to the heat absorbed by or released from the thin layer. If the time step and the thickness of the elementary layers form a refinement sequence approaching infinity, this difference converges to zero.)

The external surface temperature can be determined using the following correlation:

$$\left(T^k_{surface,1} - T^k_{outdoor\ air}\right) * \alpha_{out} = \left(T^k_1 - T^k_{surface,1}\right) * \frac{\lambda}{\frac{\Delta x}{2}} \tag{5}$$

$$T^k_{surface,1} = \frac{\frac{2\lambda}{\Delta x} * T^k_1 + \alpha * T^k_{outdoor\ air}}{\frac{2\lambda}{\Delta x} + \alpha_{out}}\ [°C] \tag{6}$$

Determination of the internal surface temperature:

$$\left(T^k_{indoor\ air} - T^k_{surface,n}\right) * \alpha_{in} = \left(T^k_{surface,n} - T^k_n\right) * \frac{\lambda}{\frac{\Delta x}{2}} \tag{7}$$

$$T^k_{surface,n} = \frac{\frac{2\lambda}{\Delta x} * T^k_n + \alpha_b * T^k_{indoor\ air}}{\frac{2\lambda}{\Delta x} + \alpha_{in}}\ [°C] \tag{8}$$

The objective of developing the model is to investigate how the indoor air temperature evolves as a function of time. Therefore, the output variable under investigation is the indoor air temperature.

Determination of the indoor air temperature:

$$T^{k+1}_{indoor\ air} = T^k_{in} + \frac{\Delta t}{V * \rho_{air} * c_{air}} * \left(\left(\sum_{j=1}^{m}\left(A_j * \alpha_{in,j} * \left(T^k_{surface,in,j} - T^k_{indoor\ air}\right)\right)\right) + Q^k\right)\ [°C] \tag{9}$$

Where:
Q^k is the additional thermal power (heat gain or loss) of the air at the *k*-th time step (e.g., heating) [W]
A_j is the surface area of the *j*-th building envelope (in contact with the indoor

air) [m^2]
$\alpha_{in,j}$ is the heat transfer coefficient at the boundary between the j-th surface and the indoor air $\left[\frac{W}{m^{2\circ}C}\right]$

The indoor air temperature depends on the magnitude of the heat flows absorbed or released by the air, the heat capacity of the air, and the duration of these heat flows. This correlation assumes that the air in the room is perfectly mixed. The thermal power output of fully convective heating can be determined based on the surface area of the heat emitter, its surface temperature, the air temperature, and the heat transfer coefficient.

The thermal power output of fully convective heating:

$$Q^k = A_{heat\ emitter} * \alpha^k_{heat\ emitter,air} * \left(T^k_{heat\ emitting\ surface} - T^k_{indoor\ air}\right)[W] \quad (10)$$

Where:
$\alpha^k_{heat\ emitter,air}$ is the heat transfer coefficient between the surface of the heat emitter and the indoor air at the k-th time step $\left[\frac{W}{m^{2\circ}C}\right]$
$T^k_{indoor\ air}$ is the bulk temperature of the indoor air at the k-th time step [°C]

In the case of commercially available radiators, manufacturers typically specify the heat output for supply and return water temperatures and indoor air temperatures as fixed in the EN 442/2 standard (specifically at 75/65/20 °C). The heat output of the radiator is typically governed by the air-side heat transfer coefficient. Under forced circulation, the water-side heat transfer coefficient is generally two orders of magnitude higher than the air-side coefficient. Furthermore, even when the radiator is turned off and the flow ceases, there is typically an order of magnitude difference between the water-side and the air-side heat transfer coefficients. Radiators are made of steel or aluminum with a thickness of a few millimeters; the thermal resistance of these materials is also orders of magnitude lower than that of the internal heat transfer coefficient. The air-side heat transfer coefficient depends on the mean surface temperature of the radiator, the indoor air temperature, and various geometric characteristics.

The air-side heat transfer coefficient can be determined as follows:

$$\alpha = \alpha_{nom.} * \left(\frac{\Delta T_{ln}}{\Delta T_{nom.ln}}\right)^{n_{kit}} \left[\frac{W}{m^2K}\right] \quad (11)$$

Where:
n_{kit} is the characteristic exponent of the heat emitter (manufacturer's catalog data) [-]
$\Delta T_{nom.ln.}$ is the logarithmic temperature difference at the temperature values specified by the manufacturer (nominal mean excess temperature of the heating

water) [°C]
ΔT_{ln} is the logarithmic temperature difference in the case under investigation (actual mean excess temperature of the heating water) [°C]
$\alpha_{nom..}$ is the heat transfer coefficient valid at the temperature values specified by the manufacturer (at $\Delta t_{nom.ln}$) $\left[\frac{W}{m^2K}\right]$
α is the heat transfer coefficient in the case under investigation (at Δt_{ln}) $\left[\frac{W}{m^2K}\right]$
Note: When the radiator is turned off and the flow ceases, the above correlation is slightly modified: the logarithmic mean temperature difference ΔT_{ln} is replaced by the temperature excess of the radiator surface.

The surface temperature of the thin-walled radiator closely approximates the temperature of the heating water.

The Logarithmic Mean Temperature Difference (LMTD) (mean excess temperature relative to the indoor air temperature):

$$\Delta T_{ln} = \frac{T_{supply} - T_{return}}{ln\frac{T_{supply} - T_{indoor\ air}}{T_{return} - T_{indoor\ air}}}\ [°C] \quad (12)$$

Where:
T_{supply} is the supply water temperature (which approximates the radiator surface temperature near the inlet connection) [°C]
T_{return} is the return water temperature (which approximates the radiator surface temperature near the outlet connection) [°C]
$T_{indoor\ air}$ is the indoor air temperature [°C]
Note: The above correlation for the logarithmic mean temperature difference is used for hot-water radiators with active fluid flow. When the radiator is turned off, the fluid flow ceases; therefore, this correlation must be slightly modified.

If the radiator surface temperature is uniform throughout, then the excess temperature of the radiator surface relative to the indoor air temperature is:

$$\Delta T_{excess} = T_{rad,surface} - T_{indoor\ air}\ [°C] \quad (13)$$

In the current model (in the initial phase of modeling), I used this simplified correlation. I modeled the radiator thermally as a thin body; therefore, I did not divide it into layers and considered the temperature of all its points to be identical. While this simplification closely approximates reality for radiators without PCM, it represents a significant simplification for radiators equipped with PCM. Therefore, the future objective is to develop a model that more closely approximates reality.

Based on these simplifications, the surface temperature of the heat emitter can be described as follows:

$$T_{rad}^{k+1} = T_{rad}^{k} - \frac{\Delta t}{C_{rad}^{k}} * \left(A_{rad} * \alpha_{rad,air}^{k} * \left(T_{rad}^{k} - T_{indoor\ air}^{k}\right)\right)\ [°C] \quad (14)$$

Where:

C_{rad}^{k} is the heat capacity of the radiator at the k-th time step (calculated based on the effective specific heat using the radiator's temperature at the *k*-th time step) $\left[\frac{J}{°C}\right]$

$\alpha_{rad,lev}^{k}$ is the heat transfer coefficient between the radiator surface and the indoor air at the *k*-th time step (which depends on the temperature conditions at the *k*-th time step) $\left[\frac{W}{m^{2}°C}\right]$

T_{rad}^{k} is the surface temperature of the radiator at the *k*-th time step [°C]

In the model, I assume that the radiator is switched off at the start of the simulation and the radiator valves are closed. Therefore, the heat is stored by the water within the radiator, the heat capacity of the radiator body, and—if present—the PCM inside the radiator.

In the case of a conventional radiator that does not contain phase change material, the heat capacity of the radiator is constant:

$$C_{conv.rad.} = m_{water} * c_{water} + m_{body} * c_{body} \left[\frac{J}{°C}\right] \quad (15)$$

Where the subscript 'water' refers to the water within the radiator, the subscript 'body' refers to the material of the radiator.

In the case where phase change material is integrated into the radiator body, we cannot assume a constant heat capacity. Applying the effective specific heat method, I have divided the specific heat of the PCM into three temperature-dependent stages; thus, the heat capacity in the model is temperature-dependent. The complete phase change of the material occurs between temperatures T_{oa} and T_{of}, within an interval of width $\Delta T = T_{of} - T_{o}$. The essence of the effective specific heat method is to treat the latent heat of fusion as an increased specific heat within the melting temperature range. A drawback of this method is that a sufficiently small time step must be chosen to ensure the melting process is not 'skipped' by the model. Therefore, in the model, the melting process takes place over a temperature change of ΔT.

Calculation of the PCM specific heat using the effective specific heat method as a function of temperature:

$$c_{eff\,PCM} = \begin{cases} c_{solid}\left[\frac{J}{kg°C}\right] & T < T_{oa} \\ \frac{c_{solid} + c_{liquid}}{2} + \frac{L_o}{\Delta T}\left[\frac{J}{kg°C}\right] & T_{oa} \le T \le T_{of} \\ c_{liquid}\left[\frac{J}{kg°C}\right] & T > T_{of} \end{cases}$$

In the case of a radiator containing phase change material, the heat capacity of the radiator based on the effective specific heat method is:

$$C(T) = m_{water} * c_{water} + m_{body} * c_{body} + m_{PCM} * c_{eff.PCM}(T) \left[\frac{J}{°C}\right] \quad (16)$$

Where:

$C(T)$ is the temperature-dependent heat capacity $\left[\frac{J}{°C}\right]$

$c_{eff.PCM}(T)$ is the temperature-dependent effective specific heat of the PCM $\left[\frac{J}{kg°C}\right]$

The subscript 'water' refers to the water within the radiator, the subscript 'body' refers to the material of the radiator (excluding PCM), and the subscript 'PCM' refers to the phase change material integrated into the radiator.

2.2. The cooling simulation

After determining the correlations, I initiated a simple simulation, which in this case consisted of populating the Excel spreadsheet I had developed.

The objective of the first simulation is to examine how the indoor air temperature evolves as a function of time when the building is left alone with a pre-heated but switched-off conventional radiator (without PCM). In the second simulation, I similarly investigated the time-dependent cooling process, but with a radiator containing PCM. The current model is still under development; the floor-on-ground and internal walls have not yet been included, so the model is not yet suitable for the complete modeling of real-world processes. Currently, the thermal storage mass consists only of the walls, the lower and upper slabs, the radiator, the water within it (and the PCM in the second case).

The 'room' under investigation is a space with internal dimensions of 5m x 5m and a ceiling height of 2.8m. It has no windows or doors, stands on stilts above the ground, and is in contact with the outdoor air on all sides. The slabs forming the floor and the ceiling of the room are 30 cm thick reinforced concrete. The external walls are made of small-sized solid ceramic bricks with a thickness of 38 cm. The internal air volume is 70 m³. The room is illustrated in Figure 1. The characteristics of the building envelopes are presented in Table 2. Geometric thermal bridges were not yet taken into account in this simulation, and there is no heat demand from infiltration. The surface thermal resistance values (the reciprocals of the heat transfer coefficients) are contained in Table 3.

Table 2. Characteristics of the building envelope of the simulated room (Source: MSZ 04-140-2)

Characteristic	Reinforced concrete slab	Small-sized solid ceramic brick masonry
Thickness [m]	0,3	0,38
Thermal conductivity $\left[\frac{W}{m°C}\right]$	1,55	0,78
Specific heat capacity $\left[\frac{J}{kg\,°C}\right]$	840	880
Density $\left[\frac{kg}{m^3}\right]$	2400	1730

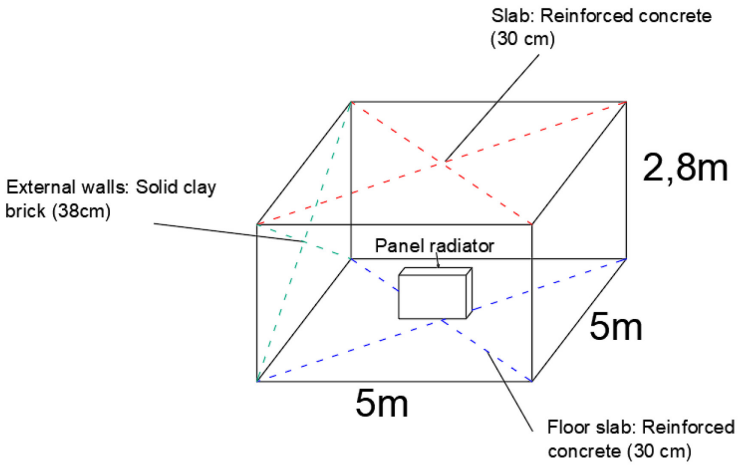


Figure 1. The "room" subject to simulation

Table 3. Surface thermal resistance values used in the simulation (Source: MSZ EN ISO 6946)

Name of structure	**External surface thermal resistance [m²K/W]**	**Internal surface thermal resistance [m²K/W]**
Ceiling slab	0,04	0,1
Floor slab	0,04	0,17
External wall	0,04	0,13

The energy flow required to maintain the room temperature in a steady-state (required heating load): Q=7774 W (excluding the effect of thermal bridges)

Initial and boundary conditions:

The outdoor air temperature throughout the simulation: -15°C.
The initial indoor air temperature: 20°C
At the start of the simulation, a steady-state temperature distribution corresponding to the -15°C outdoor and 20°C indoor air temperatures prevailed at every point of the building envelope.
The initial temperature was 70°C at every point of the radiator, including its surface.

Two versions were examined: in the first case, a radiator without PCM was considered, while in the second case, a thermally thin radiator equipped with 50 kg of PCM was taken into account during the simulation.

Based on the Purmo online selection tool, a Purmo Compact 33 (900mm x 2300mm) type radiator delivers a heat output of 7498 W at an inlet water temperature of 75°C, a return water temperature of 65°C, and an indoor air temperature of 20°C.

The logarithmic mean temperature difference (LMTD) based on the manufacturer's data:

$$\Delta T_{ln} = \frac{T_{supply} - T_{return}}{ln\frac{T_{supply} - T_{indoor\ air}}{T_{return} - T_{indoor\ air}}} = \frac{75°C - 65°C}{ln\frac{75°C - 20°C}{65°C - 20°C}} = 49{,}83\ °C \quad (17)$$

Note: The three-column radiators I have selected emit approximately 18% of their output via radiation (Baumann, 2010). For an accurate calculation of radiant heat transfer, the temperatures and emissivity factors of the surrounding surfaces must also be known. In the absence of these data, the accuracy of the calculation decreases. The thermal output values provided by the manufacturer include both radiant and convective heat transfer.

Determination of the product of the combined heat transfer coefficient and the radiator surface area:

$$Q_{nom.} = \alpha_{nom.} * A_{rad} * \Delta T_{nom.ln} \quad (18)$$

$$\alpha_{nom.} * A_{rad} = \frac{Q_{nom.}}{\Delta T_{nom.ln}} = \frac{7498W}{49{,}83\ °C} = 150{,}5\ \frac{W}{°C} \quad (19)$$

Note: Here, the combined heat transfer coefficient α accounts for both convective and radiant heat transfer. The subscript "nom." refers to the nominal state (as specified by the manufacturer).

Determination of the temperature dependence of the heat transfer coefficient:

$$A_{rad}\alpha_{rad.}^{k} = A_{rad}\alpha_{névl.} * \left(\frac{\Delta T_{excess}}{\Delta T_{nom.ln}}\right)^{n_{kit}} \quad (20)$$

In the simulation, we examine the room and radiator temperatures after the radiator is turned off. Since the radiator is treated as a thermally thin body in the model, all its points are at the same temperature at any given moment. Therefore, the difference between the radiator temperature and the indoor air temperature is substituted for the mean temperature excess.

$$A_{rad}\alpha_{rad.}^{k} = 150{,}5\,\frac{W}{^{\circ}C} * \left(\frac{T_{rad.surface}{}^{k} - T_{indoor\ air}^{k}}{49{,}83\ ^{\circ}C}\right)^{n_{kit}} \quad (21)$$

Where:

n_{kit}=0,3428 [-] (exponent derived from the Purmo radiator manufacturer catalog)

$\alpha_{rad.}^{k}$ is the value of the combined heat transfer coefficient between the radiator and the room air in the *k*-th time step $\left[\frac{W}{m^{2}{}^{\circ}C}\right]$

The mass of the selected steel panel radiator is 174 kg, and the mass of the water within is 30.6 kg. The mass of the PCM (Phase Change Material) used is 50 kg, specifically type Rubitherm RT54HC, which has a latent heat of fusion of 200,000 J/kg and a specific heat capacity of 2000 J/(kgK) in both solid and liquid states. The phase change occurs between 53°C and 54°C. Since the PCM-filled radiator was modeled as a thermally thin body, the thermal conductivity and density of the PCM were not utilized in the calculation. (This is a significant simplification that will be refined in future models.)

The structures were divided into 10 layers. The time step was set to Δt=10s, taking the stability criterion into account.(In this case, the stability criterion for the air yields a more restrictive time step than that calculated for the solid structure: $t_{min,air} = \frac{\rho_{air} * c_{air} * V_{air}}{\sum A_j \alpha_{in,j}}$=107 s > Δt=10s) Detailed calculations performed in Excel show that the PCM melting process can also be accurately tracked using the chosen Δt=10s time step.)

3. Results and discussion

The radiator containing PCM was able to maintain the indoor air temperature at 13.4 °C one hour after the heating was switched off, whereas, with the radiator without PCM, the room air temperature would have dropped to 9.1 °C.
In the model room, there were no internal walls, furniture, or equipment other than the radiator. The total thermal storage mass was provided by the external building envelopes, the radiator, and the indoor air. The room was 'elevated on stilts,' meaning it did not have a floor-on-ground construction.
The time evolution of the indoor air temperature is illustrated in Figures 2 and 3.

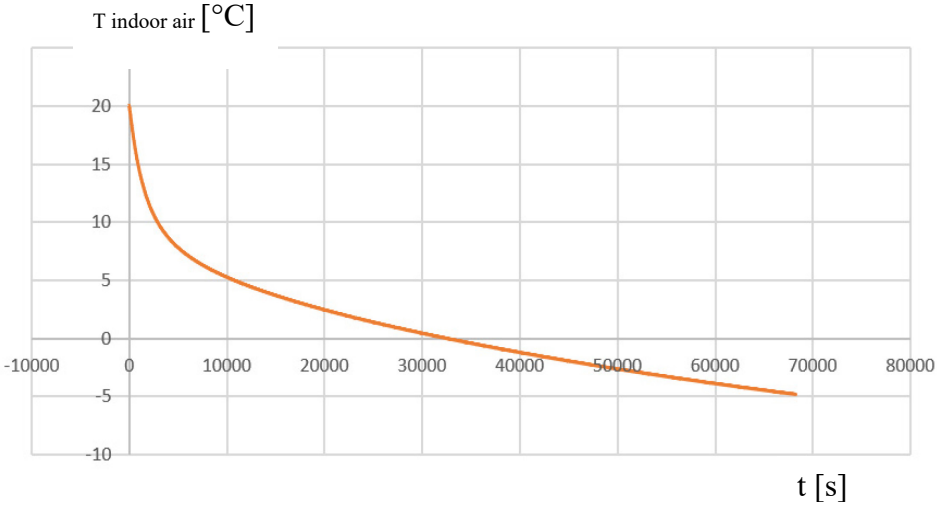


Figure 2. Evolution of the indoor air temperature for a radiator without phase change material

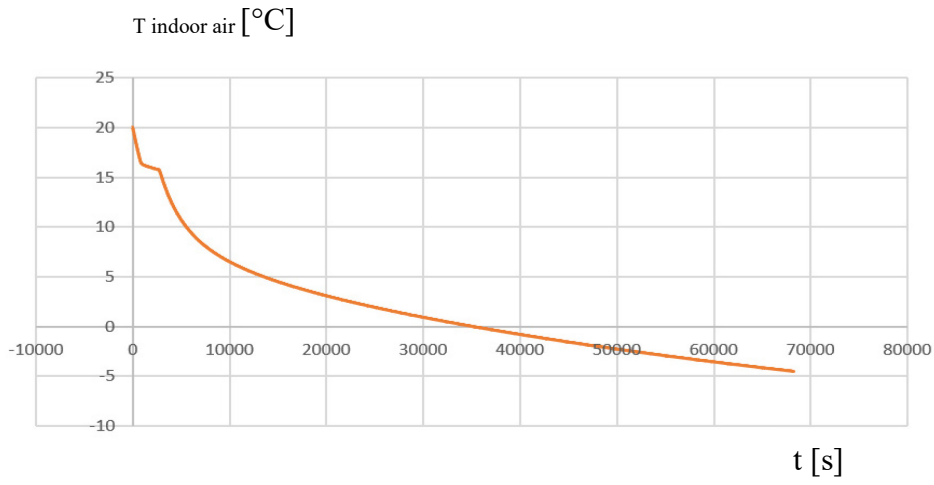


Figure3. Evolution of the indoor air temperature for a radiator containing phase change material

At the start of the simulation, the internal surface temperatures ranged between 5.3°C and 13.5°C, depending on the location of the surface. The reason for these low surface temperatures lies in the low thermal resistance of the masonry. After switching off the heating, the air—which has a small heat capacity—would quickly cool down to a value close to the surface temperature of the building envelopes in the absence of a heat source. Since the temperature difference between the air and the wall surfaces is large at the beginning of the process, higher heat fluxes flow from the air toward the walls. The cooling process slows down after the installation of the radiator, but it is still faster than the cooling rate typically observed in real buildings. The reason for this can be found in the differences between the model and actual buildings. In the case of real buildings, the envelopes typically have higher thermal resistance, and internal structures (such as internal walls and furnishings) also possess significant thermal storage mass; therefore, cooling is slower in a real building. My model did not include a floor-on-ground, nor any windows or doors. I did not take into account internal heat gains, the effects of thermal bridges, solar gains, or the heat demand from infiltration. The radiator was modeled thermally as a thin body.

My primary goal was to create a simple numerical model using the finite difference method that can be further developed in the future. Moving forward, I aim to create a model that more closely approximates reality by eliminating the aforementioned discrepancies.

6. Conclusion

Starting from the Fourier equation, it is possible to model simple, time-varying heat transfer processes—including phase changes—using the finite difference method.

The model I developed, which includes only the external building envelopes and a radiator, revealed that a radiator containing 50 kg of PCM was able to maintain the indoor air temperature at 13.4 °C one hour after the heating was turned off; in contrast, with a radiator without PCM, the room temperature would have dropped to 9.1 °C.

The model currently includes several simplifications; therefore, further development of the model is necessary to investigate the processes occurring in real buildings.

References

[1] Incropera F.P. and DeWitt D.P. (2007), Fundamentals of Heat and Mass Transfer, 6th Edition, John Wiley & Sons, New York. Wiley

[2] Carslaw H.S. and Jaeger J.C. (1959), Conduction of Heat in Solids, 2nd Edition, Clarendon Press, Oxford. Oxford University Press

[3] Al-Sanea, S. A. (2002). Thermal performance of building roof elements. *Building and Environment*, *37*(7), 665–675. https://doi.org/10.1016/s0360-1323(01)00077-4

[4] Ozel, M., & Pihtili, K. (2007). Optimum location and distribution of insulation layers on building walls with various orientations. *Building and Environment*, *42*(8), 3051–3059. https://doi.org/10.1016/j.buildenv.2006.07.025

[5] Mižáková, J., & Piteľ, J. (2017). An analytical dynamic model of heat transfer from the heating body to the heated room. *MATEC Web of Conferences*, *125*, 02047. https://doi.org/10.1051/matecconf/201712502047

[6] Jannat, N., Hussien, A., Abdullah, B., & Cotgrave, A. (2020). A comparative simulation study of the thermal performances of the building envelope wall materials in the Tropics. *Sustainability*, *12*(12), 4892. https://doi.org/10.3390/su12124892

7] Asan, H. (2006). Numerical computation of time lags and decrement factors for different building materials. *Building and Environment*, *41*(5), 615–620. https://doi.org/10.1016/j.buildenv.2005.02.020

[8] Fraisse, G., Viardot, C., Lafabrie, O., & Achard, G. (2002). Development of a simplified and accurate building model based on electrical analogy. *Energy and Buildings*, *34*(10), 1017–1031. https://doi.org/10.1016/s0378-7788(02)00019-1

[9] Baumann, M. (2010) Épületenergetika, Etudus Főiskola

[10] MSZ EN 442-2:2015 (Radiators and convectors — Part 2: Test methods and rating)

[11] MSZ-04-140-2:1991 Thermal calculations of buildings and building envelope structures: Thermal design

[12] MSZ EN ISO 6946:2017 (Building components and building elements — Thermal resistance and thermal transmittance — Calculation methods)

[13] Zeneli, M., Nikolopoulos, A., Karellas, S., & Nikolopoulos, N. (2021). Numerical methods for solid-liquid phase-change problems. *Ultra-High Temperature Thermal Energy Storage, Transfer and Conversion*, 165–199. https://doi.org/10.1016/b978-0-12-819955-8.00007-7

[14] Gao, Y., & Meng, X. (2023). A comprehensive review of integrating phase change materials in building bricks: Methods, performance and applications. *Journal of Energy Storage*, *62*, 106913. https://doi.org/10.1016/j.est.2023.106913

Materials and sensor technologies for condition monitoring and Total Productive Maintenance: a review with a case study on crop-protection equipment

Ebrahim Mirzaiee ASRAMI[1], Gábor KALÁCSKA[2] Miklós DARÓCZI[3]

[1] *Doctoral Program of Mechanical Engineering, Hungarian University of Agriculture and Life Sciences, Gödöllő*

[2] *Department of Materials Science and Engineering Processes, Institute of Technology, MATE, Gödöllő*

[3] *Department of Engineering Management, Institute of Technology, MATE, Gödöllő*

Abstract

Total Productive Maintenance (TPM) seeks to eliminate unplanned downtime and maximize Overall Equipment Effectiveness (OEE) by combining autonomous operator-level upkeep with continuous condition monitoring (CM) of the underlying equipment. While TPM is well established in discrete and process manufacturing, its translation to agricultural machinery, which operates in dusty, vibrating, and chemically aggressive environments, remains underexplored. This review pulls together the materials- and sensor-level enabling technologies that make CM and TPM viable for agricultural equipment, with particular focus on crop-protection machinery (sprayers, booms, nozzles, pumps) where agrochemical exposure, abrasive slurries, and field vibration drive accelerated wear and corrosion. We survey the principal classes of sensor materials, piezoelectric ceramics and polymers for vibration sensing, metal-oxide semiconductors and electrochemical materials for chemical sensing, optical fibres and two-dimensional materials for distributed and flexible monitoring, and nano-enabled platforms for residue detection, and map each class to the diagnostic and prognostic functions required by the eight pillars of TPM. We then examine how Internet-of-Things connectivity, edge artificial intelligence, and digital-twin frameworks integrate these sensor outputs into actionable maintenance decisions, including remaining-useful-life estimation and OEE-driven planning. A case study on crop-protection equipment illustrates how the combined materials–sensor stack is being deployed to detect nozzle wear, pump cavitation, agrochemical corrosion, and pesticide residue. Persistent barriers include sensor calibration drift in the field, materials incompatibility with aggressive agrochemicals, power and connectivity constraints, and the absence of TPM-specific standards for agricultural fleets. Future research directions point toward self-powered and self-healing sensor materials and federated-learning digital twins capable of fleet-scale predictive maintenance.

Keywords

Total Productive Maintenance; condition monitoring; sensor materials; precision agriculture; crop-protection equipment; digital twin; Overall Equipment Effectiveness; predictive maintenance.

Nomenclature

Abbreviation	Full term
Maintenance & management	
TPM	Total Productive Maintenance
CM	Condition monitoring
CBM	Condition-based maintenance
TBM	Time-based maintenance
PdM	Predictive maintenance
PHM	Prognostics and Health Management
RUL	Remaining useful life
OEE	Overall Equipment Effectiveness
TEEP	Total Effective Equipment Performance
PEE	Production Equipment Effectiveness
MTBF	Mean time between failures
MTTR	Mean time to repair
OSA-CBM	Open System Architecture for Condition-Based Maintenance
FMEA / FMECA	Failure Mode (and Effects / and Criticality) Analysis
OEM	Original equipment manufacturer
TRL	Technology readiness level
Agriculture & operations	
PA	Precision agriculture
UAV / UGV	Unmanned aerial / ground vehicle
IoT	Internet of Things
WSN	Wireless sensor network
LoRaWAN	Long-range wide-area network
NB-IoT	Narrowband IoT
GNSS / GPS / RTK	Global navigation satellite system / Global Positioning System / real-time kinematic
Sensor and device families	
MEMS	Micro-electro-mechanical systems

Abbreviation	**Full term**
AE	Acoustic emission
FBG	Fibre Bragg grating
FET	Field-effect transistor
RFID	Radio-frequency identification
RTD	Resistance temperature detector
IMU	Inertial measurement unit
HSI	Hyperspectral imaging
TIA	Transimpedance amplifier
ADC	Analogue-to-digital converter
Materials	
PZT	Lead zirconate titanate
AlN	Aluminium nitride
ZnO	Zinc oxide
PVDF	Poly(vinylidene fluoride)
P(VDF-TrFE)	Poly(vinylidene fluoride-co-trifluoroethylene)
MOS	Metal-oxide semiconductor
MOF	Metal-organic framework
MIP	Molecularly imprinted polymer
TMDC	Transition-metal dichalcogenide
2-D	Two-dimensional
VOC	Volatile organic compound
PTFE	Polytetrafluoroethylene
HDPE	High-density polyethylene
EPDM	Ethylene propylene diene monomer
FKM	Fluoroelastomer
AISI 316L	American Iron and Steel Institute austenitic stainless grade
SiC / GaN	Silicon carbide / gallium nitride (wide-band-gap)
Signal processing & AI	
FFT	Fast Fourier transform
RMS	Root mean square
DSP	Digital signal processing
AI / ML / DL	Artificial intelligence / machine learning / deep learning
CNN	Convolutional neural network
RNN	Recurrent neural network
LSTM	Long short-term memory
SVM	Support vector machine

Abbreviation	Full term
RF	Random forest
OEE accounting variables	
A	Availability factor of OEE
P	Performance factor of OEE
Q	Quality factor of OEE
Standards & methods	
ISO / IEC	International Organization for Standardization / International Electrotechnical Commission
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
API	Active pharmaceutical ingredient (§3.2 context)

1. Introduction

1.1 Agriculture 4.0 and the equipment-uptime imperative

The transition from conventional to digitized farming, variously referred to as Precision Agriculture, Smart Farming, or Agriculture 4.0, has accelerated over the past decade through the convergence of Internet of Things (IoT) connectivity, low-cost sensor networks, machine learning, and unmanned aerial and ground platforms [1–4]. In this paradigm, soil, crop, and atmospheric variables are continuously measured at field resolution, and operational decisions are taken with site-specific granularity [5–7]. Consequent gains in yield, input efficiency, and environmental footprint are now well documented [3, 7, 8], and have shifted agriculture from a labour-bound activity toward a capital- and equipment-bound one. The resulting dependence on a growing fleet of high-value assets, such as autonomous tractors, robotic harvesters, variable-rate sprayers, irrigation controllers, greenhouse climate systems, places equipment availability and reliability at the centre of farm productivity. Unplanned downtime during the narrow windows of planting, spraying, or harvest can be irrecoverable within a season [9].

1.2 From corrective to total productive maintenance

The maintenance literature has progressively shifted from corrective (run-to-failure) and time-based preventive strategies toward Condition-Based Maintenance (CBM) and, more recently, predictive maintenance underpinned by data-driven prognostics [10, 11]. Total Productive Maintenance (TPM), introduced in Japanese manufacturing in the 1970s and codified by Nakajima, builds on these foundations by binding maintenance into the wider production system through eight organizational pillars: autonomous maintenance, planned

maintenance, focused improvement, early-equipment management, quality maintenance, training, office TPM, and safety/health/environment [12]. Its headline metric, Overall Equipment Effectiveness (OEE), aggregates availability, performance, and quality losses into a single productivity score and has become the dominant industrial yardstick for maintenance effectiveness [13]. Recent case studies report substantial OEE gains and downtime reductions when TPM is implemented in oil and gas, pharmaceutical, and discrete-manufacturing settings [12, 14, 15]. The integration of TPM with Industry 4.0, sometimes termed TPM 4.0 or Maintenance 4.0, extends these gains by replacing periodic inspection with continuous, sensor-driven CM and by closing the loop between machine condition data and maintenance scheduling [15–17].

1.3 Sensors and materials as the enabling layer

The viability of CBM, predictive maintenance, and TPM 4.0 depends on the ability to instrument equipment continuously, robustly, and economically. This is a materials problem as much as a signal-processing one: the sensitivity, selectivity, durability, and form-factor of a sensor are set by the active material at its core. Piezoelectric ceramics (PZT) and polymers (PVDF) underpin vibration and strain measurement for rotating machinery [18]; metal-oxide semiconductors and chemiresistive layers transduce gas and volatile-organic-compound concentrations [19, 20]; optical fibres and two-dimensional materials (graphene, MXenes) enable distributed and flexible sensing on curved or moving parts [21, 22]; nano-engineered electrochemical interfaces detect residues at trace concentrations [20, 23]. Choices among these materials govern whether a sensor can survive on a vibrating tractor chassis, inside a pesticide-laden pump housing, or on a flexible boom subject to thermal and mechanical fatigue [24–26].

1.4 Why agricultural equipment is a hard case

Agricultural service environments stress sensors and equipment alike. Tractors and self-propelled sprayers vibrate broadband across the audible range and dissipate heat under variable load; spray booms flex and oscillate; pumps, valves, and nozzles handle abrasive, often corrosive, fluids; greenhouses and silos host condensation, ammonia, and biogenic compounds [26]. Agrochemicals, pesticides, herbicides, and liquid fertilizers accelerate corrosion of metallic components and degrade elastomeric seals [27, 28]. Field connectivity is intermittent, primary power is constrained to vehicle electrical systems or harvested ambient energy, and the operator population is heterogeneous in technical training. The same sensor materials that perform reliably in a climate-controlled factory cell often drift, foul, or fail outright in this setting. So the maintenance literature developed for industrial machinery cannot be transposed to agriculture without re-examining the underlying materials–sensor–environment interactions [26, 29].

1.5 Crop protection as a representative case

Among agricultural subsystems, crop-protection equipment, boom sprayers, orchard sprayers, UAV applicators, and their associated pumps, nozzles, and dosing controls concentrate several of the harshest service conditions. Nozzle orifices wear under abrasive suspension flow, distorting droplet size and application rate; pumps suffer cavitation and seal degradation; metallic frames corrode under repeated agrochemical wash; electronic dosing units must operate beside high-voltage spray circuits and within reach of chemical drift [30–34]. Crop-protection operations also have an outsized economic and environmental footprint: a single under-spray event compromises yield, while over-spray inflates input cost and contaminates soil and groundwater [30, 31]. The CM challenge here is multimodal vibration and acoustic sensing for rotating components, electrochemical or optical sensing for residue and concentration, strain sensing for boom dynamics, and visual or hyperspectral sensing for spray-pattern validation [32, 33, 35]. Crop protection, therefore, provides a uniquely rich setting in which to examine how sensor materials and CM strategies couple to TPM outcomes.

1.6 Scope, contributions, and structure

This review pursues three contributions. First, it pulls together the materials science of sensors used in equipment CM into a single, application-oriented taxonomy, with a focus on durability and chemical compatibility for agricultural service. Second, it bridges the otherwise-separate literatures on TPM, CBM, predictive maintenance, and digital twins by mapping their respective data and decision requirements onto sensor-material capabilities. Third, it grounds the synthesis in a structured case study of crop-protection equipment, where it identifies the highest-impact CM measurements and the most pressing materials gaps. The remainder of the paper is organized as follows. Section 2 presents the review methodology and the bibliometric basis of the corpus. Section 3 develops the conceptual background of TPM, OEE, and CM. Section 4 surveys the principal classes of sensor materials. Section 5 reviews the corresponding sensor technologies and their reported applications. Section 6 examines the data, AI, and digital-twin layer that connects CM outputs to maintenance decisions. Section 7 develops the crop-protection case study. Section 8 discusses cross-cutting challenges, and Section 9 outlines future research directions before concluding.

2 .Review methodology

2.1 Scope and research questions

The review addresses three research questions: **(RQ1)** Which classes of sensor materials are most relevant to the condition monitoring of agricultural

equipment, and what are their durability and selectivity trade-offs under field conditions? **(RQ2)** How do the resulting sensor outputs flow through IoT, AI, and digital-twin layers into the eight pillars of Total Productive Maintenance, and what does the resulting CM-to-TPM pipeline look like in agriculture? **(RQ3)** What does this integrated picture imply specifically for crop-protection equipment sprayers, booms, nozzles, pumps, and UAV applicators, and what materials and sensor gaps remain?

2.2 Information sources and search strings

A structured search was conducted across **Scopus**, **Web of Science**, **IEEE Xplore**, **ScienceDirect**, and the publisher portals of **MDPI**, **Springer Nature**, **Wiley**, and **Frontiers**. Open-access aggregators (**CORE.ac.uk**, **Unpaywall**, **Europe PMC**) were used to retrieve full text where direct access was restricted. The time window was set at **2003–2026**, with a deliberate concentration on 2018–2025 publications to capture the post-Industry 4.0 transition; a small number of earlier foundational works on condition-based maintenance, OEE, and silicon-detector radiation damage were retained as anchor references.
The search vocabulary was organized into four orthogonal axes that were combined with Boolean AND:

Axis	Representative terms
Maintenance & reliability	*"total productive maintenance"*, *"condition-based maintenance"*, *"predictive maintenance"*, *"prognostics and health management"*, *"overall equipment effectiveness"*, *"remaining useful life"*, *"digital twin"*
Sensors & materials	*"piezoelectric"*, *"metal oxide"*, *"electrochemical sensor"*, *"chemiresistive"*, *"optical fiber sensor"*, *"two-dimensional material"*, *"nanomaterial"*, *"MEMS accelerometer"*, *"chipless RFID"*
Agricultural context	*"precision agriculture"*, *"agriculture 4.0"*, *"smart farming"*, *"agricultural machinery"*, *"tractor"*, *"greenhouse"*, *"variable-rate"*
Crop-protection case	*"sprayer"*, *"nozzle wear"*, *"pesticide"*, *"glyphosate"*, *"agrochemical corrosion"*, *"UAV pesticide application"*, *"crop protection"*

2.3 Inclusion and exclusion criteria

Records were retained when they (i) were peer-reviewed journal articles, conference proceedings, or authoritative grey literature (extension-service or national-laboratory technical reports); (ii) addressed at least one of the four search axes substantively rather than tangentially; and (iii) were available in English. Records were excluded when they were duplicates, lacked methodological transparency, addressed unrelated industrial sectors without

transferable insight, or were retracted. Highly cited foundational works were retained even when older than the primary time window.

2.4 PRISMA-style flow

A structured screening process was applied to assemble the review corpus. Records were identified from five sources: Scopus, Web of Science, IEEE Xplore, ScienceDirect, and the publisher portals of MDPI, Frontiers, and Springer, yielding an initial pool of approximately 610 unique records after de-duplication. Title-and-abstract screening removed 482 records, leaving 128 for full-text retrieval (of which 5 could not be recovered). Of the 123 records assessed for eligibility, 72 were excluded for the following reasons: off-scope industry (41), methodology not transparent (18), non-English (7), and superseded by a newer review (6). The remaining studies constitute the final review corpus, predominantly peer-reviewed journal articles, complemented by a small number of conference and book-chapter contributions and two grey-literature technical reports on agricultural corrosion.

2.5 Bibliometric snapshot of the corpus

The corpus is concentrated in the period **2020–2025 (≈82 %)**, with foundational anchors from 2003, 2006, and 2017. By topic, **sensor materials and devices** account for the largest fraction (≈40 %), followed by **smart-farming IoT/AI** (≈25 %), **sensors deployed in agricultural environments** (≈18 %), **maintenance and condition-monitoring theory** (≈10 %), and **agrochemical corrosion and crop-protection-equipment materials** (≈7 %). The journal distribution is dominated by *Sensors*, *Applied Sciences*, *Agronomy*, *Processes*, and *Computers and Electronics in Agriculture*. The geographic distribution of corresponding authors shows a marked concentration in **East Asia, the European Union, and North America**, with emerging contributions from South Asia and Sub-Saharan Africa on crop-protection applications. This distribution is consistent with the global pattern of digital-agriculture R&D investment and confirms that the materials/sensor and the agricultural-maintenance communities still publish in largely disjoint venues, a structural gap that this review seeks to bridge.

3. Background: TPM, OEE, condition.based and predictive maintenance

3.1 Evolution of maintenance philosophies

Industrial maintenance practice has progressed through four broad stages, each motivated by the inadequacy of the previous one. **Corrective (run-to-failure)** maintenance accepts unplanned downtime as the operational mode, and is rational only when failures are inexpensive, equipment is redundant, or asset cost is dominated by capital rather than availability. **Time-based preventive**

maintenance (TBM) schedules service at fixed intervals derived from manufacturer recommendations and reliability statistics; it suppresses catastrophic failures but performs interventions on components that have not yet degraded, inflating both labour cost and the risk of maintenance-induced failure [10, 11]. **Condition-based maintenance (CBM)** triggers service from observable health indicators rather than calendar time, and so captures the productive life of components that age slowly while still catching outliers that age rapidly [10]. **Predictive maintenance** extends CBM by adding a forecasting layer: prognostic models project the trajectory of a health indicator forward in time to estimate remaining useful life (RUL) and to optimize the scheduling decision against operational constraints [17, 36, 37]. The empirical case for moving along this trajectory is well documented: Molęda et al. [17] report 20–40 % reductions in unplanned downtime in the power industry when shifting from corrective to predictive, and Werbińska-Wojciechowska and Winiarska [16] confirm comparable patterns across the Industry 4.0 maintenance literature.

3.2 Total Productive Maintenance and the eight pillars

Total Productive Maintenance (TPM) is an organizational complement, rather than a replacement, to CBM and predictive maintenance. Introduced in Japanese manufacturing and codified through Nakajima's eight-pillar model, TPM frames maintenance as a *shared production responsibility* rather than a separate craft. Its pillars are typically rendered as: (i) **autonomous maintenance**, transferring routine inspection, cleaning, and lubrication to operators; (ii) **planned maintenance**, executed by maintenance specialists on a CBM- or predictive-driven schedule; (iii) **focused improvement** (kaizen), targeting recurrent loss modes; (iv) **early-equipment management**, feeding operational learning back into procurement and design; (v) **quality maintenance**, binding maintenance to product-quality outcomes; (vi) **education and training**; (vii) **office TPM**, extending the philosophy to administrative processes; and (viii) **safety, health, and environment** [12, 14, 15]. The eight pillars are interdependent: autonomous maintenance reduces sensor load by enabling operator-level visual checks, planned maintenance becomes tractable only when CBM data identify true intervention needs, and early-equipment management is possible only when both autonomous and planned activities generate structured failure records.

3.3 Overall Equipment Effectiveness (OEE)

The headline performance metric of TPM is **Overall Equipment Effectiveness (OEE)**, defined as the product of three factor scores [13]:

$$OEE = A \times P \times Q$$

where **availability** A is the ratio of actual operating time to planned production time, **performance** P is the ratio of actual to ideal throughput during operating

time, and **quality** Q is the fraction of output that meets specification on first pass. World-class manufacturing benchmarks place "good" OEE near 85 % (typically A≈0.90, P≈0.95, Q≈0.999); typical baselines on legacy equipment hover around 40–60 % [13]. Ng Corrales et al. [13] catalogue more than a dozen variants and extensions of OEE, including Total Effective Equipment Performance (TEEP), Production Equipment Effectiveness (PEE), and OEE-based composites for transportation and mining, that adapt the basic formula to sector-specific loss structures. For this review, OEE is also the bridge between sensor-derived CM data and TPM decision-making: availability losses are quantified directly from downtime telemetry, performance losses from cycle-time and throughput signals, and quality losses from product-side measurement. Each of the three factors therefore, demands its own sensor pathway.

3.4 CBM, predictive maintenance, and the OSA-CBM stack

CBM systems are conventionally decomposed into the seven-layer **Open System Architecture for Condition-Based Maintenance (OSA-CBM)** stack, which separates data acquisition, data manipulation, state detection, health assessment, prognostic assessment, decision support, and presentation [10]. Each layer maps onto specific algorithmic and sensor choices: data acquisition is set by transducer physics (Section 4) and signal-conditioning electronics; state detection traditionally relies on time- and frequency-domain features (RMS, kurtosis, spectral peaks, envelope demodulation); health assessment compares those features against baselines or thresholds; prognostic assessment fits trajectory models Wiener, Gamma, Weibull, or learned recurrent neural networks to project RUL [36]. Park et al. [37] illustrate a complete OSA-CBM pipeline for cryogenic safety valves in which a deep-learning model trained on pressure and acoustic signals delivers RUL estimates with a mean absolute error of a few operating cycles. Equivalent pipelines for centrifugal pumps, rotating machinery in industrial robots, and rolling-element bearings are reviewed by Chen et al. [36] and synthesise the methodological consensus: a hybrid combination of physics-informed features and data-driven sequence models currently outperforms either approach alone.

3.5 How TPM, CBM, and predictive maintenance interact

It helps to be clear about the relationship between these three concepts because they are sometimes incorrectly treated as competitors. **CBM is a sensing-and-decision discipline**; it specifies *when* intervention is justified given measured conditions. **Predictive maintenance is a forecasting layer on top of CBM**; it specifies *how long the present condition can be safely tolerated*. **TPM is an organizational frame**; it specifies *who acts on those signals, with what skills, and against what production objectives* [12, 15]. The recent literature on *TPM 4.0* / *Maintenance 4.0* converges on the position that the three concepts are most powerful when stacked: CBM and predictive maintenance furnish the signals;

TPM furnishes the operational structure that converts those signals into OEE gains [15, 16, 17]. The remainder of this review is organized along that stack, first the materials and sensors that produce the signals (Sections 4 and 5), then the data and decision layer that processes them (Section 6), and finally the application case study in which the stack is exercised end-to-end (Section 7).

4. Materials for condition-monitoring sensors

The sensor materials used in condition monitoring divide naturally into eight classes, defined by transduction mechanism and form factor. Each class has a distinct performance envelope sensitivity, selectivity, bandwidth, temperature tolerance, chemical compatibility, fabrication cost, and durability, and therefore a distinct niche in the overall CM stack. This section surveys the eight classes with explicit attention to the durability and chemical-compatibility constraints that govern agricultural deployment.

4.1 Piezoelectric ceramics, single crystals, and polymers

Piezoelectric materials remain the workhorse of vibration and strain sensing in industrial CM. **Lead zirconate titanate (PZT)** offers piezoelectric coefficients in the range d33≈200–600 pC/N and operating temperatures to several hundred °C, making it the default choice for hard-mounted accelerometers and acoustic-emission sensors on rotating machinery [18]. Lead-free alternatives, **aluminium nitride (AlN)** and **zinc oxide (ZnO)** sacrifice an order of magnitude in coefficient but enable CMOS-compatible thin-film integration and are increasingly competitive in MEMS accelerometers [18]. On the polymer side, **poly(vinylidene fluoride) (PVDF)** and its trifluoroethylene copolymer P(VDF-TrFE) deliver lower coefficients still but exceptional mechanical flexibility, chemical inertness against most agrochemicals, and the ability to be solution-cast onto curved structural surfaces. Chen et al. [18] review the rapid recent progress in flexible piezoelectric materials for strain sensing on wearable and machine-mounted surfaces, with energy-autonomous self-powered architectures combining vibration sensing with energy harvesting on the same element.
Figure 3 summarises the principal classes of flexible and wearable piezoelectric sensor materials PVDF, P(VDF-TrFE), porous and hierarchically structured composites, and the application domains (health monitoring, human–machine interaction, structural sensing) that motivate the materials-selection trade-offs discussed above.

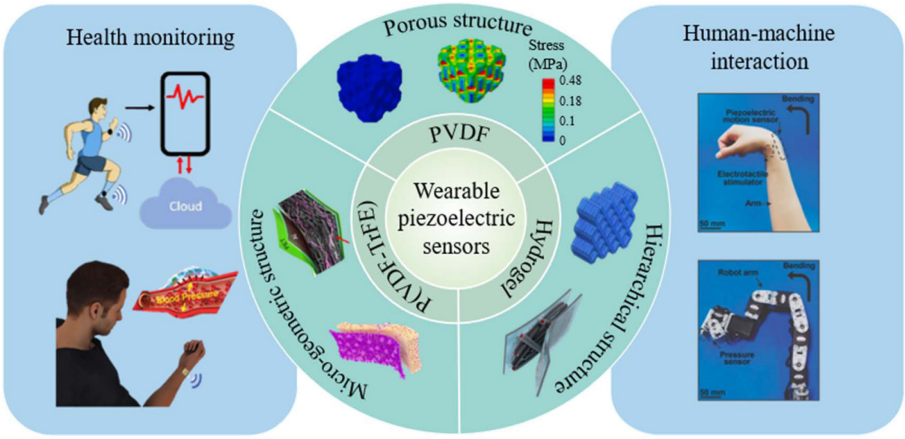


Figure 3. Overview of flexible/wearable piezoelectric sensor materials and their principal application domains. Reproduced from Ref. [18].

4.2 Metal-oxide semiconductor (MOS) materials

Metal-oxide sensors transduce the partial pressure of reactive gases into resistance changes by surface-mediated charge transfer. **SnO_2, ZnO, WO_3, and In_2O_3** dominate the literature, with selectivity tuned by dopant choice (Pd, Pt, Au, Ag), morphology (nanowires, nanosheets, hollow spheres), and operating temperature. Marikutsa et al. [19] provide a comprehensive synthesis of how the chemistry of active surface sites, Lewis acid sites, oxygen vacancies, and metal-cluster catalysts governs both selectivity and long-term stability under realistic gas mixtures. In CM applications, MOS sensors detect lubricant degradation by-products, motor-insulation outgassing, and ammonia in livestock and greenhouse environments. The principal limitations remain elevated operating power (heated substrate, typically 150–400 °C) and humidity cross-sensitivity, both of which constrain deployment on battery-powered agricultural assets.

Figure 4 illustrates why selectivity in metal-oxide sensors is a materials-chemistry problem: the surface charge-transfer reactions for an oxidising analyte (NO_2) and a reducing analyte (CO) produce opposite band-bending at the oxide surface, and the magnitude of that response is set by the density and type of active surface sites.

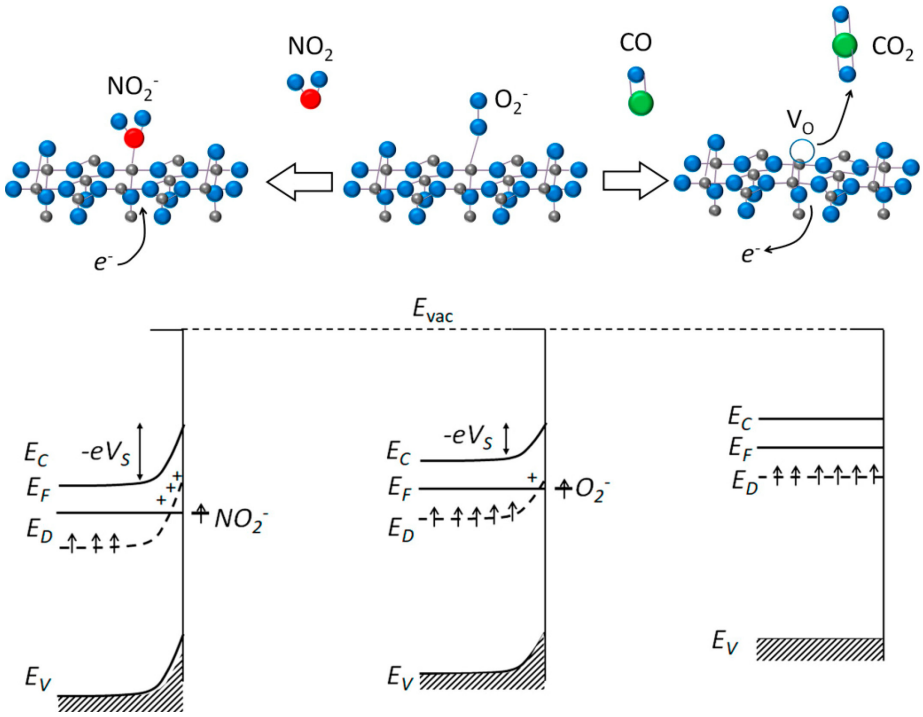


Figure 4. Surface gas-sensing mechanism of metal-oxide semiconductors and the associated energy-band changes for oxidising (NO_2) versus reducing (CO) analytes. Reproduced from Ref. [19].

4.3 Electrochemical and chemiresistive materials

Electrochemical sensors couple analyte-specific redox reactions to current, potential, or impedance signals at functionalized electrodes; chemiresistive sensors transduce analyte adsorption directly into resistance changes in an active film. Baba et al. [20] review recent advances in chemiresistive and electrochemical sensing materials for volatile organic compounds, emphasising the role of **conducting-polymer composites** (PEDOT: PSS, polyaniline, polypyrrole), **graphene and graphene-oxide derivatives**, and **metal-organic frameworks (MOFs)** in achieving sub-ppm detection limits with room-temperature operation. For agricultural CM, electrochemical platforms additionally enable in-line detection of pesticide residues at the sprayer outlet [32, 33] and continuous monitoring of soil-solution chemistry [38]. Stability under repeated wet/dry cycling and biofouling resistance remain active research areas.

4.4 Optical-fibre and photonic materials

Silica and polymer optical fibres function as both a communication medium and a distributed sensing element. Fibre Bragg gratings (FBGs) transduce strain and temperature into wavelength shifts; intrinsic Mach–Zehnder and Fabry–Pérot interferometers, evanescent-wave configurations, and surface-plasmon-resonance overlays extend the menu to vibration, pressure, refractive index, and analyte-specific chemical sensing. Jha et al. [22] review the recent surge in optical-fibre-based wearable sensors, with particular attention to polymer fibres that tolerate sharper bend radii and chemical exposure. For long agricultural booms and irrigation pipelines, distributed acoustic and strain sensing along a single fibre provides spatial resolution unmatched by point-sensor arrays at the cost of more demanding interrogator electronics.

4.5 Two-dimensional materials

Two-dimensional materials **graphene**, **transition-metal dichalcogenides (MoS_2, WSe_2)**, **MXenes (e.g. $Ti_3C_2T_x$)**, and **hexagonal boron nitride,** combine an atomically thin form factor with high carrier mobility, large surface area, and tunable surface chemistry. Sun et al. [21] review their use in wearable biodevices and flexible-electronic sensor platforms, where they enable strain, electrochemical, and photonic sensing on conformable substrates. For CM applications, the principal opportunities are (i) MXene-based flexible strain sensors for distributed monitoring of booms and frames, (ii) graphene-derivative chemiresistors for residue and gas detection, and (iii) heterostructure photodetectors for hyperspectral imaging payloads.

4.6 Flexible and wearable substrates

Polymer substrates, **polyimide (PI)**, **polydimethylsiloxane (PDMS)**, and **polyethylene terephthalate (PET)** host the active sensing layers of the preceding classes and largely determine system-level durability under bending, thermal cycling, and chemical exposure. Polyimide combines high glass-transition temperature with excellent solvent resistance and is the default choice for sensors that must survive engine-bay or pump-housing environments. PDMS offers extreme conformability for shape-irregular surfaces but is permeable to many agrochemical solvents and therefore requires barrier coatings in spray-side deployment. PET dominates large-area, low-temperature applications such as full-surface strain mapping on tank exteriors [18, 21, 22].

4.7 Nanomaterials and nano-engineered surfaces

The agricultural sensing literature has invested heavily in **nano-enabled platforms**, both for crop and soil sensing and for residue detection on equipment. Ali et al. [39], Wu and Li [40], and Nguyen et al. [41] survey nanobiotechnological advances in agriculture, ranging from nano-pesticide

delivery to nano-enabled biosensors for crop-health monitoring and pesticide residue detection. The CM-relevant subset of this literature concerns (i) noble-metal nanoparticle decorations on MOS and chemiresistive surfaces to boost catalytic activity and selectivity; (ii) molecularly imprinted polymer (MIP) layers for analyte-specific electrochemical detection; and (iii) nano-textured antifouling coatings that extend sensor life under wet, chemically aggressive service. Nanotoxicity and end-of-life considerations [39] limit deployment in food-contact and water-discharge applications and must be evaluated for each formulation.

4.8 Radiation-tolerant and harsh-environment materials

A distinct strand of materials science, driven historically by high-energy physics, addresses sensors that must operate under intense radiation or extreme temperature. Lindström's foundational review [42] of radiation damage in silicon detectors quantifies the bulk-defect kinetics that limit the operational life of silicon-based imagers under field-equivalent UV and ionising exposure, and motivates the migration toward wide-band-gap materials (4H-SiC, GaN) in photodetectors that must survive direct sunlight, agrochemical drift, and outdoor weathering over multi-year service. While radiation tolerance is rarely a real limit in agricultural CM, the same materials-science framework applies to long-term UV and chemical aging of sensor packaging.

4.9 Comparative summary

Table 1 summarizes the eight classes against the criteria most relevant to agricultural CM: target signal, sensitivity range, typical operating temperature, chemical compatibility envelope, mechanical form factor, fabrication cost, and reported technology-readiness level on agricultural assets. The headline conclusion is that no single material class dominates: vibration and strain monitoring favour piezoelectric ceramics and 2-D-material flexible films; chemical and residue sensing favour electrochemical, chemiresistive, and MOS platforms; distributed structural monitoring favours optical fibres. The integration challenge addressed in Section 6 is to combine multiple sensor families into a coherent CM stack whose outputs can be fused into the OEE accounting framework of Section 3.

Table 1. Comparative summary of the principal sensor-material classes against agricultural-CM criteria. Operating-temperature windows are quoted for continuous service; transient excursions can extend the envelope. Chemical compatibility uses a qualitative four-level scale: ◉ excellent, ● good, ◐ moderate, ○ poor. Technology-readiness levels (TRL) refer to reported field deployment on agricultural equipment specifically, not laboratory or industrial deployment.

#	Material class	Principal transduced signal	Operating T (continuous)	Acidic / alkaline / organic-solvent compatibility	Mechanical form factor	Fabrication cost	Field TRL on ag equipment
M1	PZT piezoelectric ceramic	Vibration, acoustic emission, strain	-50 to +250 °C	◐ / ◐ / ● (housing-dependent)	Bulk stack, disc, hard-mounted	Low (commodity)	7–8
M2	AlN / ZnO piezoelectric thin film (MEMS)	Vibration, ultrasonic, strain	-40 to +150 °C	● / ● / ● (encapsulated)	MEMS die, surface-mount	Low at volume	6–7
M3	PVDF / P(VDF-TrFE) piezoelectric polymer	Strain, low-frequency vibration, large-area pressure	-40 to +90 °C	◉ / ● / ◐	Thin film, conformable	Low	4–6
M4	Metal-oxide semiconductor (SnO_2, ZnO, WO_3)	Gas / VOC partial pressure	Sensor operates at 150–400 °C (heated)	◐ / ◐ / ◐ (humidity-sensitive)	Heated chip, MEMS or ceramic	Low–medium	5–6
M5	Conducting-polymer / chemiresistive (PEDOT, PANI, polypyrrole)	VOC, biomolecule, residue	-20 to +80 °C	◐ / ◐ / ○	Solution-cast film, flexible	Low	3–5
M6	Graphene / graphene-oxide derivatives	Strain, gas, electrochemical residue	-40 to +200 °C (encapsulated)	● / ● / ●	Thin film, transferable	Medium	3–5
M7	MXene / TMDC 2-D materials	Strain, electrochemical, photonic	-40 to +150 °C	● / ◐ / ●	Thin film, flexible	Medium–high	2–4
M8	Silica optical fibre (FBG, interferometric)	Distributed strain, temperature, pressure, refractive index	-50 to +300 °C	◉ / ● / ◉	Long-line distributed	Medium (fibre cheap, interrogator costly)	5–6
M9	Polymer optical fibre	Strain, refractive index, chemical	-30 to +80 °C	● / ● / ◐	Distributed, flexible	Low	3–5
M10	Functionalised electrochemical electrodes (Au, Pt, glassy C, MIP, enzyme)	Specific analyte (pesticide, glyphosate, ion)	-20 to +60 °C	● / ● / ● (analyte-specific)	Screen-printed disposable or fixed probe	Low (screen-printed)	4–6
M11	Field-effect-transistor biosensor (Si, organic, 2-D)	Specific biomolecule / residue	-20 to +80 °C	● / ● / ●	Microfabricated chip	Medium–high	3–5
M12	Nano-enabled platforms (Au/Ag NP, MOF, MIP layers)	Enhanced selectivity for the above	Class-dependent	● / ● / ● (formulation-specific)	Decoration on M4–M11	Medium	3–5
M13	Chipless RFID	Strain, T,	-40 to +125	● / ● / ●	Flat tag,	Very low	3–4

#	Material class	Principal transduced signal	Operating T (continuous)	Acidic / alkaline / organic-solvent compatibility	Mechanical form factor	Fabrication cost	Field TRL on ag equipment
	passive resonators	moisture, corrosion	°C	(substrate-dependent)	battery-free		
M14	Wide-band-gap photodetectors (SiC, GaN)	UV / visible / IR	-40 to +250 °C	◉ / ◉ / ◉	Discrete die	High	4–6

The table makes the integration argument explicit: vibration and strain at the asset level are best served by M1–M3, M6–M9, M13; gas and VOC monitoring by M4–M6; pesticide-residue and lubricant-chemistry sensing by M5, M10, M11, M12; structural distributed sensing by M8, M9; UV / weathering-tolerant imaging by M14. The harsh-environment penalty falls hardest on the chemiresistive and conducting-polymer families and lightest on optical-fibre and chipless-RFID platforms; this hierarchy should govern materials-selection decisions in agricultural CM design.

5. Sensors for condition monitoring

This section moves from the material classes of Section 4 to the integrated sensor devices that implement them, and to the measurements those devices contribute to a CM-to-TPM pipeline. The taxonomy below follows the dominant signal modality.

5.1 Vibration sensors

Vibration sensing is the single most widely used CM modality, justified by the fact that the majority of mechanical failure modes, bearing wear, gear-tooth pitting, shaft imbalance, misalignment, looseness, and cavitation broadcast distinctive signatures in the acceleration spectrum [24]. The dominant device families are (i) **piezoelectric accelerometers** based on PZT or single-crystal shear or compression modes, offering broad bandwidth (up to >10 kHz), wide dynamic range, and survival to ≥10 kHz vibration without internal resonance; (ii) **piezoelectric MEMS accelerometers** based on AlN or PZT thin films, which combine moderate sensitivity with low noise floor and CMOS integration; and (iii) **capacitive MEMS accelerometers**, which dominate the low-cost, low-frequency end of the market and are increasingly competitive at moderate bandwidth as wireless IoT installations proliferate [18, 24]. Hassan et al. [24] document the methodological progression in feature extraction from envelope demodulation and order tracking to wavelet-packet decomposition and convolutional neural network classifiers and emphasize that the choice of sensor

family must match the analysis bandwidth: detecting an incipient bearing fault below 100 Hz needs different transduction physics than detecting cavitation broadband above 10 kHz.

5.2 Acoustic-emission sensors

Acoustic-emission (AE) sensing extends the bandwidth above ~100 kHz, into the regime where transient elastic waves from crack initiation, plastic deformation, and seal failure radiate before they perturb the macroscopic vibration signal. AE sensors are almost exclusively PZT-based, with resonant or broadband-flat designs depending on the target frequency. In agricultural CM, AE sensing is principally relevant to early detection of pump-impeller pitting, seal degradation in cavitating service, and incipient gear failure in transmissions [18, 24]. The principal limitation is the very high data rate (sample rates of 1–10 MHz are typical), which constrains continuous streaming and motivates on-sensor event-driven processing.

5.3 Temperature, pressure, and strain sensors

Thermocouples and resistance temperature detectors (RTDs) remain the dominant temperature sensors at the component level; integrated bandgap or thermistor circuits dominate at the electronics level. Pressure sensing on hydraulic and pneumatic systems uses **silicon piezoresistive** or **capacitive MEMS diaphragm** transducers; on aggressive media, the diaphragm is isolated by a stainless or Hastelloy capsule with silicone oil fill. Strain measurement on structural elements has historically used foil gauges and increasingly uses fibre Bragg gratings [22] and 2-D-material flexible films [21] for distributed and large-area coverage. For agricultural booms and frames, the most informative deployments are long-baseline strain sensing along the structure rather than point measurements at presumed-critical sections [26].

5.4 Chipless RFID and passive wireless sensors

A distinct strand of CM sensor research targets battery-free, low-cost devices that can be densely distributed and interrogated occasionally rather than streamed continuously. Bilic et al. [25] survey chipless RFID sensors that encode physical state (strain, temperature, moisture, corrosion progression) into the electromagnetic signature of a passive resonator. The key advantage for agricultural deployment is the absence of an onboard battery or wired power, which removes the dominant failure mode of distributed CM systems in the field. Read range is currently limited to ~1 m for most published designs, which suits boom-mounted or pump-mounted tags interrogated by a roving operator-held reader rather than continuous gateway monitoring.

5.5 Chemical, electrochemical, and biosensors

Chemical sensing closes the loop between equipment health and process quality. **Lubricant degradation** is monitored by viscosity sensors, dielectric-constant probes, and chemiresistive detection of degradation by-products. **Coolant contamination** is monitored by conductivity and pH electrodes. **Pesticide residue** at the sprayer outlet is monitored by electrochemical and field-effect-transistor biosensors functionalised with enzymes, antibodies, aptamers, or molecularly imprinted polymers [32, 33, 38, 43]. Akkaş et al. [23] review the increasingly central role of machine learning in transforming biosensor outputs from raw current/potential traces into actionable classifications, and document the rapid convergence of sensor hardware, on-edge inference, and cloud-based interpretation. Elli et al. [43] additionally review field-effect-transistor architectures for environmental and agricultural monitoring, which combine analyte selectivity with the manufacturability of standard semiconductor processes.

5.6 Optical and hyperspectral sensing

Beyond fibre-optic point and distributed sensors (§5.3), CM and crop-protection systems are increasingly equipped with **camera-based vision** (RGB, multispectral, hyperspectral, thermal) for spray-pattern validation, leak detection, and structural defect inspection. Terentev et al. [35] review the state of hyperspectral remote sensing for early plant-disease detection, with implications for closed-loop crop-protection scheduling that depends on equipment readiness. Hardie [44] reviews proximal soil-moisture sensors, several of which (dielectric, time-domain reflectometry, neutron) double as indicators of irrigation-system condition.

5.7 Wireless sensor networks and energy harvesting

Sensors must be powered, networked, and synchronized. Jawad et al. [45] review energy-efficient wireless sensor networks (WSNs) specifically for precision agriculture, surveying low-power-wide-area technologies (LoRaWAN, NB-IoT, Sigfox), medium-access protocols, and energy-harvesting strategies (solar, piezoelectric, RF). For CM specifically, the real limits are duty cycle (CM data has higher bandwidth requirements than soil/weather sensing), latency tolerance (alarm signals must propagate within seconds), and synchronisation (vibration and AE signals must be time-aligned across nodes). Aarif et al. [1] and Soussi et al. [2] survey the broader precision-agriculture sensor ecosystem that the CM layer must coexist with, and Prakash et al. [46] map the integration of IoT, wireless communication, and sensors in smart-farming deployments.

5.8 Sensors deployed in agricultural environments

A growing literature documents sensor performance specifically in agricultural service conditions [1, 2, 9, 26, 44, 38, 47]. Maraveas and Bartzanas [26] review structural-health-monitoring sensors for agricultural structures — greenhouses, silos, livestock buildings — and identify dust, condensation,

ammonia, and biological growth as the principal field-stressors. Kim and Lee [38] review electrochemical sensors for sustainable precision agriculture, with emphasis on the materials-stability constraints addressed in Section 4. Laveglia et al. [47] consolidate proximal-sensing techniques for sustainable crop management. These environment-specific findings should be read jointly with the CM-platform discussion of this section: a sensor architecture validated in a clean laboratory or factory cell rarely transfers unmodified to the field.

6. From sensors data to maintenance decisions: pipeline, AI, and digital twins

A modern CM-to-TPM pipeline is a layered system that transforms raw sensor signals into maintenance actions, organised as a vertical stack of seven layers in which each layer corresponds to one tier of the OSA-CBM architecture of Section 3.4 and feeds the next. Layer 1 sensors and front-end electronics — comprises the piezoelectric, MEMS, fibre-Bragg-grating, electrochemical, field-effect-transistor, chipless-RFID, and imaging devices of Section 5, conditioned by charge amplifiers, bridges, potentiostats, transimpedance amplifiers, and analogue-to-digital converters. Layer 2 edge processing extracts RMS, kurtosis, FFT, envelope, and wavelet-packet features at the sensor and upstreams raw waveforms only when a deviation is detected. Layer 3 diagnostics converts those features into a fault class, severity, and location through support-vector-machine, random-forest, convolutional, or recurrent neural-network classifiers. Layer 4 prognostics projects the trajectory forward in time using physics-based, data-driven, or hybrid remaining-useful-life models that report a point estimate together with its uncertainty. Layer 5, the digital twin integrates as-built geometry, component genealogy, exposure history, current and projected state, and rolls all of these up to an asset-level OEE = Availability × Performance × Quality score. Layer 6, the TPM decision frame consumes those outputs at the cadence dictated by each of the eight pillars: real-time alerts for autonomous maintenance and for safety/health/environment, daily-to-weekly scheduling for planned maintenance, monthly-to-quarterly reviews for focused improvement, batch-level reporting for quality maintenance, annual aggregates for early-equipment management, and continuous case-history feeds for training and office TPM. Layer 7 action closes the loop with operator alerts, scheduled interventions, design feedback to OEMs, and OEE reporting; the twin in turn updates sensor priors and edge-processing thresholds, completing the bidirectional coupling between physical asset and virtual representation.

6.1 Signal acquisition, conditioning, and edge processing

Sensor signals enter the pipeline through analogue front-ends matched to the transducer (charge amplifiers for piezoelectric devices, Wheatstone bridges for strain gauges, potentiostats for electrochemical electrodes, transimpedance

amplifiers for photodetectors). Anti-alias filtering, gain ranging, and ADC conversion produce digital streams that are either transmitted raw or pre-processed on-edge. Edge processing, increasingly performed by low-power microcontrollers with hardware DSP and neural-accelerator blocks, is necessary whenever the raw bandwidth exceeds the available link capacity, as is generally the case for vibration and AE sensing. Typical on-edge operations include RMS and crest-factor extraction, FFT and short-time spectral features, envelope demodulation around bearing characteristic frequencies, and threshold-based event detection that triggers up-streaming of raw waveforms when a deviation is detected [24, 36].

6.2 Feature extraction and diagnostics (state detection and health assessment)

The OSA-CBM layers of state detection and health assessment have historically relied on hand-crafted features, time-domain statistics, frequency-domain peaks, wavelet-packet energies fed into classical classifiers such as support vector machines, random forests, or shallow neural networks [36]. The recent literature on data-driven PHM has shifted toward **end-to-end deep learning** in which convolutional and recurrent networks operate directly on segmented time-series or spectrogram inputs and learn discriminative features implicitly [36, 37]. Park et al. [37] illustrate this with a long short-term memory (LSTM) model applied to cryogenic safety-valve signals, achieving fault classification accuracy above 95 % across multiple fault modes. The trade-off is data hunger: deep models require either large labelled fault datasets or carefully constructed transfer-learning paths from simulation or adjacent assets.

6.3 Prognostics and remaining-useful-life estimation

Prognostic assessment fits the trajectory of one or more health indicators forward in time. The methodological literature is divided into three families. **Physics-based prognostics** integrate a damage-evolution model (Paris–Erdogan crack growth, electrochemical corrosion kinetics, lubricant degradation) over operating conditions; they require explicit damage models but extrapolate well outside the training distribution. **Data-driven prognostics** learn the trajectory from historical run-to-failure datasets, typically using recurrent networks, attention models, or particle filters [36]. **Hybrid prognostics** combine the two, anchoring a data-driven model on physics-informed features or constraints. Chen et al. [36] survey the rapid recent progress in deep-learning RUL prediction, including the migration toward transformer architectures and toward uncertainty-aware models that report confidence intervals rather than point estimates. For agricultural assets, the principal challenges remain the scarcity of run-to-failure data and the wide variability of operating conditions across crops, seasons, and regions [29].

6.4 Decision support and AI integration

Maintenance decisions are not made on RUL alone; they trade against production schedules, parts and labour availability, agronomic calendars, and safety constraints. The decision-support layer therefore integrates CM outputs with operations data and either presents them to human planners through dashboards or, increasingly, optimizes maintenance actions automatically using reinforcement-learning or mixed-integer-programming solvers [16, 17]. Hassan et al. [48] survey monitoring and advanced control strategies in smart agriculture, and Subeesh and Mehta [8] consolidate the AI and IoT methods most relevant to agricultural decision support. Werbińska-Wojciechowska and Winiarska [16] additionally provide a bibliometric and systematic review map of how maintenance performance has been studied across the Industry 4.0 transition, identifying the convergence on data-driven decision support as the dominant research direction.

6.5 Digital twins for maintenance decision support

A **digital twin** is a structured, dynamically updated virtual representation of a physical asset that is bidirectionally coupled to it through sensor data and control feedback. For maintenance, the twin holds (i) the asset's geometry, material genealogy, and component-level bill of materials; (ii) the as-operated history including cumulative load, vibration, temperature, and chemical exposure; (iii) the current and projected state of each instrumented subsystem; and (iv) the planned maintenance schedule against which RUL forecasts are evaluated. Zhang et al. [29] review digital-twin approaches specifically for agricultural machinery damage prediction and maintenance, documenting both the substantial methodological progress in linking physics-based and data-driven twins and the persistent gap in agricultural-specific standardized twin schemas. The digital twin is the natural home for the OEE accounting framework of Section 3: availability, performance, and quality losses can be attributed to component-level events in the twin and rolled back up to asset-level OEE in real time.

Figure 5 shows a representative digital-twin architecture for agricultural-machinery maintenance: a physical layer of on-machine sensors feeds a data-processing layer, which drives a virtual layer performing real-time monitoring, intelligent diagnosis, and maintenance decisions, closed by a real-time feedback loop to the physical asset.

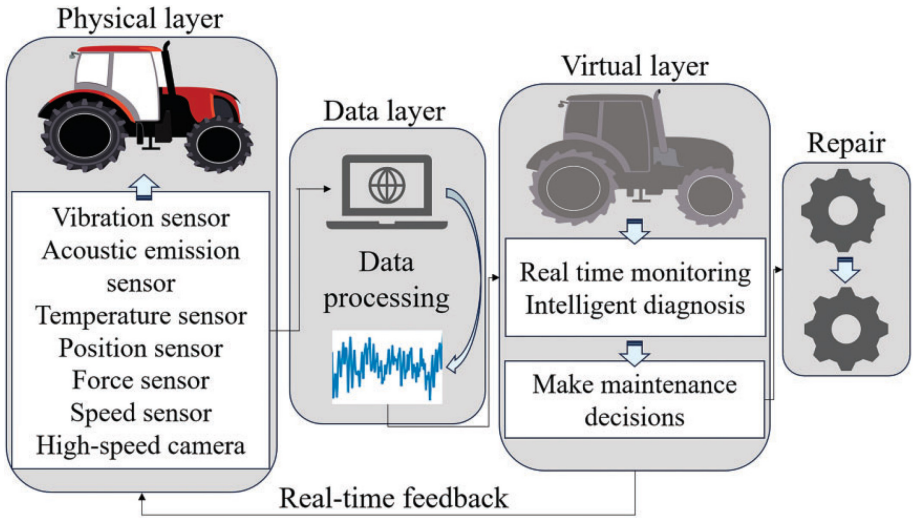


Figure 5. Digital-twin architecture for agricultural-machinery condition monitoring and maintenance, showing the physical → data → virtual layers with a real-time feedback loop. Reproduced from Ref. [28].

6.6 Mapping CM outputs to the TPM eight pillars

The final integration step is to align CM-to-twin outputs with the TPM organizational frame. The mapping is asymmetric; each pillar consumes different CM information at different cadences, as summarized in Table 3.

Table 3. Mapping of CM-pipeline outputs onto the eight TPM pillars and their characteristic decision cadences.

TPM pillar	Principal CM inputs	Decision cadence
Autonomous maintenance	Visual checks; operator-facing alerts from edge processing	Per shift / per pass
Planned maintenance	Diagnostic state; prognostic RUL	Daily / weekly
Focused improvement	Pareto of recurrent failure modes from twin history	Monthly / quarterly
Quality maintenance	Process-side measurements tied to equipment condition	Per batch / per pass
Early-equipment management	Cross-asset failure-mode aggregates fed back to OEM/design	Annual
Education and training	Anonymised case histories from twin	Continuous
Office TPM	Maintenance-cost and OEE roll-ups	Monthly
Safety, health, environment	Anomaly alerts from chemical and pressure sensors	Real time

Reported case studies in process and discrete manufacturing, Shannon et al. [14] in pharmaceuticals, Drewniak and Drewniak [12] in crude-oil processing, Mendes et al. [15] on a conveyor, consistently document substantial OEE gains when the CM-to-TPM stack is implemented end-to-end. Translating these gains to agricultural fleets remains a research opportunity addressed in Sections 7 and 9.

7. Case study: Condition monitoring and total productive maintenance of crop-protection equipment

7.1 Why crop-protection equipment is a representative case

Crop-protection operations are simultaneously among the most economically consequential and the most failure-prone activities in modern agriculture [30, 31, 34]. A sprayer pass conducted with an out-of-specification nozzle, a degraded pump, or a misregulated dosing controller can under-treat the crop and cause yield loss, over-treat it and inflate input costs, or scatter agrochemical drift into adjacent ecosystems and water bodies [30, 31, 33]. The same operations expose equipment to abrasive suspensions, corrosive liquids, mechanical vibration, and ultraviolet weathering, a combination that few industrial assets encounter so concentrated [26, 27, 28]. The CM-to-TPM problem is therefore both economically motivated and technically demanding, and provides a natural integration test for the materials and sensor layers reviewed in Sections 4–5 and the data and digital-twin layers reviewed in Section 6.

7.2 Equipment taxonomy and service environment

For the purpose of this case study, "crop-protection equipment" denotes four operational classes that share substantial subsystem commonality:

- **Self-propelled and tractor-mounted field sprayers** with wide booms (commonly 18–36 m), centrifugal or diaphragm pumps, hydraulic boom suspensions, and tens to hundreds of nozzles on a single rail [30].
- **Orchard and vineyard air-blast sprayers** axial fans, lower flow rates per nozzle but higher angular variability, and frequent canopy-aware variable-rate operation [31].
- **Unmanned-aerial-vehicle (UAV) and unmanned-ground-vehicle (UGV) applicators** use small electric pumps, lightweight nozzles, and high duty cycles per battery pack; CM is constrained by payload and energy budgets [31].
- **Greenhouse and protected-environment dosing systems,** venturi or piston metering, recirculating spray loops, and tight regulatory limits on residue [34].

In all four classes, the wetted path comprises a tank, a pump, a filtration stage, a pressure-regulation valve, a distribution manifold, and nozzles; the dry side adds a chassis, boom, or fan structure, dosing electronics, and a navigation / positioning subsystem. Each of these has a distinct failure signature and demands a different sensor modality, as developed below.

7.3 Dominant failure modes and degradation pathways

7.3.1 Nozzle wear and droplet-size drift

Nozzle orifices erode under abrasive suspension flow. Fertilizer slurries and clay-based formulations are particularly aggressive, and orifice enlargement of 10 % typically inflates flow by a comparable margin while shifting droplet-size distribution toward larger volumes (the agronomic-extension threshold for nozzle replacement is conventionally a flow drift above 10 % of nominal). Wear is invisible to the eye and progresses slowly, so it is a textbook target for sensor-based CM: differential-flow measurement across symmetric boom segments, in-line droplet-size estimation from acoustic signatures, or visual analysis of spray pattern.

7.3.2 Pump cavitation, seal failure, and elastomer swelling

Centrifugal and diaphragm pumps in crop-protection service cavitate when filter pressure drop rises or when surfactant-rich formulations lower effective vapor pressure. Cavitation broadcasts characteristic high-frequency acoustic emissions and a distinct vibration spectrum that is readily detected by piezoelectric or MEMS accelerometers [18, 24]. Elastomeric seals additionally suffer chemical swelling and embrittlement under repeated agrochemical exposure; the resulting leakage paths are detectable by pressure-decay tests during idle periods and by capacitive or chemiresistive moisture sensors at the pump housing [20].

7.3.3 Corrosion of metallic frames, fittings, and tanks

Steels, aluminium alloys, and brasses used in sprayer frames, fittings, and tank fittings are simultaneously attacked by acidic pesticide formulations, alkaline liquid fertilisers, salt-laden irrigation water, and biofilm-mediated microbial corrosion [27, 28]. Mitigation relies on a combination of (i) stainless-steel grades, polyolefin or PTFE-lined wetted paths, and EPDM/FKM elastomer selection for chemical compatibility; (ii) protective paints and powder coatings on dry-side structure; and (iii) corrosion inhibitors injected in flush cycles [27, 28]. CM of corrosion progression remains under-instrumented in field equipment; promising approaches include chipless-RFID corrosion-sensitive tags [25] and embedded electrochemical impedance probes.

7.3.4 Boom structural fatigue

Long booms oscillate vertically and laterally under field traverse and wind loading. Repeated strain reversals fatigue welds and folding joints; a partial failure mid-pass can scatter spray and damage the crop. Distributed strain sensing along the boom, historically by foil gauges, increasingly by optical fibres or 2-D-material-based flexible films [21, 22], provides both real-time alarm and accumulated-damage histories suitable for prognostics.

7.3.5 Dosing-electronics and metering-valve failures

Electronic flow controllers and solenoid metering valves drift with vibration, temperature, and chemical attack. Self-test routines (open-loop flow against commanded duty cycle) and redundancy with at least one independent flow sensor are the standard CM responses [30, 31].

7.4 Materials selection for agrochemical compatibility

The materials side of the case study is, in practice, the real limit. The wetted-path materials must resist (i) localised acidic attack from sulphur-based and chlorine-based active ingredients; (ii) swelling under aromatic-solvent emulsion concentrates; (iii) galvanic effects at brass/aluminium interfaces in tank fittings; and (iv) cavitation erosion at pump impellers. Field experience reported in agricultural-corrosion surveys [27, 28] favours **AISI 316L** and duplex stainless grades for pump bodies and pressure manifolds, **PTFE and high-density polyethylene** liners for tanks and hoses, and **EPDM** or **FKM** elastomers for valve seats matched to the formulation chemistry. Coatings, zinc-rich primers, powder-cured polyester, and increasingly polyurethane-ceramic hybrids protect the dry-side structure. The corresponding sensor envelope must match these wetted-path constraints: piezoelectric stacks for vibration sensing should be housed in stainless or polymer-lined enclosures, electrochemical sensing electrodes must be chosen for compatibility with the analyte chemistry, and any chemiresistive film exposed to the spray atmosphere must tolerate transient peak concentrations several orders of magnitude above its target sensitivity range [19, 20].

7.5 Sensor-material integration for crop-protection condition monitoring

7.5.1 Flow, pressure, and droplet quality

Inline turbine, ultrasonic, or magnetic-inductive flow meters are now standard on commercial variable-rate sprayers and provide the per-section flow data needed both for application control and for nozzle-wear diagnosis. Pressure transducers based on silicon piezo-resistive elements or MEMS capacitive diaphragms close the loop on pump performance. Real-time droplet-size estimation, traditionally limited to laboratory laser-diffraction systems, is

increasingly achievable in-field via acoustic signatures, machine-vision droplet imaging, or impedance changes across droplet-impacted sensing arrays [30, 31].

7.5.2 Pesticide-residue and active-ingredient sensing

Electrochemical and chemiresistive sensors functionalised with **molecularly imprinted polymers**, **enzyme layers**, and **nanostructured metal-oxide or graphene-derivative films** now detect organophosphate, carbamate, neonicotinoid, and glyphosate residues at sub-ppm levels [32, 33]. These devices complement field-effect-transistor biosensors that target specific pesticide analytes with high selectivity. In a CM context, in-line residue sensors serve two functions: they verify that the dispensed formulation matches specification, and they signal cross-contamination between tanks switched mid-season.

Figure 6 maps the contemporary landscape of electrochemical pesticide-residue sensing, in which nanomaterials (MXenes, graphene, MOFs, molecularly imprinted polymers, nanozymes), miniaturised and wearable device formats, and AI-driven data analysis are converging on the same building blocks that an in-line crop-protection CM sensor would draw on.

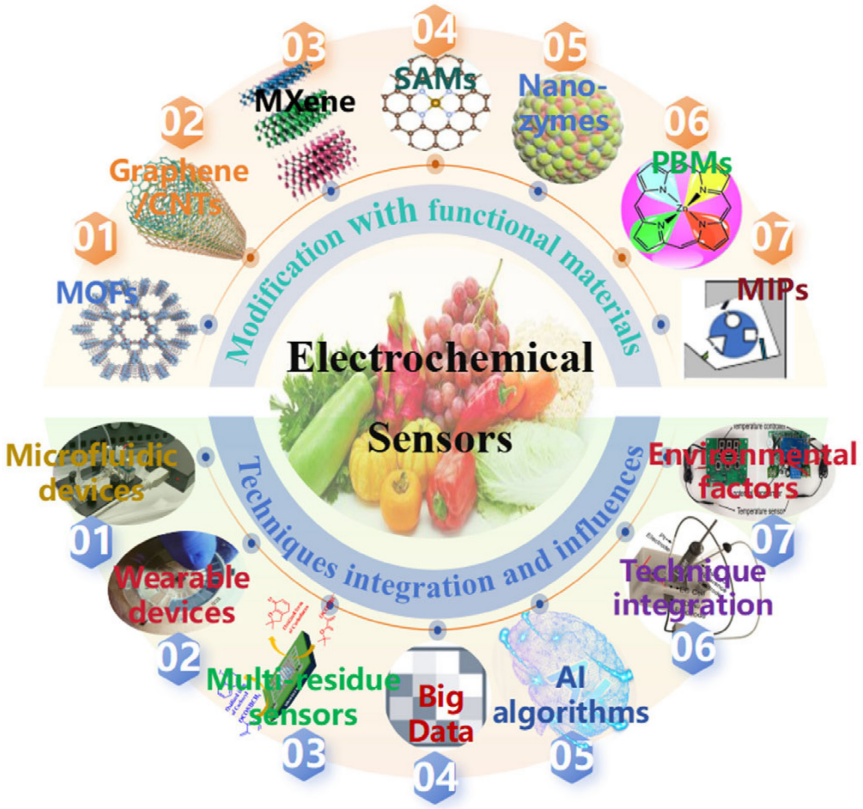


Figure 6. Landscape of nanomaterials, device formats, and data methods for electrochemical pesticide-residue detection. Reproduced from Ref. [31].

7.5.3 Vibration and acoustic sensing on pumps and booms

Piezoelectric and MEMS accelerometers mounted on pump housings yield the cavitation and bearing-wear signatures discussed in §7.3.2. Acoustic-emission sensors push the detection bandwidth above 100 kHz, where incipient pitting and seal failure radiate before becoming hydraulically significant [24]. Chipless-RFID resonant sensors offer a battery-free, low-cost form factor that can be distributed along a boom or inside a tank for occasional interrogation rather than continuous streaming [25].

7.5.4 Visual and hyperspectral validation of spray pattern

Downward-looking RGB and multispectral imagers on UAVs and on the sprayer itself validate spatial coverage at the field level. Hyperspectral imaging additionally allows early detection of disease and weed pressure that drives the application decision in the first place [35]. From a CM perspective, departure of the observed spray pattern from the commanded one is a powerful integrative diagnostic, summing nozzle wear, pressure drift, boom dynamics, and dosing errors into one observable.

7.5.5 Summary mapping of failure modes to sensor modalities

The cross-tabulation in Table 2 consolidates §7.3 and §7.5: each dominant failure mode of crop-protection equipment is mapped to the sensor modalities that detect it, together with the appropriate sensor-material classes from Table 1. ◉ marks a primary detection modality (sensitive, established, field-deployable); ● a useful secondary modality; ◐ an emerging but not yet field-deployed approach; blank indicates no significant role.

Table 2. Failure-mode × sensor-modality matrix for crop-protection equipment. Material codes M1–M14 refer to Table 1.

Failure mode	Vibration / AE	Pressure / flow	Strain / SHM	Electrochemical / FET	Chipless RFID	Visual / HSI	Principal material classes
Nozzle orifice wear	● (acoustic droplet signature)	◉ (per-section flow drift)	◉ (spray-pattern imaging)	M1, M2, M14			
Pump cavitation	◉ (broadband AE, low-freq vib.)	● (suction-side pressure)	M1, M2				
Seal failure / leakage	◐ (AE on housing)	● (pressure decay test)	● (chemiresistive moisture)	●	◐ (thermal imaging)	M1, M5, M13	
Agrochemical corrosion of frame/fittings	● (FBG corrosion-coupon)	◉ (impedance probe)	◉ (corrosion-sensitive tag)	● (visual coating defects)			
Boom	● (modal-shift	◉ (FBG / 2-	● (passive	●			

Failure mode	Vibration / AE	Pressure / flow	Strain / SHM	Electrochemical / FET	Chipless RFID	Visual / HSI	Principal material classes
structural fatigue	detection)	D film distributed)	strain tag)				
Dosing electronics / valve drift	◉ (closed-loop flow check)	● (residue cross-check)	M2, M11				
UAV battery / motor health	◉ (IMU vibration)	M2					
Pesticide-residue / formulation off-spec	◉ (electrochemical / FET)	● (VOC headspace)	● (HSI of spray cloud)	M4, M5, M10, M11, M12			
Tank biofouling / contamination	● (turbidity, flow drop)	● (impedance)	●	●			
UV / weathering of dry-side coatings	● (distributed FBG)	● (HSI / multispectral)	M8, M14				

The matrix surfaces two patterns. *First*, every failure mode has at least one **field-deployable (◉) primary modality**, confirming the §7.8 observation that transduction physics is not the real limit. *Second*, the matrix is sparsely populated rather than dense, meaning that an end-to-end CM stack for crop-protection equipment can be assembled from a relatively small set of sensor families (vibration/AE, flow/pressure, distributed strain, electrochemical, chipless-RFID, imaging) rather than requiring exotic per-failure-mode hardware.

7.6 UAV and robotic-sprayer condition monitoring

UAV applicators add their own failure modes: battery degradation, brushless-motor bearing wear, propeller imbalance, and GPS/RTK drift, but they also remove some constraints (vibration is now airframe-borne, agrochemical exposure is lower-volume) and enable new sensing geometries. Onboard inertial measurement units provide free vibration data, while battery-management systems supply state-of-health estimates that map directly onto OEE availability and performance components [31].

7.7 Digital-twin and TPM integration for crop-protection fleets

Recent reviews have begun to map digital-twin architectures specifically onto agricultural machinery [29]. For a crop-protection fleet, the digital twin must represent (i) the per-machine wetted-path geometry and component genealogy, (ii) the cumulative agrochemical exposure history of each wetted component, (iii) the current and prognosed condition of every CM-instrumented subsystem, and (iv) the agronomic schedule that defines available maintenance windows. Coupled with remaining-useful-life models [36, 37] and the OEE accounting

framework [13], such a twin supports the TPM eight-pillar workflow: autonomous maintenance (operator-level pre- and post-pass inspections informed by twin-flagged risk), planned maintenance (windowed against the agronomic calendar), focused improvement (root-cause analysis on recurrent failure modes), and early-equipment management (twin-informed feedback to OEM design). Reported case studies in process and discrete manufacturing [14, 15, 17] consistently document OEE gains of 10–30 percentage points when CM-driven TPM displaces calendar-based preventive maintenance; comparable studies for agricultural fleets are still sparse and represent a clear research opportunity.

7.8 Reported outcomes and remaining limitations

The literature reviewed in §7.5–§7.7 supports four observations for crop-protection equipment. *First*, sensor materials and devices for the principal failure modes are individually mature; integration into field-deployable CM systems is the bottleneck rather than transduction physics. *Second*, agrochemical compatibility both of the wetted-path equipment and of the sensors themselves remains the dominant durability constraint and is not yet adequately captured by standardized qualification protocols. *Third*, the connection between in-field CM data and TPM-grade decisions is still typically manual or vendor-specific; open digital-twin schemas for agricultural maintenance are an open standards gap. *Fourth*, the economic case for crop-protection-equipment TPM is strong but under-quantified, with most reported numbers extrapolated from adjacent manufacturing rather than measured on agricultural fleets. Closing these four gaps is the practical agenda that motivates the future-directions section of this review.

8. Cross-cutting challenges and gaps

The integrated picture developed in Sections 3–7 surfaces six cross-cutting challenges that constrain the deployment of a CM-driven TPM stack on agricultural equipment in general, and on crop-protection equipment in particular. Each cuts across the materials, sensor, and decision layers and therefore cannot be resolved within a single discipline.

8.1 Material durability under agrochemical and field stress

The first and arguably most binding challenge is that sensor materials qualified in clean industrial environments routinely fail under prolonged agrochemical exposure, UV, biofouling, and freeze–thaw cycling. Electrochemical electrodes lose calibration through analyte poisoning [20, 32, 38]; MOS sensors drift under humid, dusty operation [19]; piezoelectric stacks delaminate at adhesive interfaces under high-vibration combined with thermal cycling [18]; optical

fibres develop sensitivity loss at chemically attacked claddings [22]. Standardized durability qualification protocols for sensors in agricultural service do not yet exist and would represent a high-leverage research and standards contribution [26, 27, 28].

8.2 Calibration drift and recalibration logistics

Even durable sensors drift. The drift problem is qualitatively different in agriculture than in factory environments: assets are remote, recalibration windows are narrow, and a calibration excursion that would be caught within hours in a factory may persist for an entire spraying season in the field. Self-calibrating sensors, those that use redundant measurement principles, in-situ reference cells, or model-based fault detection on the calibration trend itself, are an active research area [23, 37]. The decision layer must also be drift-aware: prognostic models trained on a calibrated baseline degrade quietly when the calibration baseline shifts.

8.3 Power, connectivity, and synchronisation in rural settings

CM systems have higher bandwidth and lower latency tolerance than the broader precision-agriculture sensor mix [45]. Sustained vibration streaming at ≥10 kHz exceeds the practical capacity of LoRaWAN and NB-IoT; AE data is more demanding still. Continuous on-edge feature extraction with event-triggered up-streaming is the standard compromise but pushes power and thermal load onto the sensor node. Energy harvesting — solar, vibration, RF — partially addresses the power problem but introduces its own reliability concerns. Time-synchronisation across distributed nodes, essential for correlated vibration analysis on multi-bearing assets, remains an open practical problem on agricultural fleets.

8.4 Standardisation and interoperability

The corpus reviewed in this paper documents a proliferation of vendor-specific, application-specific implementations of CM, prognostics, and digital twins for agricultural assets. Open data schemas extending OSA-CBM, MIMOSA, and the emerging Asset Administration Shell to the agricultural-machinery domain and open OEE accounting frameworks aligned with the eight-pillar TPM ontology-would substantially accelerate cross-vendor learning [16, 29]. The absence of an agreed ontology specifically for **crop-protection equipment** is particularly acute: failure-mode taxonomies, severity gradings, and intervention codes are currently OEM-specific.

8.5 Cost–benefit and adoption barriers for heterogeneous fleets

The agricultural-equipment user population is heterogeneous, ranging from large contract sprayer operators in industrial farming to small-holders in low-income

settings. The TPM business case is strong at the high end of this distribution but weak at the low end, where the per-asset cost of CM instrumentation, connectivity subscription, and decision-support software exceeds the marginal value of avoided downtime. Closing this gap will require not only cheaper sensor materials and devices (notably chipless-RFID and ultra-low-power MEMS [25]) but also new business models, sensor-as-a-service, cooperative twin operation, and contractor-shared CM data that distribute the fixed cost of the decision layer [3, 5, 7, 9].

8.6 Sustainability and end-of-life of sensor electronics

A final cross-cutting consideration is environmental: a dense CM deployment introduces non-trivial electronic-waste, battery, and rare-earth-material footprints. Several sensor-material classes reviewed in Section 4 notably nano-enabled platforms and lead-containing piezoelectrics, face increasingly tight end-of-life regulation [39, 40, 41]. The CM-design discipline must therefore extend beyond performance to lifecycle impact, favouring lead-free piezoelectrics, biodegradable substrates where service life permits, and recoverable architectures elsewhere.

9. Future directions and conclusions

9.1 Future research directions

Five research directions emerge as both technically tractable and practically high-impact:

- **Self-powered and self-healing sensor materials.** Piezoelectric energy harvesting paired with vibration sensing on the same element [18], MXene- and graphene-derivative coatings that retain function after micro-damage [21], and self-passivating electrochemical electrodes would together resolve the power and durability challenges of §8.1 and §8.3.
- **Sensor-fusion and physics-informed prognostics for agricultural assets.** Combining vibration, AE, electrochemical, and visual modalities into joint health indicators — and constraining the fusion by physics-based damage models would close the data-scarcity gap that currently limits deep-learning RUL methods in agriculture [36, 29].
- **Federated learning and fleet-level digital twins.** Privacy-preserving aggregation of CM data across machines, farms, and OEMs would allow rare failure modes to be learned at fleet scale without centralizing sensitive operational data [23, 29]. Federated digital twins of crop-protection fleets, in particular, would enable benchmarking of OEE and intervention efficacy across regions and crop systems.

- **Open standards for agricultural CM-to-TPM.** The standards gap of §8.4 will not be closed by individual research papers; it will require coordinated proposals from OEMs, research institutions, and regulators. A useful starting point is a domain extension of ISO 17359 and ISO 13374 for agricultural machinery, complemented by an OEE-with-eight-pillars accounting schema specific to crop-protection operations.
- **Quantified economic evaluation of TPM on agricultural fleets.** Rigorous before-and-after studies with measured OEE, MTBF, MTTR, and total cost of ownership would replace the current reliance on extrapolation from adjacent manufacturing [12, 14, 15] and clarify the conditions under which TPM is economically justified at different fleet scales.

9.2 Conclusions

This review has consolidated the materials, sensor, and decision-layer literatures relevant to condition monitoring and Total Productive Maintenance of agricultural equipment, with particular focus on crop-protection machinery. Three conclusions are warranted.

First, the sensor-material toolbox required for agricultural CM is largely mature at the device level: piezoelectric, MOS, electrochemical, optical-fibre, 2-D-material, nano-enabled, and radiation-tolerant classes each cover a well-defined niche, and the integration challenge is now durability under field exposure rather than transduction physics. *Second*, the maintenance-philosophy stack of CBM, predictive maintenance, and TPM is well established in industrial manufacturing, with OEE and the eight pillars providing a coherent accounting and organizational frame; the translation to agriculture is hindered less by missing concepts than by the absence of domain-specific data schemas, durability standards, and economic evidence. *Third*, crop-protection equipment offers a uniquely concentrated case study in which materials, sensors, and decision systems must work jointly under aggressive service conditions; advances on this case generalize readily to tillage, harvest, and storage equipment.

The strongest research opportunity surfaced by the review is the construction of open, federated digital twins of crop-protection fleets, instrumented with chemically-tolerant sensor materials, organized by an agriculture-specific eight-pillar TPM ontology, and benchmarked against measured OEE gains. This is the integration point at which materials science, sensor engineering, data and AI engineering, and operations management converge — and at which the next generation of agricultural-mechanical-engineering research is best directed.

References

[1] M. Aarif K. O., A. Alam, and Y. Hotak, "Smart Sensor Technologies Shaping the Future of Precision Agriculture: Recent Advances and Future Outlooks", *Journal of Sensors*, vol. 2025, no. 1, 2025, doi: 10.1155/js/2460098.

[2] A. Soussi, E. Zero, R. Sacile, D. Trinchero, and M. Fossa, "Smart Sensors and Smart Data for Precision Agriculture: A Review", *Sensors*, vol. 24, no. 8, pp. 2647, 2024, doi: 10.3390/s24082647.

[3] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big Data in Smart Farming – A review", *Agricultural Systems*, vol. 153, pp. 69-80, 2017, doi: 10.1016/j.agsy.2017.01.023.

[4] M. A. DAYIOĞLU and U. TURKER, "Digital Transformation for Sustainable Future - Agriculture 4.0 : A review", *Tarım Bilimleri Dergisi*, 2021, doi: 10.15832/ankutbd.986431.

[5] E. M. B. M. Karunathilake, A. T. Le, S. Heo, Y. S. Chung, and S. Mansoor, "The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture", *Agriculture*, vol. 13, no. 8, pp. 1593, 2023, doi: 10.3390/agriculture13081593.

[6] N. Khan, R. L. Ray, G. R. Sargani, M. Ihtisham, M. Khayyam, and S. Ismail, "Current Progress and Future Prospects of Agriculture Technology: Gateway to Sustainable Agriculture", *Sustainability*, vol. 13, no. 9, pp. 4883, 2021, doi: 10.3390/su13094883.

[7] S. Getahun, H. Kefale, and Y. Gelaye, "Application of Precision Agriculture Technologies for Sustainable Crop Production and Environmental Sustainability: A Systematic Review", *The Scientific World Journal*, vol. 2024, no. 1, 2024, doi: 10.1155/2024/2126734.

[8] A. Subeesh and C. Mehta, "Automation and digitization of agriculture using artificial intelligence and internet of things", *Artificial Intelligence in Agriculture*, vol. 5, pp. 278-291, 2021, doi: 10.1016/j.aiia.2021.11.004.

[9] S. Mansoor, S. Iqbal, S. M. Popescu, S. L. Kim, Y. S. Chung, and J. H. Baek, "Integration of smart sensors and IOT in precision agriculture: trends, challenges and future prospectives", *Frontiers in Plant Science*, vol. 16, 2025, doi: 10.3389/fpls.2025.1587869.

[10] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance", *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510, 2006, doi: 10.1016/j.ymssp.2005.09.012.

[11] B. de Jonge, R. Teunter, and T. Tinga, "The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance", *Reliability Engineering & System Safety*, vol. 158, pp. 21-30, 2017, doi: 10.1016/j.ress.2016.10.002.

[12] R. Drewniak and Z. Drewniak, "Improving business performance through TPM method: The evidence from the production and processing of crude oil",

PLOS ONE, vol. 17, no. 9, pp. e0274393, 2022, doi: 10.1371/journal.pone.0274393.

[13] L. d. C. Ng Corrales, M. P. Lambán, M. E. Hernandez Korner, and J. Royo, "Overall Equipment Effectiveness: Systematic Literature Review and Overview of Different Approaches", *Applied Sciences*, vol. 10, no. 18, pp. 6469, 2020, doi: 10.3390/app10186469.

[14] N. Shannon, A. Trubetskaya, J. Iqbal, and O. McDermott, "A total productive maintenance & reliability framework for an active pharmaceutical ingredient plant utilising design for Lean Six Sigma", *Heliyon*, vol. 9, no. 10, pp. e20516, 2023, doi: 10.1016/j.heliyon.2023.e20516.

[15] D. Mendes, P. D. Gaspar, F. Charrua-Santos, and H. Navas, "Integrating TPM and Industry 4.0 to Increase the Availability of Industrial Assets: A Case Study on a Conveyor Belt", *Processes*, vol. 11, no. 7, pp. 1956, 2023, doi: 10.3390/pr11071956.

[16] S. Werbińska-Wojciechowska and K. Winiarska, "Maintenance Performance in the Age of Industry 4.0: A Bibliometric Performance Analysis and a Systematic Literature Review", *Sensors*, vol. 23, no. 3, pp. 1409, 2023, doi: 10.3390/s23031409.

[17] M. Molęda, B. Małysiak-Mrozek, W. Ding, V. Sunderam, and D. Mrozek, "From Corrective to Predictive Maintenance—A Review of Maintenance Approaches for the Power Industry", *Sensors*, vol. 23, no. 13, pp. 5970, 2023, doi: 10.3390/s23135970.

[18] Y. Chen, X. Zhang, and C. Lu, "Flexible piezoelectric materials and strain sensors for wearable electronics and artificial intelligence applications", *Chemical Science*, vol. 15, no. 40, pp. 16436-16466, 2024, doi: 10.1039/d4sc05166a.

[19] A. Marikutsa, M. Rumyantseva, E. A. Konstantinova, and A. Gaskov, "The Key Role of Active Sites in the Development of Selective Metal Oxide Sensor Materials", *Sensors*, vol. 21, no. 7, pp. 2554, 2021, doi: 10.3390/s21072554.

[20] T. Baba et al., "Advancements in Chemiresistive and Electrochemical Sensing Materials for Detecting Volatile Organic Compounds in Potato and Tomato Plants", *AgriEngineering*, vol. 7, no. 6, pp. 166, 2025, doi: 10.3390/agriengineering7060166.

[21] Y. Sun, W. He, C. Jiang, J. Li, J. Liu, and M. Liu, "Wearable Biodevices Based on Two-Dimensional Materials: From Flexible Sensors to Smart Integrated Systems", *Nano-Micro Letters*, vol. 17, no. 1, 2025, doi: 10.1007/s40820-024-01597-w.

[22] R. Jha, P. Mishra, and S. Kumar, "Advancements in optical fiber-based wearable sensors for smart health monitoring", *Biosensors and Bioelectronics*, vol. 254, pp. 116232, 2024, doi: 10.1016/j.bios.2024.116232.

[23] T. Akkaş, M. Reshadsedghi, M. Şen, V. Kılıç, and N. Horzum, "The Role of Artificial Intelligence in Advancing Biosensor Technology: Past, Present, and

Future Perspectives", *Advanced Materials*, vol. 37, no. 34, 2025, doi: 10.1002/adma.202504796.

[24] I. U. Hassan, K. Panduru, and J. Walsh, "An In-Depth Study of Vibration Sensors for Condition Monitoring", *Sensors*, vol. 24, no. 3, pp. 740, 2024, doi: 10.3390/s24030740.

[25] H. G. Bilic, T. Buyukoztekin, and S. Ozdemir, "The Use of Chipless Sensors with RFID for Condition Monitoring", *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, pp. 1-4, 2018, doi: 10.1109/idap.2018.8620810.

[26] C. Maraveas and T. Bartzanas, "Sensors for Structural Health Monitoring of Agricultural Structures", *Sensors*, vol. 21, no. 1, pp. 314, 2021, doi: 10.3390/s21010314.

[27] G. E. M. Nasr, Z. Abdel Hamid, and M. Refai, "Agricultural Machinery Corrosion", in *Introduction to Corrosion — Basics and Advances*, IntechOpen, 2023, doi: 10.5772/intechopen.108918.

[28] National Physical Laboratory, *Corrosion Control of Agricultural Equipment and Buildings*, Guides to Good Practice in Corrosion Control, NPL, Teddington, UK. Available: http://resource.npl.co.uk/docs/science_technology/materials/life_management_of_materials/publications/online_guides/pdf/corrosion_control_of_agricultural_equipment_and_buildings.pdf [Accessed: 13 May 2026].

[29] C. Zhang, J. Song, X. Yin, J. Cai, Y. Liang, and J. Lu, "Digital twin-based approaches for agricultural machinery damage prediction and maintenance: A review", *Journal of Computational Design and Engineering*, vol. 12, no. 10, pp. 87-117, 2025, doi: 10.1093/jcde/qwaf097.

[30] G. A. Mesías-Ruiz, M. Pérez-Ortiz, J. Dorado, A. I. de Castro, and J. M. Peña, "Boosting precision crop protection towards agriculture 5.0 via machine learning and emerging technologies: A contextual review", *Frontiers in Plant Science*, vol. 14, 2023, doi: 10.3389/fpls.2023.1143326.

[31] W. Li et al., "A sustainable crop protection through integrated technologies: UAV-based detection, real-time pesticide mixing, and adaptive spraying", *Scientific Reports*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-19473-x.

[32] Y. Han, L. Zhu, R. Wen, J. Liao, B. Li, and L. Wu, "Innovative Electrochemical Sensors for Pesticide Residue Detection: Nanomaterials, Miniaturization, and Intelligent Data Analysis", *The Chemical Record*, vol. 25, no. 8, 2025, doi: 10.1002/tcr.202500070.

[33] J. Mazuryk, K. Klepacka, W. Kutner, and P. S. Sharma, "Glyphosate Separating and Sensing for Precision Agriculture and Environmental Protection in the Era of Smart Materials", *Environmental Science & Technology*, vol. 57, no. 27, pp. 9898-9924, 2023, doi: 10.1021/acs.est.3c01269.

[34] E. Mamabolo et al., "Application of precision agriculture technologies for crop protection and soil health", *Smart Agricultural Technology*, vol. 12, pp. 101270, 2025, doi: 10.1016/j.atech.2025.101270.

[35] A. Terentev, V. Dolzhenko, A. Fedotov, and D. Eremenko, "Current State of Hyperspectral Remote Sensing for Early Plant Disease Detection: A Review", *Sensors*, vol. 22, no. 3, pp. 757, 2022, doi: 10.3390/s22030757.

[36] F. Wu, Q. Wu, Y. Tan, and X. Xu, "Remaining Useful Life Prediction Based on Deep Learning: A Survey", *Sensors*, vol. 24, no. 11, pp. 3454, 2024, doi: 10.3390/s24113454.

[37] M. Kim, H. Seong, and D. Kim, "Deep Learning-Based Prognostics and Health Management Model for Pilot-Operated Cryogenic Safety Valves", *Sensors*, vol. 24, no. 6, pp. 1814, 2024, doi: 10.3390/s24061814.

[38] M. Y. Kim and K. H. Lee, "Electrochemical Sensors for Sustainable Precision Agriculture—A Review", *Frontiers in Chemistry*, vol. 10, 2022, doi: 10.3389/fchem.2022.848320.

[39] S. S. Ali et al., "Nanobiotechnological advancements in agriculture and food industry: Applications, nanotoxicity, and future perspectives", *Science of The Total Environment*, vol. 792, pp. 148359, 2021, doi: 10.1016/j.scitotenv.2021.148359.

[40] H. Wu and Z. Li, "Recent advances in nano-enabled agriculture for improving plant performance", *The Crop Journal*, vol. 10, no. 1, pp. 1-12, 2022, doi: 10.1016/j.cj.2021.06.002.

[41] N. N. Nguyen, N. T. Nguyen, P. T. Nguyen, Q. N. Phan, T. L. Le, and H. D. K. Do, "Current and emerging nanotechnology for sustainable development of agriculture: Implementation design strategy and application", *Heliyon*, vol. 10, no. 10, pp. e31503, 2024, doi: 10.1016/j.heliyon.2024.e31503.

[42] G. Lindström, "Radiation damage in silicon detectors", *Nuclear Instruments and Methods in Physics Research Section A*, vol. 512, no. 1–2, pp. 30–43, 2003, doi: 10.1016/S0168-9002(03)01874-6.

[43] G. Elli et al., "Field-Effect Transistor-Based Biosensors for Environmental and Agricultural Monitoring", *Sensors*, vol. 22, no. 11, pp. 4178, 2022, doi: 10.3390/s22114178.

[44] M. Hardie, "Review of Novel and Emerging Proximal Soil Moisture Sensors for Use in Agriculture", *Sensors*, vol. 20, no. 23, pp. 6934, 2020, doi: 10.3390/s20236934.

[45] H. Jawad, R. Nordin, S. Gharghan, A. Jawad, and M. Ismail, "Energy-Efficient Wireless Sensor Networks for Precision Agriculture: A Review", *Sensors*, vol. 17, no. 8, pp. 1781, 2017, doi: 10.3390/s17081781.

[46] C. Prakash, L. P. Singh, A. Gupta, and S. K. Lohan, "Advancements in smart farming: A comprehensive review of IoT, wireless communication, sensors, and hardware for agricultural automation", *Sensors and Actuators A: Physical*, vol. 362, pp. 114605, 2023, doi: 10.1016/j.sna.2023.114605.

[47] S. Laveglia, G. Altieri, F. Genovese, A. Matera, and G. C. Di Renzo, "Advances in Sustainable Crop Management: Integrating Precision Agriculture

and Proximal Sensing", *AgriEngineering*, vol. 6, no. 3, pp. 3084-3120, 2024, doi: 10.3390/agriengineering6030177.

[48] S. I. Hassan, M. M. Alam, U. Illahi, M. A. Al Ghamdi, S. H. Almotiri, and M. M. Su'ud, "A Systematic Review on Monitoring and Advanced Control Strategies in Smart Agriculture", *IEEE Access*, vol. 9, pp. 32517-32548, 2021, doi: 10.1109/access.2021.3057865.

Initial design and performance prediction of solar ORC system for hungarian climate conditions

Mannir USMAN[1], János BUZÁS[2], Istvan FARKAS[2]
[1] *Doctoral Program of Mechanical Engineering,*
Hungarian University of Agriculture and Life Sciences MATE, Gödöllő
[2] *Institute of Technology Hungarian University of Agriculture and Life Sciences,*
MATE, Gödöllő

Abstract

The solar organic Rankine cycle (ORC) is a promising technology that generates electricity by converting low to medium-solar energy sources into mechanical work. This study presents an initial design and performance prediction for a solar organic Rankine cycle based on Hungarian climatic conditions. A precise thermal analysis was conducted to investigate the influence of evaporation temperature on the power output and thermal efficiency. The analysis shows that increasing the evaporation temperature results in steady improvements in the entire cycle performance. The system produces approximately 32 kW and 13% thermal efficiency at the highest evaporation temperature, with 24 kW and 10% net power output and thermal efficiency at the lowest temperature considered. The study provides preliminary design parameters for the deployment of small-scale solar ORC in Hungary which will be investigated experimentally in the subsequent work.

Keywords

Solar organic Rankine cycle, working fluids, solar thermal collectors, thermal analysis, power generation

1. Introduction

The demand for energy continues to grow due to significant population growth, improved living standards, and the need for cutting-edge technological advancement. The adverse effects of burning fossil fuels for energy necessitate a transition to renewable energy sources that can utilize low-temperature heat for power generation (Tovar-Facio et al., 2022). Renewable energy technology plays a crucial role in meeting energy demands, reducing climate change, enhancing energy security, and promoting sustainable development (Ighalo et al., 2025). The organic Rankine cycle ORC) is one of the reliable technologies that generate electricity from low to medium temperature heat sources, such as solar thermal energy (Matuszewska., 2024).

Hungary is distinguished by medium solar irradiation, receiving annual direct normal irradiance (DNI) of 1022 to 1241 kWh/m^2 annually (Solar Atlas., 2024). This amount of radiation is more than enough for small scale solar ORC implementation. In this regard, Solar ORC technology demonstrates viable alternative for decentralized electricity generation based on Hungarian weather conditions.

This paper presents an initial stage design and performance prediction of solar ORC system for small scale applications based on Hungarian climate conditions. A steady-state analysis was carried out to investigate the system's performance under operating conditions, emphasizing key critical performance parameters, namely net power output and thermal efficiency. The objective is to evaluate the technical viability of implementing the system, the selection of a well-suited working fluid, and the intended output of the system. The paper further presents the design methodology followed, component selection, results and discussion, conclusion, and recommendations for future research.

1.1. Overview of solar ORC

The organic Rankine cycle (ORC) is a thermodynamic process for generating electricity or heat from medium- to low-temperature heat sources, utilizing an organic compound as the working fluid rather than water due to its lower boiling point and operating pressure (Babatunde & Sunday, 2018; Pethurajan et al., 2018). The ORC system consists of an expander, a condenser, a working pump, and an evaporator (Rahbar et al., 2017). This technology is flexible, enabling electricity generation from renewable sources such as solar, geothermal, waste heat, biomass, and ocean thermal energy conversion (OTEC) (Quoilin et al., 2013).

Solar-powered organic Rankine cycle (ORC) is an effective method for converting low-grade solar thermal energy into electricity with minimal environmental impact, as it emits no greenhouse gases. The solar ORC system integrates two basic cycles, as shown in Fig. 1a. The ORC cycle and a solar cycle. Heat from the solar thermal collector is transferred to a working fluid via the heat exchanger, transforming the fluid into vapor, which drives a turbine connected to a generator to produce electricity, as illustrated in the T-S diagram in Fig. 1b (Permana et al., 2022, Hossin, 2015).

Moreover, heating methods in solar ORC can be direct or indirect. Indirect systems use heat exchangers to heat the working fluid, while direct systems heat it directly in the solar collector (Matuszewska, 2024). Indirect solar ORC is being selected in this analysis for its thermal management and safety benefits.

a) b)

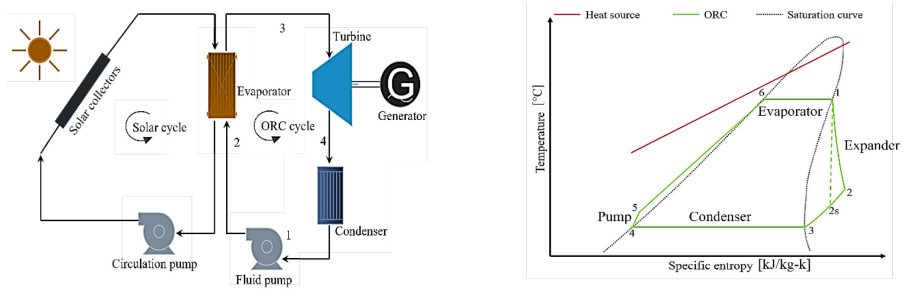


Figure. 1. Schematic diagram of a) solar ORC, b) T-S diagram

2. Material and method

2.1. Assessment of solar thermal resources in Hungary

According to the global solar atlas database, Hungary receives a daily average direct normal irradiance (DNI) of 2.8 to 3.4 kWh/m^2, representing 1022 to 1241 kWh/m^2 annually. This reveals the average level of solar energy availability in the region, which is compatible with solar thermal applications and photovoltaic systems (Permana et al., 2023). Several studies have shown that Hungary has not fully utilized its solar energy resources compared to other European countries (Talamon & Hartmann, 2017; Atsu et al., 2021). However, the region is committed to increasing its share of renewable energy by a greater percentage in the coming years (Said et al., 2023).

2.2. System configuration

The proposed solar ORC systems integrated five basic components: the solar collector, turbine, condenser, pump, and evaporator, as shown in Fig. 1 (Lee et al., 2017). Solar thermal collectors are used to absorb energy from the sun and convert it to thermal energy, which is used to heat the working fluid in the evaporator. The vaporized working fluid expands in the turbine to produce mechanical work, which is converted into electricity by the generator attached to it. The turbine outlet vapor is then condensed and pumped back to the evaporator to complete the cycle (Baral et al., 2015; Hijriawan et al., 2022).

2.3. Solar thermal collector selection

Evacuated tube collectors (ETC) were carefully chosen for this study due to their higher thermal efficiency and lower heat loss compared to flat plate collectors (FPC) (Yadav et al., 2024). ETC operates effectively in hot, cold, and cloudy

conditions, making it adaptable to various weather conditions and applications (Kumar et al., 2021). It is made up of vacuum tube heat pipe as shown in Fig. 2. Though ETC is slightly more expensive than FPC, in the long run, its higher thermal efficiency and greater thermal output compensate for these costs.

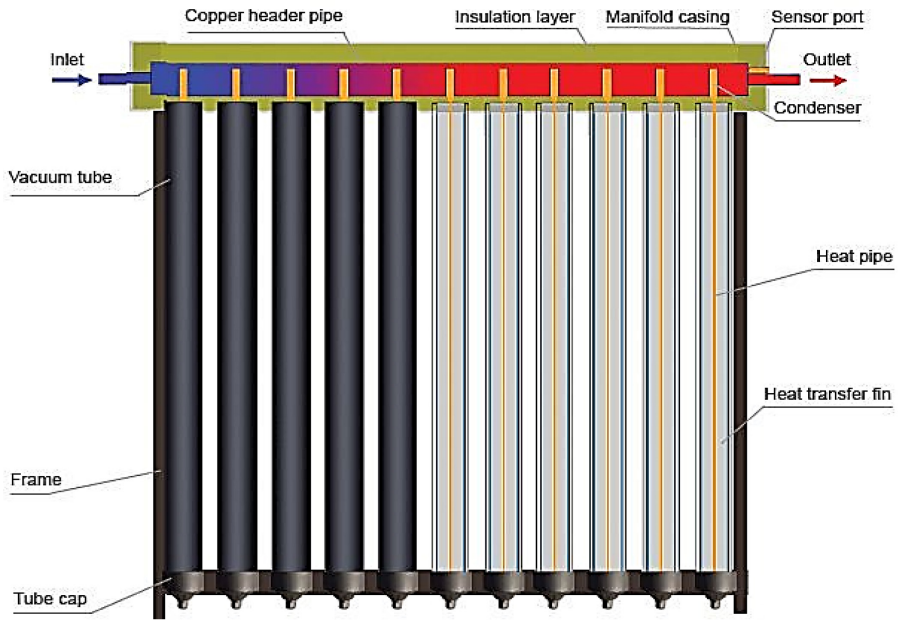


Figure. 2. Schematic diagram of evacuated tube collectors

2.4. Working fluid selection

Selection of a suitable working fluid for ORC systems plays a critical role in determining the cycle's performance, cost, and environmental footprint (Markides et al., 2025). Several criteria and factors are considered for selecting appropriate working fluids, such as thermophysical properties, stability, safety, availability, cost, and environmental data, which is the most important nowadays (Thangavel et al., 2021). R1233zd(E) is a suitable candidate for this analysis due to its ability to offset between thermal properties and environmental footprints. It serves as a substitute for R245fa; it is modern, safe, stable, and eco-friendly.

2.5. Performance prediction

The analysis presumed steady-state conditions and neglected heat loss, pressure drop, and friction in the system components (Akkaya, 2017; Lago et al., 2025; Wang et al., 2019). The performance of the systems is evaluated by measuring thermodynamic performance indicators such as net power output and thermal

efficiency, which can be expressed using Eq. 1:

$$\eta_{th} = \frac{\dot{W}_{net}}{\dot{Q}_{evap}} \quad (1)$$

where, $\dot{W}_{net}$ indicates net power output in (kW) and $\dot{Q}_{in}$ represents heat input to the system in (kW), the net power output is computed by Eq. 2:

$$\dot{W}_{net} = \dot{W}_{turb} - \dot{W}_{pump} \quad (2)$$

where, $\dot{W}_{turb}$ and $\dot{W}_{pump}$ indicate the work output of the turbine and pump respectively. The mathematical expressions for each component of the ORC cycle are presented in the subsequent equations:

Turbine:

$$\dot{W}_{Turb} = \dot{m}_f (h_1 - h_2) \quad (3)$$

Condenser:

$$\dot{Q}_{cond} = \dot{m}_f (h_3 - h_4) \quad (4)$$

Pump work:

$$\dot{W}_{pump} = \dot{m}_f (h_4 - h_3) \quad (5)$$

Evaporator:

$$\dot{Q}_{evap} = \dot{m}_f (h_1 - h_4) \quad (6)$$

where $\dot{m}_f$ represents the mass flow rate of the working fluid in (kg/s) and h indicates the specific enthalpy of the working fluid at inlet and outlet of each component in (kJ/kg.K) which are used to determine the turbine work and pump work. The standard properties correlation software like CoolProp are used to evaluate the thermal properties of the working fluid at each node of the cycle under varying environmental conditions.

2.6. Operating parameters

The proposed solar ORC was analysed at different evaporating temperatures ranging from 90 to 120 °C, representing the realistic maximum heat energy produced by the solar collectors across four seasons: winter, spring, summer, and autumn. The condenser temperature was fixed at 25 °C, which is equal to the Hungarian average environmental conditions. The efficiencies of the turbine and pump were taken as 80%. The evaporating and condensing pressures were determined from the saturation properties table of the working fluid. The detailed input parameters for this analysis were summarized in Table 1.

Table 1. Basic operating parameters

Parameters	Specification
Working fluid	R1233zd(E)
Solar collector type	Evacuated tube
Evaporation temperature (T_{evap}]	90 - 120 °C
Evaporation pressure [P_{evap}]	8 - 16 bar
Condensation temperature [T_{cond}]	25°C
Condensation pressure [P_{cond}]	1.3 bar
Turbine and pump efficiency [η_t , η_P]	70%

3. Results and discussion

Fig. 3 presents the change of net power output with evaporation temperature for the proposed solar ORC. The analysis indicates that the net power output increases steadily as the evaporator temperature increases. The power output rises from 24 kW at 90 °C to a maximum of 32 kW at 120 °C. This rise is primarily due to higher inlet temperature and pressure, which increase enthalpy across the turbine.

Fig. 4 shows the effect of evaporation temperature on the cycle's thermal efficiency. The proposed solar ORC yields thermal efficiency of approximately 10% at a lower evaporator temperature of 90 °C and 13% at the highest evaporator temperature of 120 °C. The rise in efficiency is primarily due to improved utilization of the heat input.

The analysis indicates that increasing the evaporation temperature significantly influences the performance measures of the systems, such as net power output and thermal efficiency. It also proved the suitability of solar ORC for small-scale applications.

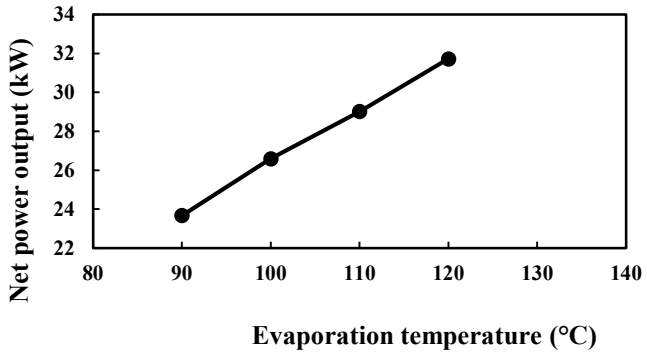


Figure. 3. Net power output as a function of evaporation temperature

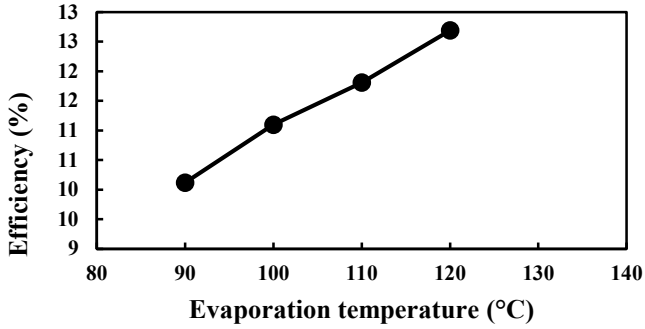


Figure. 4. Thermal efficiency as a function of evaporation temperature

Conclusions

Overall, this study presented predicted performance and initial design parameters of the solar ORC system in Hungary. Based on the thermal analysis conducted, the effect of evaporation temperature on thermal performance indices was analysed. The result shows that the net power output and thermal efficiency increase significantly with increasing evaporation temperature, with the highest power output of 32 kW and thermal efficiency of 13 % at the maximum evaporation temperature.

The result provides a technical background for solar ORC design and expected thermal output. It also serves as a framework for sizing, design, and performance predictions of small-scale solar ORC systems before experimental setup and testing. Future work will focus on building and testing the laboratory-scale solar ORC test rig to validate the design-predicted performance.

Acknowledgements

This work was supported by the Stipendium Hungaricum Programme and by the Doctoral School of Engineering Sciences, Hungarian University of Agriculture and Life Sciences, Gödöllő, Hungary.

References

[1] Akkaya, A. v. (2017). performance analysis of an organic Rankine cycle under different ambient conditions. In *Journal of Thermal Engineering* (Vol. 3, Issue 5). Yildiz Technical University Press.

[2] Atsu, D., Seres, I., & Farkas, I. (2021). The state of solar PV and performance analysis of different PV technologies grid-connected installations in Hungary. *Renewable and Sustainable Energy Reviews*, *141*.

https://doi.org/10.1016/j.rser.2021.110808

[3] Babatunde, A. F., & Sunday, O. O. (2018). A Review of Working Fluids for Organic Rankine Cycle (ORC) Applications. IOP Conference Series: Materials Science and Engineering, 413(1). https://doi.org/10.1088/1757-899X/413/1/012019

[4] Baral, S., Kim, D., Yun, E., & Kim, K. C. (2015). Energy, exergy and performance analysis of small-scale organic rankine cycle systems for electrical power generation applicable in rural areas of developing countries. *Energies*, *8*(2), 684–713. https://doi.org/10.3390/en8020684

[5] Global Solar Atlas. (2026). *Interactive solar map* [Web map]. Global Solar Atlas. Retrieved February 1, 2026, from https://globalsolaratlas.info/map

[6] Hijriawan, M., Himawanto, D. A., Pambudi, N. A., & Arifin, Z. (2022). organic Rankine cycle (ORC) system in renewable and sustainable energy development: a review of the utilization and current conditions for small-scale application. In *Journal of Applied Engineering Science* (Vol. 20, Issue 3, pp. 957–970). Institute for Educational Research. https://doi.org/10.5937/jaes0-36319

[7] Hossin, K., & Mahkamov, K. (2015). *Performance evaluation for a 10 kW solar Organic Rankine Cycle power system to operate in the UK climate conditions.* In *Proceedings of The European Conference on Sustainability, Energy & the Environment (ECSEE2015)*. IAFOR. https://papers.iafor.org/submission15907/

[8] Ighalo, J. O., Sun, Z., Osaro, E., Zhou, C., Nnadozie, E. C., Chen, Z., Ashaolu, S. E., & Yap, P.-S. (2025). *Renewable energy with a step forward towards environmental sustainability.* In A green vision towards a renewable energy future (pp. 233–262). Springer. https://doi.org/10.1007/978-3-031-93760-6_8

[9] Kumar, A., Said, Z., & Bellos, E. (2021). An up-to-date review on evacuated tube solar collectors. In *Journal of Thermal Analysis and Calorimetry* (Vol. 145, Issue 6, pp. 2873–2889). Springer Science and Business Media B.V. https://doi.org/10.1007/s10973-020-09953-9

[10] Lago, T. G. S., Padilha, B. R. P., Teixeira, F. S., Lima, J. A., Marques, A. S., & Santos, C. M. S. (2025). Thermodynamic modelling of a power generation plant using solar concentrators assisted by organic Rankine cycle for João Pessoa city, Brazil. *Archives of Thermodynamics*, *46*(1), 123–134. https://doi.org/10.24425/ather.2025.154187

[11] Lee, Y. R., Liu, L. W., Chang, Y. Y., & Hsieh, J. C. (2017). Development and application of a 200 kW ORC generator system for energy recovery in chemical processes. *Energy Procedia*, *129*, 519–526. https://doi.org/10.1016/j.egypro.2017.09.176

[12] Markides, C. N., Bardow, A., de Paepe, M., de Servi, C., Groß, J., Haslam, A. J., Lecompte, S., Papadopoulos, A. I., Oyewunmi, O. A., Seferlis, P., Schilling, J., Linke, P., Tian, H., & Shu, G. (2025). Working fluid and system optimisation of organic Rankine cycles via computer-aided molecular design: A

review. In *Progress in Energy and Combustion Science* (Vol. 107). Elsevier Ltd. https://doi.org/10.1016/j.pecs.2024.101201

[13] Matuszewska, D. (2024). Solar Organic Rankine Cycle (ORC) Systems: A Review of Technologies, Parameters, and Applications. In Energies (Vol. 17, Issue 20). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/en17205106

[14] Permana, D. I., Rusirawan, D., & Farkas, I. (2022). A bibliometric analysis of the application of solar energy to the organic Rankine cycle. Heliyon, 8(4), e09220. https://doi.org/10.1016/j.heliyon.2022.e09220

[15] Permana, D. I., Rusirawan, D., & Farkas, I. (2023). The theoretical approach of the solar organic Rankine cycle integrated with phase change material for the Hungarian region. *Energy Science and Engineering*, *11*(12), 4429–4445. https://doi.org/10.1002/ese3.1589

[16] Pethurajan, V., Sivan, S., & Joy, G. C. (2018a). Issues, comparisons, turbine selections and applications – An overview in organic Rankine cycle. In Energy Conversion and Management (Vol. 166, pp. 474–488). Elsevier Ltd. https://doi.org/10.1016/j.enconman.2018.04.058

[17] Quoilin, S., Broek, M. Van Den, Declaye, S., Dewallef, P., & Lemort, V. (2013). Techno-economic survey of organic rankine cycle (ORC) systems. In Renewable and Sustainable Energy Reviews (Vol. 22, pp. 168–186). https://doi.org/10.1016/j.rser.2013.01.028

[18] Rahbar, K., Mahmoud, S., Al-Dadah, R. K., Moazami, N., & Mirhadizadeh, S. A. (2017). Review of organic Rankine cycle for small-scale applications. In Energy Conversion and Management (Vol. 134, pp. 135–155). Elsevier Ltd. https://doi.org/10.1016/j.enconman.2016.12.023

[19] Talamon, A., Hartmann, B. (2017). Optimal Utilization of Solar Energy Resources in Hungary. In: Azzopardi, B. (eds) Sustainable Development in Energy Systems. Springer, Cham. https://doi.org/10.1007/978-3-319-54808-1_9

[20] Thangavel, S., Verma, V., Tarodiya, R., & Kaliyaperumal, P. (2021). Comparative analysis and evaluation of different working fluids for the organic rankine cycle performance. *Materials Today: Proceedings*, *47*, 2580–2584. https://doi.org/10.1016/j.matpr.2021.05.064

[21] Tovar-Facio, J., Cansino-Loeza, B., & Ponce-Ortega, J. M. (2022). Management of renewable energy sources. In M. Martín (Ed.), *Sustainable design for renewable processes* (pp. 3–31). Elsevier. https://doi.org/10.1016/B978-0-12-824324-4.00004-4

[22] Wang, R., Jiang, L., Ma, Z., Gonzalez-Diaz, A., Wang, Y., & Roskilly, A. P. (2019). Comparative analysis of small-scale organic Rankine cycle systems for solar energy utilisation. *Energies*, *12*(5). https://doi.org/10.3390/en12050829

[23] Yadav, D. K., Malakar, S., Arora, V. K., & Sinhmar, N. (2024). Evacuated Tube Solar Collector-Based Drying System: Analytical Modeling, Influencing Factors, and Recent Progress in Drying of Agri-Commodities. *Food Engineering Reviews*, *16*(4), 567–594. https://doi.org/10.1007/s12393-024-09382-6

The evolution and future of ISO 14001: Strengthening environmental management for a sustainable future

Sahab ALKHOLI[1], Miklós DARÓCZI[2]
[1] *Doctoral Program of Mechanical Engineering*
Hungarian University of Agriculture and Life Sciences, MATE, Gödöllő
[2] *Department of Engineering Management, Institute of Technology,* MATE, Gödöllő

Abstract

ISO 14001 has been around since 1996, acting as the go-to standard for organizations trying to get a handle on their environmental footprint. It hasn't stood still, though. The 2004 update mostly smoothed integration with other systems, while the 2015 revision pushed harder on strategy—bringing in lifecycle thinking and putting environmental issues on the desks of senior managers. More recently, in 2024, the standard edged further toward climate-related risks and stakeholder expectations.
Another big revision is now on the horizon for 2026. Early signs point to stronger attention on climate change, greenhouse gas emissions, and circular economy principles, paired with structural changes like a shift to the new Harmonized Structure, a section on "Planning of Change," and tougher expectations around transparency and reporting. This article looks back at how ISO 14001 has developed, how it interacts with related standards, and what the next round of changes might mean. It also suggests some practical steps—such as rethinking risk documentation or improving stakeholder communication—that could make the transition a little smoother, though the effort required will probably vary depending on an organization's size and resources.

Keywords

ISO 14001, Environmental Management System, revision, changes

1. Introduction

In the world where protecting the environment is getting increased interest, many organizations are trying their best to protect the environment and reduce their ecological footprint and align with sustainable development goals. According to (ISO, What you need to know about ISO environmental management standards, 2025), the International Organization for Standards (ISO) invested a quite efforts into the environmental topics. For that, ISO

created a group of standards presenting a list of requirements for an Environmental Management System, these standards are named with ISO 14000 family and designed to help organizations develop and maintain effective Environmental Management Systems (EMS) (Magerholm Fet & Michelsen, 2023) & (Gasiorowski-Denis, 2013).

Among these, ISO 14001 stands out as the foundational framework for Environmental Management Systems (EMS), offering a set of requirements that help organizations reduce environmental impact, improve resource efficiency, and ensure compliance with applicable laws. As of 2025, more than 500,000 certifications to ISO 14001 within 180 countries all over the world, allowing more efficient usage of resources and reducing waste in the organizations and companies, helping them with improving their environmental performance. According to (ISO, Strategy 2030, 2025), ISO is trying to improve its environmental standards in a way that fits more environmental sides aligning with the 17 Sustainable Development Goals, which are the Global Agenda for 2030 have been set by the United Nations.

The initial publication of ISO 14001 was in 1996 and revised in 2004 and 2015. It has undergone significant revisions to remain relevant in an ever-changing environmental landscape. The 2004 revision improves the compatibility with other management systems (Fonseca, 2015) & (ISO, ISO 14001:2004, 2004). While the 2015 version introduced new concepts such as strategic environmental management, life cycle thinking, and enhanced stakeholder engagement (Fonseca, 2015). A further amendment in 2024 enriched the importance of climate change by requiring organizations to consider climate-related risks and stakeholder expectations (ISO, ISO 14001:2015/Amd 1:2024, 2024).

As ISO prepares for a major revision of ISO 14001 in 2026, the focus shifts to enhancing the standard's relevance in the context of climate change, greenhouse gas emissions, and circular economy principles (Hub, 2025) & (Austria, 2024). This article explores the development and structure of ISO 14001, the interrelated standards within the ISO 14000 family, and the anticipated changes that aim to align environmental management practices with the global sustainability agenda.

However, the article proposed several recommended actions in terms of keeping the organization updated with the requirements of ISO 14001, and in terms of reacting and preparing to align and fulfill these upcoming requirements.

2. Experimantal

2.1. ISO 14000 family background

With the increase in environmental awareness and regulatory pressure, it is expected that organizations will demonstrate a proactive approach to environmental responsibility. The International Organization for Standardization (ISO) developed a set of standards, called ISO 14000 family of standards, to

assist in managing the environmental responsibilities within the organizations. These standards have been initiated in 1996, and aiming to ensure the compliance of regulatory laws, minimizing the negative environmental impacts, in addition to continuing improving the environmental performance (Magerholm Fet & Michelsen, 2023) & (Gasiorowski-Denis, 2013). Where **Figure (1)** presents the timeline for publishing these standards.

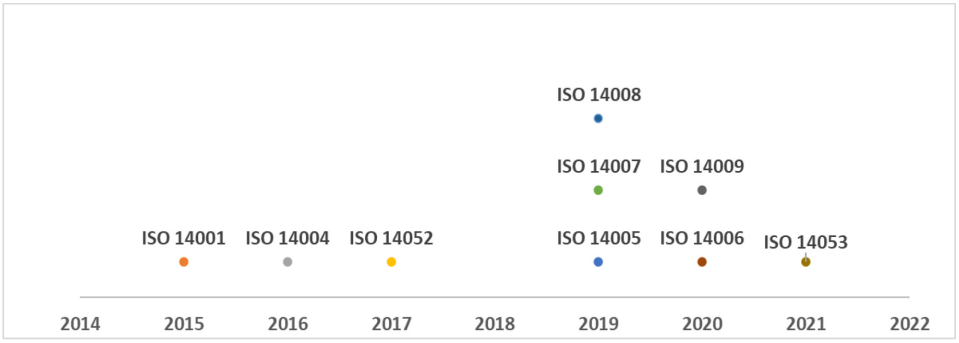


Figure (1): The timeline of ISO standards publications. (Self-made)

2.2. Interrelated standards in ISO 14000 family

2.2.1. ISO 14001

The main standards and most popular one is ISO 14001 (Environmental Management Systems — Requirements with Guidance for Use). The standard which is used now is published in 2015, and provides a structure safeguard the environment and stay in balance with socio-economic needs by responding to changing environmental conditions. ISO 14001 provides the requirements to achieve the intended outcomes which have been set by an organisation for its Environmental Management System (ISO, ISO 14001:2015, 2015).

2.2.2. ISO 14004

One year after publishing ISO 14001, the international standard ISO 14004 has been published, named with Environmental Management Systems — General Guidelines on Implementation. The standard gives the general guidelines on the standard's implementation, establishing, maintaining, and continually improving the system to support the environmental management (ISO, ISO 14004:2016, 2016).

2.2.3. ISO 14005

Environmental Management Systems — Guidelines for a Flexible Approach to

Phased Implementation or ISO 14005 standard which served in 2019; helps organizations -especially small and medium-sized enterprises (SMEs)- gradually implement an Environmental Management System (EMS) that aligns with ISO 14001 requirements, using a flexible, phased approach. In addition to address environmental risks such as pollution, resource scarcity, and climate change (ISO, ISO 14005:2019, 2019). ISO 14005 makes EMS implementation more accessible and manageable, especially for organizations with limited resources, while supporting long-term sustainability and continual improvement.

2.2.4. ISO 14007

In the same year the ISO 14007, which is named Environmental Management — Guidelines for Determining Environmental Costs and Benefits has been served. The main aim of it is providing guidelines for organizations to identify, assess, and express the environmental costs and benefits related to their environmental aspects and dependencies on nature, such as resource use and ecosystem services. In addition to promoting an anthropocentric approach, focusing on human well-being and the total economic value of nature, including both use and non-use values (ISO, ISO 14007:2019, 2019).

2.2.5. ISO 14008

Also the International Organization for Standards (ISO) published ISO 14008 in 2019, calling it by Monetary Valuation of Environmental Impacts and Related Environmental Aspects, gives a methodological framework for the monetary valuation of environmental impacts and aspects such as effects on human health, ecosystems, and the built environment, helping organizations assign economic value to environmental issues such as emissions or natural resource usage, in a structured and consistent way (ISO, ISO 14008:2019, 2019).

2.2.6. ISO 14006

In 2020 the Environmental Management Systems — Guidelines for Incorporating Eco-design or in other way ISO 14006 standard has been distributed. The standard provides organizations with the guidelines in establishing, documenting, implementing, maintaining and continually improve the eco-design management as part of the Environmental Management System. This document can be applicable to any organisation despite its type, size, products, or services provided, or even having EMS implemented already or other management systems (ISO, ISO 14006:2020, 2020).

2.2.7. ISO 14009

At the same time, ISO distributed ISO 14009, named with Environmental Management Systems — Guidelines for Incorporating Material Circulation in Ddesign and Development. It guides organizations to establish and improve

material circulation during the design and development of products, using an Environmental Management System (EMS) framework, especially one based on ISO 14001. In addition to support resource efficiency and circular economy principles (ISO, ISO 14009:2020, 2020).

In terms of material flow cost accounting ISO published 3 documents as the following:

2.2.8. ISO 14051

In 2011 and starting with ISO 14051, known with Environmental Management — Material Flow Cost Accounting — General Framework. It gives a general framework for Material Flow Cost Accounting (MFCA), it is a method can be used to trace and quantify material and energy flows in physical and monetary terms within an organization. Aiming to improving organizations' resources efficiency and reducing environmental impacts (ISO, ISO 14051:2011, 2011).

2.2.9. ISO 14052

Based on ISO 14051, the international organisation developed ISO 14052 (Environmental Management — Material Flow Cost Accounting — Guidance for Practical Implementation in a Supply Chain) to guide organizations with applying MFCA in supply chains. For that they extended the use of MFCA beyond an individual organization to promote collaborative resource efficiency across upstream and downstream partners. The standard was published in 2017 (ISO, ISO 14052:2017, 2017).

2.2.10. ISO 14053

While in 2021, the standards of Environmental Management — Material Flow Cost Accounting — Guidance for Phased Implementation in Organizations, or in another word ISO 14053; the standard helps organizations and especially small and medium-sized enterprises (SMEs) improve material efficiency and environmental performance during the implementation (ISO, ISO 14053:2021, 2021).

Figure (2) shows the connections and how the previous standards work together (ISO C. , 2025).

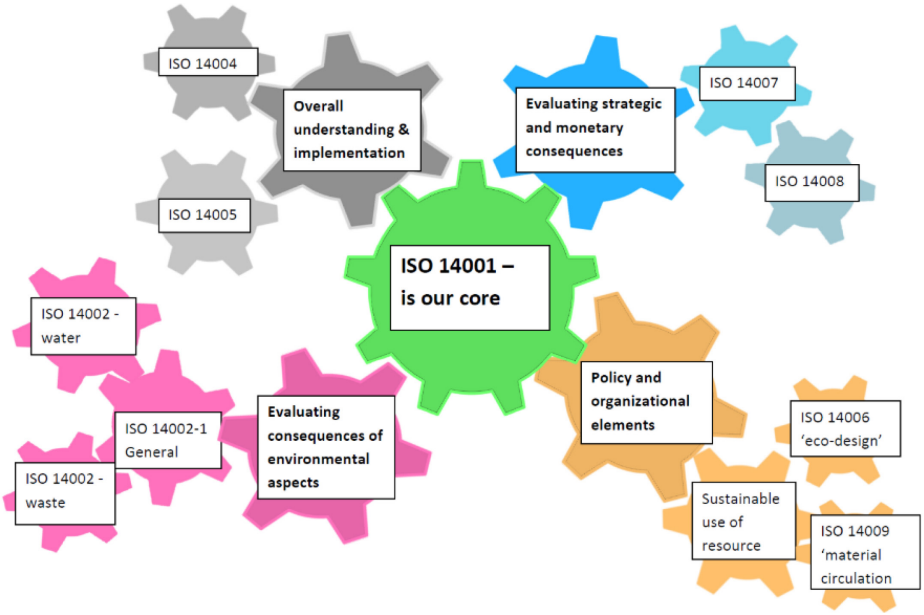


Figure (2): The connection between ISO14000 family standards (ISO C. , 2025).

2.3. ISO Standard Lifecycle

To have a standard, the International Standard Organization specified a lifecycle, Figure (3). Each standard must pass by the followed stages in this lifecycle:

1. Preliminary stage:
   In this stage an idea for a new standard is received, reviewed, and either abandoned or moved forward.
2. Proposal stage:
   Here, the new project proposal (NP) is registered, balloted (voted on by members), and either rejected or approved.
3. Preparatory stage:
   A Working Draft (WD) is prepared by a working group of experts. Comments are gathered, revisions made, and eventually the draft can be approved to move forward.
4. Committee stage:
   A Committee Draft (CD) is circulated to ISO members for comments. This is where consensus-building happens. The draft may be revised several times before approval to advance.
5. Enquiry stage:

The Draft International Standard (DIS) is circulated to all ISO members for a 12-week ballot. It may be approved, sent back for revision, or undergo another DIS ballot.

6. Approval stage:
   The Final Draft International Standard (FDIS) is issued for an 8-week formal vote. If approved, it moves to publication.
7. Publication stage:
   The International Standard is published and made official.
8. Review stage:
   Every standard is reviewed periodically (usually every 5 years). It may be confirmed, revised, or proposed for withdrawal.
9. Withdrawal stage:
   If outdated or no longer needed, the standard can be withdrawn after a ballot.

2.4. ISO 14001 History

The widest diffuse environmental standard between ISO 14000 family is the ISO 14001 Environmental Management Standard. ISO 14001 has evolved significantly since its first publication in 1996. According to (Lazarte, 2015), all ISO standards undergo a review every three to five years ensuring that they are relevant and up to date. In 2004 and 2015, the standard ISO 14001 underwent revisions, and with each resumption addressing emerging environmental concerns and business needs (Fonseca, 2015). Creating the foundational model for ISO's Environmental Management Standards, ISO starts with specifying the requirements that enable organizations to formulate a policy and objectives for an Environmental Management System taking into consideration applicable laws in addition to information about significant environmental impacts (ISO, ISO 14001:1996, 1996).

In 2004, the version emphasis the first changes in the standards exceeding the environmental aspects that the organization determines and can control to the ones which can influence. In addition to improving the compatibility with other management systems such as quality standards ISO 9001 (Fonseca, 2015) & (ISO, ISO 14001:2004, 2004).

While the used version nowadays is the one of 2015. The standard compasses various aspects, starting with resources usage and waste management reaching the environmental performance measuring and monitoring (ISO, ISO 14001:2015, 2015). And as Fonseca mentioned in her article (Fonseca, 2015), ISO 14001:2015 affirmed sustainability, context of the organization, and stakeholder engagement, in addition to considering life cycle perspective of its products, services, and activities. The responsible committee of modifying ISO

14001 published the key changes of the version since 2004 edition (ISO/TC 207/SC 1, 2025) & (da Fonseca, 2015):

1. Strategic Environmental Management
   ISO 14001:2015 ensures the integration of Environmental Management into strategic planning of the organization.
2. Leadership
   The standard assigns specific responsibilities to Top Management in order to develop Environmental Management within the organization.
3. Protecting the environment
   The commitment to proactive initiatives to protect the environment has been expanded by the version of 2015.
4. Environmental performance
   Shifting the focus in the organization from merely improving the management system to enhancing actual environmental performance.
5. Lifecycle perspective
   ISO 14001:2015 prompts organizations to look after the environmental impacts throughout the entire lifecycle of products and services.
6. Outsourced processes
   Encourage organizations to consider outsourced processes and how can be controlled or influenced by environmental management.
7. Communication
   The standard evaluates setting a communication strategy internal and external the organisation equally.
8. Documentation
   Documented information is a push from the standard to reflect the evolution of digital and cloud-based systems.

Figure (3): The lifecycle of an ISO standard (Self-made based on (ISO, International harmonized stage codes, 2025))

The previous changes aimed to ensure the standard remained relevant and effective in addressing contemporary environmental challenges. And for that reason, ISO published an amendment to ISO 14001:2015 in 2024 considering the importance of climate changes. Requesting organizations to determine whether climate change is a relevant issue in the organization, in addition to highlighting the possibility of interested parties having requirements related to climate change as well (ISO, ISO 14001:2015/Amd 1:2024, 2024).
According to (Hikichi, Salgado, & Beijo, 2016), these revisions have contributed to the standard's global implementation, with the number of ISO 14001 certifications persistently increasing across countries and economic sectors, especially in the US. Despite this growth, ISO 14001 certifications still fall behind ISO 9001, however the gap is narrowing (Hikichi, Salgado, & Beijo, 2016). Studies have shown that implementing ISO 14001 can lead to improved environmental performance and business benefits, supporting its role in promoting sustainability (Vries, Bayramoglu, & Wiele, 2012).

3. Results and discussion.

3.1. Expected changes in ISO 14001 for 2026

Looking ahead and driven by several strategic and practical factors reflecting the development of environmental management, further revisions are anticipated by the beginning of 2026. Where in the fall of 2023 the development of the new ISO 14001 standard began, and the draft international standard (DIS) has been published in June 2025, while the final draft international standard (FDIS) is expected to be published in October 2025; with the final version anticipated to be published in January 2026 (SGS, 2025) & (DNV, 2025). Organizations are increasingly viewing sustainability not just as a regulatory requirement, but as a source of competitive advantage, prompting the need for more flexible, data-driven frameworks that emphasize measurable and impactful outcomes. To ensure the standard remains relevant and effective, ISO has gathered inclusive feedback from global businesses, governments, and stakeholders, confirming the importance of proactive climate action and sustainable supply chain management. Moreover, insights from ISO technical committees and study groups (concentrating on risk management, digital innovation, and climate-related challenges) have played a crucial role in shaping the incoming updates. However, these drivers aim to align ISO 14001 with contemporary sustainability priorities while maintaining global consistency and long-term relevance.

Accordingly, the new revision is expected to place greater emphasis on climate change adaptation, greenhouse gas management, and the integration of environmental considerations into organizational purpose and reporting practices (Hub, 2025). Based on (Hub, 2025) & (Austria, 2024) publications' feedback, the most expected changes are the followings:

1. Switching from the original High-Level Structure (HLS) to a new Harmonized Structure, with enhancing alignment with other ISO management system standards.
2. Including a new section 6.3 "Planning of Change", which reflects the growing importance of managing environmental transitions proactively.
3. Replacing the term "outsourcing" with "externally provided" to better reflect modern business practices.
4. In accordance with the ISO London Declaration (September 2021), the life cycle perspective, climate change adaptation, and the role of greenhouse gases will be more conspicuously featured.
5. Environmental aspects will be expanded as technical topics by the ISO 14002-1 series, providing deeper, sector-specific guidance.
6. Placing a stronger emphasis on reporting and transparency, responding to increased demand for demonstrable environmental performance.
7. To help with implementation, the standard will introduce clarifications and practical explanations in appendices or notes.
8. There is a discussion among embedding the concept of an organization's "purpose", but still not emphasized.

By these changes, ISO aims to ensure that ISO 14001 remains practical, future-oriented, and aligned with the global sustainability agenda.

In **Table (1)** we are presenting the main points and focuses of ISO 14001:2004, and ISO 14001:2015, in addition to the expected focuses of ISO 14001:2026.

ISO 14001:2004	**ISO 14001:2015**	**ISO 14001:2026**
Improved Compatibility with other management systems.	Strategic Environmental Management.	Harmonized Structure.
Expanded Environmental Scope.	Leadership and assigning specific responsibilities.	Planning for Change.
	Commitment to protecting the environment.	Sustainability Focus (life cycle perspective, climate change adaptation, and the role of greenhouse gases).
	Enhancing actual environmental performance.	Expanded environmental aspects.
	Lifecycle perspective.	Reporting and transparency focus.
	Consider outsourced processes.	Clarifications and practical explanations for the implementation.
	Communication strategy.	
	Documented information.	

*Table (1)***:** The main points of ISO 14001 in 2004, 2015, and 2026 versions. (Self-made)

3.2. How should organizations react

Considering what has been discussed in this article and based on site observations, there are several actions that are recommended:

1. In terms of keeping the organization updated with the requirements of ISO 14001, the organization should:
   - Wait and stay tuned to the upcoming version of ISO 14001 and become familiar with ISO management system standards and the expected concepts like "Planning of Change".
   - Be prepared for more details and explanations regarding topics like life cycle perspective, climate change adaptation, and the role of greenhouse gases.
   - Regularly, check www.iso,org since periodic updates will be made available by ISO/TC207/SC1.
2. In terms of reacting and preparing to align and fulfill the upcoming requirements, the following are recommended:
   - For the organizations currently certified to ISO 14001:2015, they should review draft changes in the 2025 revision to ensure timely compliance and alignment with future requirements.
   - Highlighting climate change and lifecycle thinking should motivate organizations to begin assessing their environmental impacts and dependencies more holistically.
   - Give great importance to reporting and enhancing the environmental data management systems to ensure transparency and traceability of performance metrics.
   - Encouraging organizations to adopt relevant ISO 14000 family standards to maximize the effectiveness of their Environmental Management System.

For organizations which are planning to establish an Environmental Management System based on ISO 14001, it would be highly recommended to check and study the latest version of the standard "ISO 14001:2015" (ISO, ISO 14001:2015, 2015). Taking the recommendations into consideration will strengthen both environmental and operational performance of the organizations.

What worth to mention is that for certified organizations, there will be around three years as a transition period from the publication date of the standard. So, in case the new standard is published in January 2026, the organizations will have time until January 2029 to transition.

4. Conclusion

- In this article, we examined the historical development, structure, and ongoing evolution of the ISO 14001 standard within the ISO 14000 family.
- There were light changes in 2004, by improving the standard's harmony with other management systems.
- The 2015 version marked a considerable shift in environmental management by integrating strategic planning, stakeholder engagement, lifecycle thinking, and enhanced leadership requirements.
- With the 2024 amendment affirming climate change, and a comprehensive revision set for 2025, ISO 14001 is adapting to the urgent environmental imperatives of our time.
- The anticipated changes include transitioning to a Harmonized Structure, the introduction of a new section on "Planning of Change", enhanced reporting obligations, deeper consideration of climate-related risks, and the modernization of terminology.
- Integral standards within the ISO 14000 family, such as ISO 14004, 14005, 14006, and the MFCA-related standards (ISO 14051–14053), give essential support for specific aspects of environmental and resource efficiency management.
- In general, these developments place ISO 14001 as a central tool for achieving sustainable development across sectors.

Acknowledgements

This research was supported and supervised by an agricultural manufacturing company.

References

[1]Austria, Q. (2024, May 28). *ISO 14001 is being revised - what can we expect?* Retrieved from Quality Austria: https://www.qualityaustria.com/en/news/iso-14001-is-being-revised/?utm

[2]da Fonseca, L. M. (2015). ISO 14001:2015: An Improved Tool for Sustainability. *Journal of Industrial Engineering and Management*, 14. doi:http://dx.doi.org/10.3926/jiem.1298

[3]DNV. (2025, February 03). *The ISO 14001 draft update released: What to expect in the new version*. Retrieved from DNV: https://www.dnv.com/news/2025/iso-14001-standard-draft-update-

released/

[4]Fonseca, L. M. (2015). ISO 14001: 2015: An improved tool for sustainability. *Journal of Industrial Engineering and Management*, 37-50. Retrieved from https://doi.org/10.3926/JIEM.1298

[5]Gasiorowski-Denis, E. (2013, April 29). *ISO 9000 and ISO 14000: two flagship collections now available online*. Retrieved from ISO: https://www.iso.org/news/2013/04/Ref1732.html

[6]Hikichi, S. E., Salgado, E. G., & Beijo, L. A. (2016). Characterization of dissemination of ISO 14001 in countries and economic sectors in the Americas. *Journal of Environmental Planning and Management*, 1554-1574.

[7]Hub, B. (2025, June 2). *Latest Changes in ISO 14001: Understanding the 2025 Revision*. Retrieved from BPR Hub: https://www.bprhub.com/blogs/iso-14001-changes-2025-revision#:~:text=The%20standard%20will%20transition%20from%20the%20original%20High-Level,currently%20under%20revision,%20with%20new%20amendments%20in%20development.

[8]ISO. (1996). *ISO 14001:1996*. Retrieved from ISO: https://www.iso.org/standard/23142.html

[9]ISO. (2004). *ISO 14001:2004*. Retrieved from ISO: https://www.iso.org/standard/31807.html

[10] ISO. (2011). *ISO 14051:2011.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14051:ed-1:v1:en

[11] ISO. (2015, Mar). *ISO 14001:2015.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14001:ed-3:v1:en

[12] ISO. (2015). *ISO 14001:2015*. Retrieved from ISO: https://www.iso.org/standard/60857.html

[13] ISO. (2016, Mar). *ISO 14004:2016.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14004:ed-3:v1:en

[14] ISO. (2017). *ISO 14052:2017.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14052:ed-1:v1:en

[15] ISO. (2019). *ISO 14005:2019.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14005:ed-2:v1:en

[16] ISO. (2019). *ISO 14007:2019.* Retrieved from ISO: https://www.iso.org/standard/70139.html?browse=tc

[17] ISO. (2019). *ISO 14008:2019.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:43243:en

[18] ISO. (2020). *ISO 14006:2020.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14006:ed-2:v1:en

[19] ISO. (2020). *ISO 14009:2020.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14009:ed-1:v1:en

[20] ISO. (2021). *ISO 14053:2021.* Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14053:ed-1:v1:en

[21] ISO. (2024). *ISO 14001:2015/Amd 1:2024*. Retrieved from ISO: https://www.iso.org/obp/ui/en/#iso:std:iso:14001:ed-3:v1:amd:1:v1:en

[22] ISO. (2025, September 22). *International harmonized stage codes*. Retrieved from ISO: https://www.iso.org/stage-codes.html#60_60

[23] ISO. (2025, May 12). *Strategy 2030*. Retrieved from ISO: https://www.iso.org/strategy2030.html

[24] ISO. (2025, May 12). *What you need to know about ISO environmental management standards*. Retrieved from ISO: https://www.iso.org/standards/popular/iso-14000-family#:~:text=There%20are%20more%20than%20500%2C000%20certifications%20to%20ISO,a%20wide%20range%20of%20benefits%20of%20ISO%2014001.

[25] ISO, C. (2025, May 31). *ISO/TC 207/SC 1.* Retrieved from Committee ISO: https://committee.iso.org/home/tc207sc1

[26] ISO/TC 207/SC 1. (2025, June 1). *ISO 14001:2015 - main changes since 2004 edition*. Retrieved from ISO/TC 207/SC 1: https://committee.iso.org/sites/tc207sc1/home/projects/published/iso-14001---environmental-manage/main-changes.html?utm

[27] Lazarte, M. (2015, January 22). *ISO 14001 revision proceeds to final stage*. Retrieved from ISO: https://www.iso.org/news/2015/01/Ref1928.html

[28] Magerholm Fet, A., & Michelsen, O. (2023). Environmental Management Systems. In A. Magerholm Fet, *Business Transitions: A Path to Sustainability* (pp. 67-76). Springer. doi:https://doi.org/10.1007/978-3-031-22245-0

[29] SGS. (2025, September 18). *ISO 14001:2026 – Key Updates and Transition Guidance*. Retrieved from SGS: https://www.sgs.com/en/showcases/iso-14001-2026-key-updates-and-transition-guidance

[30] Vries, H. J., Bayramoglu, D., & Wiele, T. V. (2012). Business and environmental impact of ISO 14001. *International Journal of Quality & Reliability Management*, 425-435.

Colour measuring of the thickness of solder resist layer to improve the manufacturability of electronic devices

Bernadett VARGA[1], Róbert KERESZTES[2], János BUZÁS[3]
[1] *Doctoral Program of Mechanical Engineering, Hungarian University of Agriculture and Life sciences,* MATE, Gödöllő
[2] *Department of Materials Science and Engineering Processes, Institute of Technology, MATE Gödöllő*
[3] *Department of Mechatronics, Institute of Technology, MATE Gödöllő*

Abstract

In the process of the surface mounting of electronic panels, the solder resist layer, its material and thickness largely determine the manufacturability of Printed Circuit Board (PCB). Soldering errors can also occur during the manufacturing of the electronic unit, and during the use of the electrical device, an inadequate thickness and quality of the solder resist layer can cause the electronic panel to fail. As the first step in surface mounting, laser marking of electronic circuits is usually used for product traceability. Electronics manufacturers most commonly use data matrix codes (DMC) for product identification, which are created using a laser beam in the solder mask layer of the printed circuit board. Different materials do not absorb laser beams of the same wavelength equally. The result of the marking process is greatly influenced by the optical properties of the solder resist, which have not been studied so far and depend on the material composition. In addition, solder resist coatings also have insulation properties due to their good electrical and dielectric properties. Anti-solder mask plays a very important role in preventing electron corrosion of copper wiring running on the surface. It is important to mention determining the exact thickness of layer measurement in reducing the effect of Electrochemical Migration (ECM) on the surface of printed circuit boards (PCBs) and in eliminating it. Galvanic Corrosion or bimetallic corrosion is when dissimilar metals (such as copper and aluminium) encounter each other, forming a primary cell in the electrolyte. The more active metal (in this case aluminium) acts as an anode, which accelerates the corrosion process. The following article reviews the literature on methods for measuring the thickness of the solder resist layer. Also, how the change in the effect of the layer thickness affects the quality of laser marking and what effect it has in the context of the development of Electrochemical Migration and Galvanic Corrosion.

Keywords

Printed Circuit Board (PCB), Electrochemical Migration (ECM), Galvanic Corrosion (GC), Data Matrix Codes (DMC), Solder Resist Layer (SRL), Laser Marking (LM), Surface Mounting Technology (SMT)

1. Introduction

The purpose of the solder mask (SM) on the PCB is to prevent the solder from wetting and thus the formation of solder bridges, further prevent oxidation, reduce contamination, inhibit whisker growth, etc. (Brewin and Willis, 2007), (Chowdhury et al., 2018). According to He et al. (2010) and Bušek et al. (2018), SMs should also suppress ECM. Both works investigated the effect of the presence of SM on ECM and showed that its presence reduces the risk of ECM in both cases.

Susceptibility to ECM is also influenced by the interface between the electrodes, which can be coated with a SM.

The various PCBs SM types differ in how they are applied, their composition, and of the price.

PCB are sufficient parts of electronic and electronics appliances. Mass production of these devices requires identification of these products both during technological steps (Traceability) and the services. The most effective way for marking these plates is to form data matrix code (DMC) on the surface and/or in-depth profiling of SMs up to the metallization. This is the first step in the surface mounted technology.

Both material composition and layer thickness determine the proper manufacturability, laser marking capability of SM and quality use of PCBs, which includes avoiding the development of Electrochemical Migration and Galvanic Corrosion.

The colouring measure system was developed for the PCB's identification, namely measuring the colouring in XYZ-colour system. The XYZ values of SM layers are measured, and software was developed for calculating automatically the colour coordinates x and y and the brightness value L.

In the course of the present work different PCBs were analysed originating from more manufacturers. The data matrix was formed using different lasers with different power and duration.

Furthermore, the application of the colorimetric method for measuring the solder resist varnish layer is presented. The importance of measuring the solder layer during surface mounting for better product traceability, to avoid soldering errors, and to prevent the development of possible electrical corrosion during the final use of electronic devices.

2. Material and methods

The following subsections present laser marking solutions for solder resist coatings. A detailed description of the factors influencing the manufacturability of solder resist coatings of different quality and thickness from manufacturers.

2.1. Laser marking methos

Two laser marking methods are known: either by discolouring SM, the laser beam penetrates the layer only to a depth of a few microns, or by completely removing SM down to the surface of the copper layer in accordance with the code design.

In both cases, the choice of exact laser processing parameters can only be determined by preliminary testing. However, it is often a problem during production if the raw material, in this case the circuit board, changes. The change should be understood as a change in the thickness and/or material of the solder resist layer.

Since the goal during laser marking is not only the proper removal of the coating layer, but also the subsequent work process, the printing of the solder is largely determined. Because if this layer is outside the normal range, either very below or above the expected mask layer thickness, soldering errors can occur, such as short circuits or smearing of the solder.

In the industrial praxis two laser marking methods are used:

1. Discolouration of the SM, the laser beam penetrates the layer only down to a depth of several μm's (Fig. 1),
2. The SM is completely removed up to the copper layer, according to the pattern of Data Matrix Code (Fig. 2).

Figure 1. Rectangle DMC: it is made by means of a CO_2 laser source. The dimensions of DMC are 3x8 mm (rows x columns: 12 x 26). The thickness of SM was 21 μm.

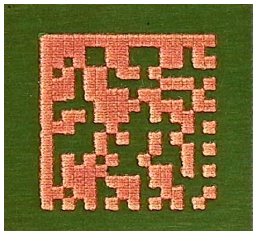


Figure 2. Square DMC: it is made by Nd:YAG laser source. The dimensions of DMC are of 4 x 4 mm (rows x columns: 16 x 16) The thickness of SM layer was 36 μm.

2.2. The role of SM

The SM aids inspection provides protection for conductors and prevents visual fatigue during hand assembly. The various PCB SM types differ in how they are applied, their composition, and the price.
In addition, solder resist coatings also have insulating properties due to their good electrical and dielectric properties.

The PCB: a basic electronic component made of 1 to 16 layers of bakelite or fiberglass resin, produced in small or large-scale production, but also in individual production. The insulating board carries a conductive layer a few microns thick on its surface.

The conductive layer is usually copper, often tin-plated. Many different coated layer systems are used in the production of printed circuit boards. DIN 40 804 standard defines the most important quality requirements for solder resist coatings.

Corrosion problems are particularly common among electronic boards and components, especially in high humidity and dirt conditions.
Exposure to humidity can trigger a variety of corrosion-related failures. The formation of a water film of variable thickness on the surface of a PCB depends on several factors, including the surface and structure of the PCB, the presence of hygroscopic residues, and fluctuations in ambient humidity or temperature. The presence of a water film is a significant contributor to several operational failures, including corrosion modes such as leakage current (LC), short circuits, and ECM (Piotrowska et al., 2018).

2.3. Laser marking problems

During SMT production, it often occurs that the SM becomes too thick or thinner than that specified (possibly beyond tolerance limits). The parameters set at the time of previous testing are unable to handle each layer thickness in an appropriate quality. If the layer is too thin, the applied laser beam melts the copper sheet after removal of the SM with the consequence that the code format is distorted. If the SM is too thick, it cannot be removed and colouring, respectively, at its full depth by using the preset laser processing parameters. The problem is that, if the layer thickness varies frequently, or the PCB's are made by different manufacturers, nevertheless, with the same pattern, it is often necessary to apply marking with every setting on the test area of the PCB. All these involve the cycle time to be increased which is disadvantageous in respect of production (Cucu et al., 2008), (Steen and Mazumder, 2010).

2.4. The operation of the colour measuring device

With the application of the colour measuring device, SM layers can be separated by manufacturers by measuring the colour coordinates and colour depth of the layer, and the change in layer thickness can also be measured by changing the

colour depth. By quickly and non-destructively checking the thickness of the coating mask, the number of rejects can be reduced or even eliminated. Since the goal during laser marking is not only the proper removal of the coating layer, but also the subsequent work process, the printing of the solder is largely determined. Because if this layer is outside the normal range, either very below or above the expected mask layer thickness, soldering errors can occur, such as short circuits or smearing of the solder.

Measuring material composition is necessary because in the case of the colouring method, the ratio of chemical elements in the solder resist layer and the particle size of the compounds formed by them are decisive. The material composition and the particle size and distribution of the solder resist layer are also decisive parameters in layer removal with laser marking.

The colouring equipment works in accordance with XYZ-colour system. In Fig. 3 can be seen the CIE XYZ colour space diagram, the tristimulus values of a colour are the amounts of three primary colours in a three-component additive colour model needed to match that test colour. The tristimulus values are not the *S* (short wavelength), *M* (middle wavelength), and *L* (long wavelength) responses of the human eye, but rather a set of tristimulus values called *X*, *Y*, and *Z*, which are roughly red, green and blue. The *X*, *Y*, *Z* values are a sort of 'derived' parameters from the red, green, blue colours (Hunt and Pointer, 2011).

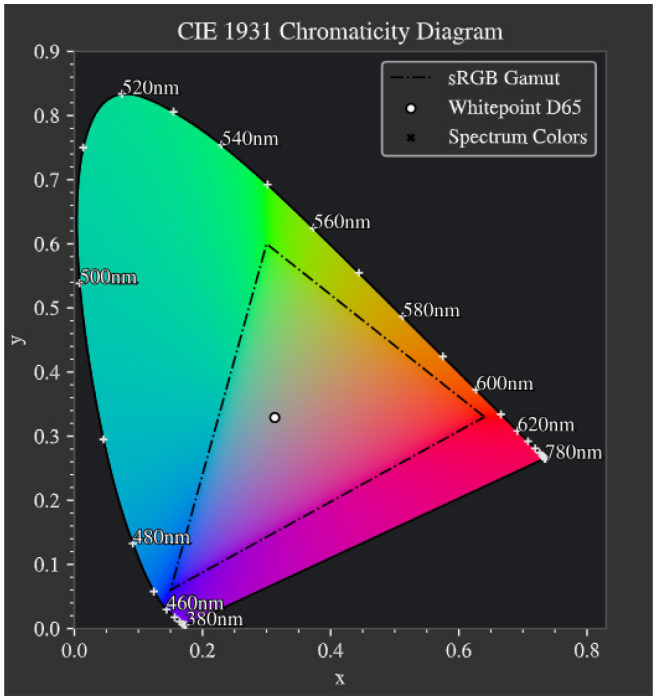


Figure 3. CIE XYZ colour space
https://drocheam.github.io/optrace/_images/chroma_1931.svg

The laser software is suitable to receive the measured values of the colouring laser machines automatically. The different colours of SM layer are categorized according to the supplier or the thickness of SM layer. According to the suppliers' PCB are separated is very easy, because the colours and the identify number of PCB are different. If the thickness of SM layers is changed less than with ±1 μm, the layer can be separated by the colour measuring system.

2.5. Possible causes of EMC

Electronic manufacturing technology is constantly evolving, with miniaturization as a key goal. The increased density of surface-mounted components in a smaller area, coupled with the reduction in the distance between the conductive traces or leads of the components, increases the risk of electrochemical migration (ECM) (Minzari et al., 2011), (Illés et al., 2014). Exposure to humid, polluted environments also increases the risk of ECM, especially when the printed circuit board (PCB) is in direct contact with the environment (Zhong, 2020).

ECM is a corrosion phenomenon that occurs in a high humidity environment and due to the applied bias voltage. In this case, it can lead to a decrease in the surface insulation resistance between adjacent electrodes, or even a short circuit (He, 2011), (Liao et al., 2020).

The presence of ECM can reduce the life of electronic components and can even lead to partial or permanent failure (Illés et al., 2022).

In a previous study the results indicated that the higher roughness of the solder mask significantly contributes to ECM inhibition through the creation of a mechanical barrier for the dendrites.

Furthermore, lower ECM susceptibility was also observed for copper-defined pads, where a similar effect is presumed. The use of a rough solder mask and a copper-defined pad design is recommended if the PCB is to be washed from flux residues after the soldering process (Klimtová et al., 2024).

2.6. Possible causes of types of PCB corrosion

Galvanic PCB corrosion occurs when dissimilar metals, such as gold and copper, come into contact in the presence of an electrolyte like moisture. To identify galvanic corrosion visually, look for **erosion or severe pitting in one metal.** Corrosion of a variety of types can be identified by examining the performance of circuit boards.

Atmospheric Corrosion: occurs due to prolonged exposure to oxygen and moisture in the air. Over time, oxidation leads to the formation of copper oxides, which degrade conductivity.

Galvanic Corrosion: When two dissimilar metals come into contact in the presence of an electrolyte, an electrochemical reaction occurs, causing one metal to corrode faster than the other.

Electrolytic Corrosion: an external power source causes metal migration,

leading to the formation of conductive dendrites that can create unintended electrical connections and shorts.
Fretting Corrosion: This type results from repeated mechanical motion, such as vibrations, which gradually remove protective layers and expose the underlying metal to oxidation.
Pitting Corrosion: appears as small holes in the metal surface, weakening the PCB structure over time.
Crevice Corrosion: Occurs in confined spaces where contaminants such as flux residues accumulate, creating a localized reaction that corrodes the metal.
Filiform Corrosion: moisture gets trapped beneath protective coatings, leading to thread-like corrosion paths along PCB traces (Khanna, 2023).5. Conclusions

3. Results and discussion

3.1. Surface of SM layers

In the Fig. 4 series, there are 3 diverse materials of the SM layer from different manufacturers. The thickness of the SM layers changes between 39 μm to 42 μm. One processing laser parameters (marking velocity, laser output power, resolution, and delays) was used for the removal solution of the SM. The laser marking print growth parameters of symbol were different, therefore the reading quality of DMC in diverse qualification. The colour of SM layers from different manufacturers can be featly distinguished by eye.

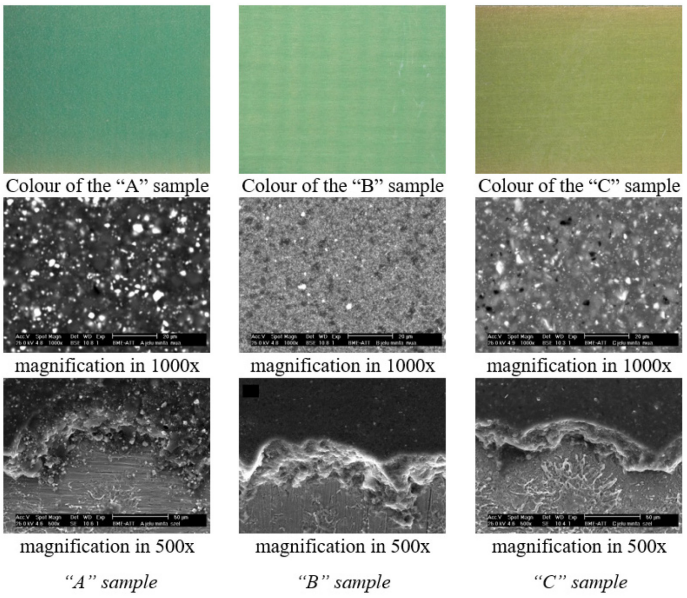


Figure 4. Solder resist coating samples from different manufacturers

3.2. Colour measuring of SM

The XYZ values of SM layers are measured, the software calculates automatically the x and y colour coordinates and L brightness values. These parameters appeared in a coordinate system. According to the values the PCB's can be separated quite simply.

The advantage of this software, it sends values to laser software. If the laser software is programmed with these values, the correct processing laser parameters can be run automatically.

In Fig. 5 the colour results of "A", "B" and "C" samples can be seen in the colour measure software. On these samples not the changing of the layers' thickness is caused the bad marking quality. In this case the bad quality is caused by different suppliers which used different materials in SM layers.

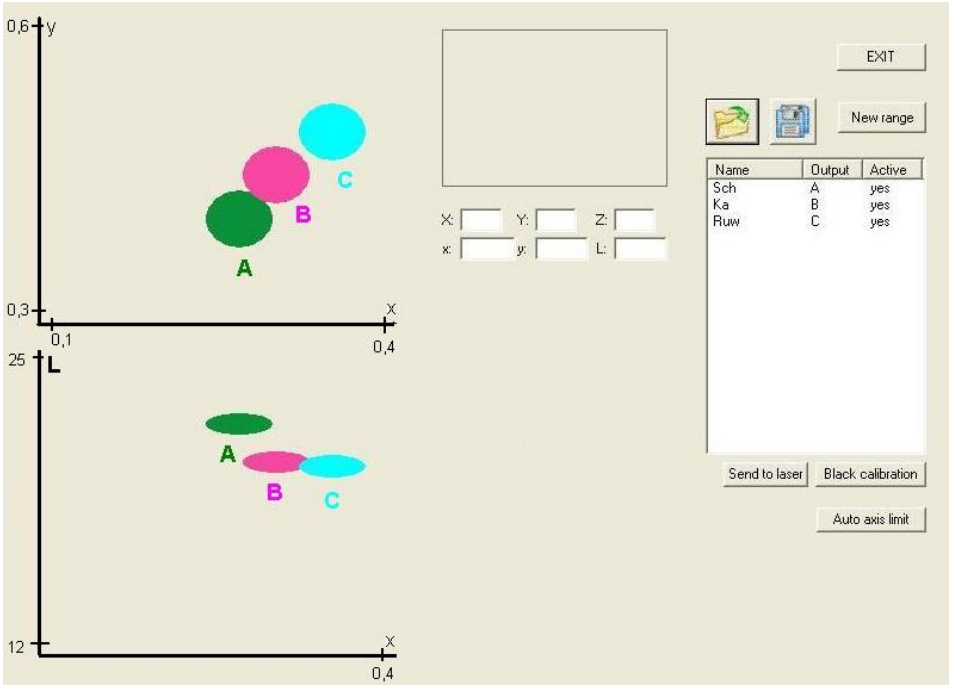


Figure 5. Colouring measure software

In Fig. 6 the colour values x, y, L of "A", "B" and "C" samples can be seen.

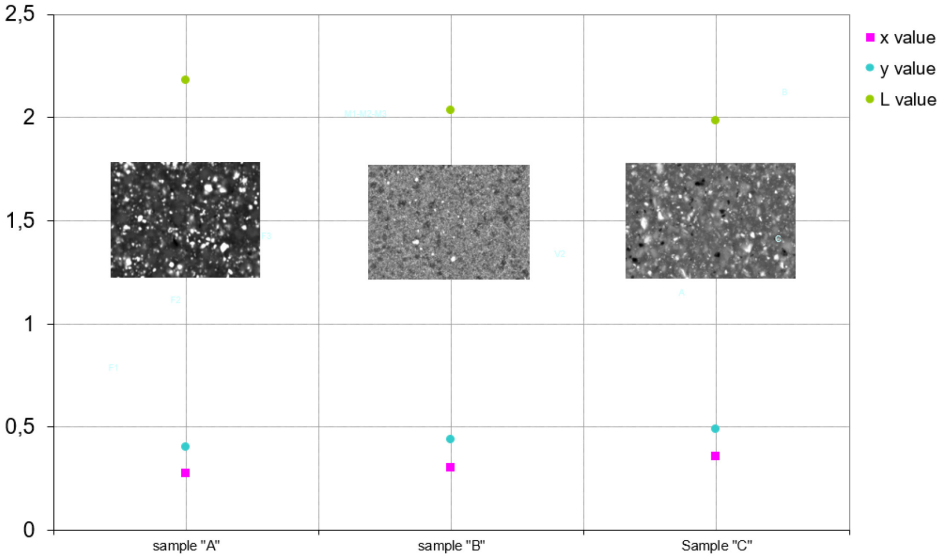


Figure 6. Colour coordinates and brightness of the samples with the colour measure software

In Fig. 6 it is also illustrated the values shown in the software of the colour measuring device (Table 1). The thickness of SM was measured at five points for each sample. Due to the grain structure and material composition, the measurement program calculated the average of five measurement points for each sample. In Fig. 6, the brightness value (L) is divided by 10. The SM thickness of all three samples was the same, 20 microns with a difference of 1-2 microns due to surface roughness. However, the quality of the copper under SM was different, so the *x*, *y* and *L* values also gave different results.

Table 1. Colour measurement results

Sample name	**Value *L*/10**	**Value *x***	**Value *y***
"A"	2,17943	0,271597	0,400578
"B"	2,03489	0,301941	0,438485
"C"	1,98118	0,353735	0,487435

4. Conclusions

During our present scientific summary, the raw material of electronic products, the surface of PCB, the SM layer. For identifying the electronic product and to trace the production processes, a product identification code is produced in the SM using laser beam marking. The condition of product identification is that DMC produced by the selected laser equipment can be read in the best possible quality.

It was found that the results of marking processes are greatly influenced by the optic properties of SM that depend on the composition, grain structure of the layer, surface roughness and the thickness of SM layer.

I conducted further research and based on scientific articles I found correlations that the varnish layer thickness is also a determining property in relation to the manufacturability of the PCB. The formation of ECM and various GC also depends on the properties and thickness of the varnish layer. My research also pointed out that colour measurement can be used to measure the quality and thickness of the layers and avoid the formation of manufacturing defects in electronic devices. By measuring the colour of the varnish layer, the manufacturing parameters can be optimized and thus the corrosion caused during the use of electronic devices can be reduced.

Acknowledgements

This work was supported by the Doctoral School of Engineering Sciences, Hungarian University of Agriculture and Life Sciences, Gödöllő, Hungary.

References

[1] Brewin, A., Willis, B. (2007). Characterising solder mask performance. *Report/Guide, National Physical Laboratory (NPL) Report*. https://eprintspublications.npl.co.uk/4043/

[2] Bušek, D., Dušek, K., Kulhavý, J. (2018). Dendritic growth and its dependence on various conditions. *Proceedings of the 2018 41st International Spring Seminar on Electronics Technology (ISSE)*, Zlatibor, Serbia, pp. 1–5. https://doi.org/10.1109/ISSE.2018.8443654

[3] Chowdhury, P.R., Suhling, J.C., Lall, P. (2018). Mechanical characterization of solder mask materials. *Proceedings of the 2018 17th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, IEEE: San Diego, CA, USA, pp. 1133–1141. https://doi.org/10.1109/ITHERM.2018.8419469

[4] Cucu, T.C., Varzaru, G., Turcu, C., Codreanu, N.D., Plotog, I., Fuica, R. (2008). 1D and 2D solutions for traceability in an Electronic Manufacturing Services company, *2008 31st International Spring Seminar on Electronics Technology*, pp. 585-588. https://doi.org/10.1109/ISSE.2008.5276596

[5] He, X., Azarian, M.H., Pecht, M.G. (2010). Effects of solder mask on electrochemical migration of tin-lead and lead-free boards. *IPC Printed Circuit Expo, APEX & Designer Summit Proceedings*. https://www.circuitinsight.com/pdf/effects_solder_mask_electromigration_ipc.pdf

[6] He, X., Azarian, M.H., Pecht, M.G. (2011). Evaluation of electrochemical migration on printed circuit boards with lead-free and tin-lead solder. *Journal Electronic Materials*, 40, 1921–1936. https://doi.org/10.1007/s11664-011-1672-3

[7] Hunt, R.W.G., Pointer, M.R. (2011). Measuring colour. 4th edition, The Wiley-IS&T Series in Imaging Science and Technology. John Wiley & Sons. ISBN: 978-1-119-97537-3.

[8] Illés, B., Medgyes, B., Horváth, A. (2014). Electrochemical migration behaviour of surface finishes after vapour phase reflow soldering. *Proceedings of the 2014 IEEE 20th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, Bucharest, Romania, pp. 253–257. https://doi.org/10.1109/SIITME.2014.6967039

[9] Illés, B., Medgyes, B., Dušek, K., Bušek, D., Skwarek, A., Géczy, A. (2022). Numerical simulation of electrochemical migration of Cu based on the Nernst-Plank equation. *International Journal of Heat Mass Transfer*, 184, 122268. https://doi.org/10.1016/j.ijheatmasstransfer.2021.122268

[10] Khanna, V.K. (2023). Humidity and contamination effects on electronics. In *Extreme-Temperature and Harsh-Environment Electronics (Second Edition)*, IOP Publishing, ISBN 978-0-7503-5072-3, Pages 15-1 to 15-12. https://doi.org/10.1088/978-0-7503-5072-3ch15

[11] Klimtová, M., Veselý, P., Králová, I., Dušek, K. (2024). Influence of solder mask on electrochemical migration on printed circuit boards. *Materials*, 17 (17), 4242. https://doi.org/10.3390/ma17174242

[12] Liao, B., Wang, H., Wan, S., Xiao, W., Guo, X. (2020). Electrochemical migration inhibition of tin by disodium hydrogen phosphate in water drop test. *Metals*, 10 (7), 942. https://doi.org/10.3390/met10070942

[13] Minzari, D., Grumsen, F.B., Jellesen, M.S., Møller, P., Ambat R. (2011). Electrochemical migration of tin in electronics and microstructure of the dendrites. *Corrosion Science*. 53 (5), 1659–1669. https://doi.org/10.1016/j.corsci.2011.01.009

[14] Piotrowska, K., Din, R.U., Jellesen, M.S., Ambat, R. (2018). Effect of solder mask surface chemistry and morphology on the water layer formation under humid conditions, *IEEE Trans. Compon. Packag. Manuf. Technol.* 8 (10), 1756–1768. https://doi.org/10.1109/TCPMT.2018.2792047

[15] Steen, W.M., Mazumder, J. (2010). Laser material processing (4th edition). ISBN: 978-1-84996-061-8, Springer London, https://doi.org/10.1007/978-1-84996-062-5

[16] Zhong, X. (2020). In situ study of the electrochemical migration of tin in the presence of H_2S. *Journal of Materials Science: Materials in Electronics*, 31, 8996–9005. https://doi.org/10.1007/s10854-020-03433-4

A process-based framework for efficient heat recovery in solar dryers for resource-strained environments

Abraham Kiprop JUMA [1], Zoltán KURJÁK [2] Janos BEKE [2]
[1] *Doctoral Program of Mechanical Engineering,*
Hungarian University of Agriculture and Life Sciences, MATE, Gödöllő
[2]*Department of Vehicle Technology, Institute of Technology,*
Hungarian University of Agriculture and Life Sciences, MATE, Gödöllő

Abstract

Solar drying is a sustainable method for post-harvest preservation, particularly in resource-constrained and off-grid environments. However, solar dryers often exhibit low thermal efficiency due to structural heat losses, exhaust air losses, air leakage, and non-uniform airflow. Existing research primarily emphasizes experimental and numerical modeling, which are time-consuming, resource-intensive, and impractical for low-income smallholder farmers. This paper proposes a process-based framework that systematically maps heat loss pathways and identifies cost-effective, practical heat recovery strategies. All major heat losses are evaluated, and the simplest, most affordable interventions are selected. The study finds that simple mechanisms such as air recirculation, baffles, multi-pass airflow design, low-cost sealants, and improved insulation using locally available materials are feasible. The aim of the study is to enhance product quality, minimize drying time, and improve overall thermal efficiency without relying on complex numerical simulations or experimental procedures. This work bridges the gap between theoretical research and practical implementation, thereby supporting sustainable solar drying technologies in resource-constrained contexts.

Keywords

Resource-constrained environments, process-based framework, solar drying, heat recovery

1. Introduction

Post-harvest losses remain a structural inefficiency in global agricultural systems, particularly in developing countries where preservation facilities are scarce and ineffective. A significant portion of perishable goods is lost due to limited storage and processing capacities (Nyamai et al., 2025). Drying is the dominant food conservation method, as it effectively removes moisture and inhibits microbial growth (Wang et al., 2023).

Solar drying is an attractive solution for post-harvest management because it relies on clean and abundant solar energy and has the potential for decentralized deployment. Despite these benefits, solar dryers often suffer from inefficient thermal performance. Studies indicate that only a limited fraction of solar energy is harvested, with a large portion lost to the environment (Tiwari, 2016).

Current research trends focus on experimental optimization and numerical modeling. Although these methods are technically robust, they are resource-demanding and typically remain at the pilot-testing stage, making them difficult to deploy in resource-constrained environments (Goel et al., 2023). There is a need to overcome the technical and economic barriers that limit the wide-scale adoption of solar dryers.

Most studies have focused on experimental optimization or complex numerical modeling to enhance dryer performance. While these approaches are valuable, they are resource-intensive and impractical for large-scale implementation, especially in low-income settings. This gap highlights the need for an alternative approach that leverages locally available materials while minimizing system cost and complexity. Hence, a process-based design framework is urgently needed to integrate heat recovery strategies with minimal reliance on complex modeling tools or expensive experimental setups.

The goal of this paper is to formulate a framework that links thermal losses to practical, cost-effective heat recovery interventions while highlighting key trade-offs and specific applicability in resource-constrained environments.

2. Solar drying fundamentals

Drying is a process that relies on simultaneous heat and mass transfer to remove moisture from products. Drying is a process that relies on simultaneous heat and mass transfer to remove moisture from products. Solar drying uses abundant solar energy to heat agricultural products and vaporize moisture, which is then carried away by airflow and exhausted from the system (Amer, 2022). The heat can be absorbed directly within the drying chamber or harnessed via a solar collector that heats the drying air. Mixed-mode solar dryers can combine both direct and indirect energy harnessing.

The drying process is complex due to several technical constraints. Drying efficiency depends on factors such as air temperature, velocity, and relative humidity, which must be optimized collectively rather than individually to maximize drying rates. Increasing air velocity can improve moisture removal but simultaneously increases exhaust heat losses. Similarly, high relative humidity reduces the drying air's capacity to carry moisture from the crops. Another key factor is airflow uniformity. Poor air distribution causes localized over- or under-heating of the drying chamber, reducing overall efficiency and product quality (Venkatesh et al., 2025).

Drying requires a delicate balance of these variables to enhance thermal performance. A process-based approach is proposed to account for the dynamic interactions among various drying variables, which are often overlooked in conventional studies. The aim is to optimize airflow, relative humidity, and temperature in solar dryers.

3. Process-based framework for heat recovery

This section presents a six-step guide for designing and fabricating solar dryers with thermal losses in mind. The proposed framework emphasizes real-world applicability over pilot-scale or experimental implementations. Figure 1 provides a schematic flow of the framework

Step 1: Identification of Heat Loss Pathways
Map all potential thermal loss pathways, including conduction, convection, radiation, air leakages, exhaust air, and non-uniform airflow.

Step 2: Define Evaluation Criteria
Determine criteria for evaluating interventions for each identified pathway, including ease of integration, cost, heat recovery efficiency, and maintenance requirements.

Step 3: Identification of Heat Recovery Strategies
List all possible interventions for each identified pathway.

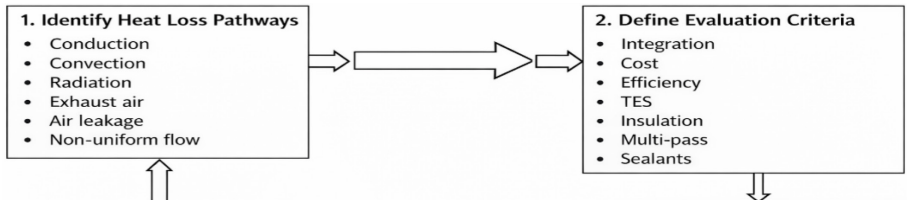


Figure 1. The schematic flow of the process-based framework for heat recovery in solar dryers.

Step 4: Selection of Suitable Solutions Based on Contex
Select strategies based on contextual relevance. For this study, priority is given to cost-effective solutions that require minimal technical expertise and utilize locally available materials, suitable for resource-constrained environments.

Step 5: Integration of Selected Solutions
Implement the chosen interventions for each pathway and ensure they function effectively within the system.

Step 6: Validation
Compare expected improvements with existing literature, simulation results, or small-scale experiments.

Overall, this framework provides a scalable, efficient, affordable, and practical guide to enhancing the thermal performance of solar drying systems without relying on complex modeling tools or capital-intensive experimental procedures.

4. Sources of thermal energy losses

The first step in enhancing the performance of solar dryers is identifying the heat loss pathways, which aids in developing cost-effective interventions.

4.1 Exhaust Air Losses
Exhaust air carries away approximately 30–45% of the input thermal energy (Murthi et al., 2025). In most conventional solar dryers, exhaust air is released into the environment without recovering its heat content. The temperature of the exhaust air is significantly higher than the ambient temperature. Exhaust heat loss is influenced by air velocity, but it is difficult to eliminate these losses completely. As a result, the most practical approach is to recover this heat and reuse it within the dryer or for secondary applications.

4.2 Structural Heat Losses
Conduction and radiation are the major sources of heat loss in the drying chamber and air heaters. Conduction losses occur through the walls, while radiation losses mainly occur through the transparent cover to the atmosphere. Heat dissipation through the walls is a major contributor within this category (Li et al., 2020). Structural losses depend on material properties, surface area, and environmental exposure. Although insulation can reduce conduction losses, many effective insulation materials are costly or inaccessible to smallholder farmers.

4.3 Air Leakage
Air leakage is a major issue that is often overlooked in the local fabrication of solar dryers. Bansal et al. (2016) noted that dryers incorporating air recirculation are particularly prone to leakage. Most existing studies have not conducted leakage tests on fabricated solar dryers. Air leakage typically occurs through door joints, structural joints, and during periodic opening of the dryer for product inspection. Leakage testing is often costly and impractical in resource-constrained regions. Therefore, the most feasible solution is to properly train local fabricators and identify cost-effective sealing materials.

4.4 Non-Uniform Airflow
Non-uniform airflow can significantly reduce system efficiency and result in uneven drying, thereby affecting product quality (Kidane et al., 2025). It

influences temperature and velocity distribution within the drying chamber and is largely dependent on system design and integration. Vertical drying configurations may lead to over-drying of lower layers and under-drying of upper layers. Additionally, certain regions, such as corners of the dryer, may be poorly heated, resulting in localized hot and cold zones.

5. Heat recovery strategies based on thermal loss pathways

Exhaust air accounts for the largest share of thermal losses in solar drying systems, making it a key target for heat recovery. Several approaches are available to recover this energy. Air-to-air heat exchangers transfer heat from exhaust air to incoming fresh air without mixing the two streams. Studies have reported heat recovery efficiencies of approximately 78–81% (Aacharya et al., 2024). However, the incorporation of heat exchangers increases system cost and complexity, making them less suitable for resource-constrained environments.

Active mechanical systems, such as solar-assisted heat pumps, have also been used to recover exhaust heat. In such systems, exhaust heat is used to preheat the working fluid in the evaporator, thereby improving overall drying efficiency. However, these systems are capital-intensive and technically complex.

Hybrid heat recovery systems, which combine two or more technologies, have also been explored. Despite their high thermal performance, they are generally expensive and complex, limiting their applicability in low-resource settings.

One of the simplest and most practical methods of exhaust heat recovery is partial recirculation of exhaust air. Existing studies have implemented this approach by recirculating a controlled portion of air, typically around 50%, using valves or dampers (Murthi et al., 2025). A key challenge with this method is the increased humidity of the recirculated air, which can reduce drying efficiency. However, recent studies suggest that adsorbent materials, such as silica gel, can be integrated into the system to absorb moisture and reduce relative humidity. This approach is cost-effective, simple to implement, and highly suitable for resource-constrained environments.

Table 1. A comparative analysis of existing exhaust heat recovery mechanisms in solar dryers.

Method	Heat Recovery Efficiency	East of Integration	Cost Level	Limitation
Recirculation	Moderate		low	High humidity
Heat exchanger	Moderate		high	Cost
Active mechanical	High		Very high	Cost and complexity
Hybrid	High		Very high	Cost and complexity

Hybrid systems and heat exchangers offer higher thermal recovery but are capital-intensive. In contrast, recirculation using manual dampers or valves provides a simple and low-cost solution, making it suitable for resource-constrained environments.

5.2 Structural Heat Loss Reduction

A significant portion of heat loss in solar dryers occurs due to conduction. While good conductors are essential for effective heat transfer from the dryer bed or air heater absorber plates, the system must be properly insulated to prevent heat loss to the environment. The most effective approach to minimizing conduction losses is the use of materials with low thermal conductivity. However, a balance must be achieved between insulation performance and material availability. In resource-constrained regions, the priority is to reduce conduction losses in a cost-effective manner.

Insulation materials can be broadly classified into high-performance materials, such as polyurethane and fiberglass, and low-cost alternatives, such as sawdust, clay, straw, wood, and polystyrene. Table X presents commonly available insulation materials along with their thermal conductivity and relative cost.

Radiation losses also occur in solar dryers, particularly through transparent covers. Several improvements have been proposed to minimize these losses. Previous studies show that the use of double glazing in solar air heaters provides higher heat retention compared to single glazing (Efendi, 2024). Double glazing creates an insulating air layer that reduces both radiative and convective heat losses. Table 2 provides a comparative analysis of structural heat reduction mechanisms

Table 2. A comparative analysis of the interventions to reduce structural heat losses in solar dryers.

Method	Leakage reduction	Ease of Integration	Cost Level	Limitations
High-perforamance insulation	High	High	High	Cost
Local insulation	Moderate	Medium	Low	Durability
Double glazing	Moderate	Medium	Low	Cost

Although high-performance materials offer superior thermal resistance, they are often expensive and less accessible in resource-constrained environments. In contrast, low-cost materials provide moderate thermal resistance but are affordable and readily available, making them more practical for small-scale applications.

5.3 Air Leakage Prevention

Existing studies indicate that air leakage is a major limitation in solar drying systems, particularly when air recirculation is implemented, as it leads to significant heat losses. The most effective solution is precise airtight construction. However, this approach is often impractical in resource-constrained environments due to its cost and the need for technical expertise.
Alternatively, the use of basic sealants such as foam tapes, rubber strips, and silicone provides a simple and cost-effective solution. In addition, proper structural design—especially at joints and door interfaces—plays a crucial role in minimizing air leakage. These approaches are practical and suitable for locally fabricated solar dryers.

Table 3. A comparative analysis of the interventions to minimize air leakage losses in solar dryers.

Method	Leakage reduction	East of Integration	Cost Level	Limitations
Airtight construction	High	High	High	Cost High skills
High-performance sealants	High	moderate	High	Cost
Low-cost sealants	Moderate	low	Low	Durability
Proper structural design	Moderate	moderate	moderate	High skills

Precise airtight construction is not feasible for smallholder drying systems. Basic sealing methods provide a high return on investment and require no specialized training. These sealing approaches can be effectively integrated with proper structural design to minimize air leakage.

5.4 Non-Uniform Airflow

Non-uniform airflow results in uneven temperature distribution in solar drying systems, which reduces drying efficiency and leads to inconsistent product quality. The most advanced solution is the use of automated control systems that incorporate sensors and feedback mechanisms. These systems regulate airflow and promote uniform drying rates. However, such control systems are complex and expensive, limiting their applicability in resource-constrained environments.
An alternative approach is the use of baffles and improved ducting. This method is cost-effective and enhances airflow distribution within the drying chamber (Kidane et al., 2025). However, it requires technical expertise in the design and fabrication of solar drying systems. Nevertheless, local fabricators can adopt and integrate existing baffle designs available in the literature.
Multi-pass airflow design presents a more practical and cost-effective solution for improving uniform heating and overall dryer performance, as it increases the

residence time of air within the system. However, this approach may result in pressure drops and increased air saturation levels. Therefore, airflow must be optimized to ensure sufficient residence time without exceeding the moisture-carrying capacity of the air. Proper multi-pass configuration can significantly enhance thermal performance without the need for complex and expensive interventions.

Table 4. A comparative analysis of the interventions to foster uniform airflow in solar dryers.

Method	Efficiency	East of Integration	Cost Level	Limitations
Automation	Very high	High	High	Complex
Baffles and Ducting	Moderate	Low	Low	Design
Multi-pass	High	Moderate	Moderate	Pressure drop

5.5. Selected Cost-Effective Solutions for Resource-Constrained Environments

Table 5 summarizes the selected heat recovery strategies across the different thermal loss pathways for resource-constrained environments. The selection criteria are based on ease of integration, low cost, use of local materials, and minimal technical expertise.

Table 5. Cost-effective heat recovery solutions for each thermal loss pathway in solar drying for resource-constrained environments

Heat Loss pathway	Solution	Justification
Exhaust Air	Air recirculation	Low cost Easy integration Significant heat recovery
Structural losses	Low-cost isulation Double glazing	Better heat retention with locally available and cost-effective materials
Air leakage	Low-cost sealants	Low cost Easy integration
Non-uniform airflow	Use of baffles Multi-pass airflow design	Better uniformity at a low cost

The best solution for exhaust air losses is the partial recirculation of air using dampers or valves. This technology is cost-effective and easy to integrate into solar dryers. Other alternatives, such as hybrid systems, offer higher heat recovery efficiency but are complex and capital-intensive, making them impractical in low-resource environments.

The use of low-cost materials, such as sawdust coupled with double glazing, is practical for minimizing structural losses. This intervention provides a fair

trade-off between performance and affordability. Although high-performance insulation materials offer superior thermal resistance, they are expensive and scarce in low-resource settings.
There are various alternatives for addressing air leakage, but the most practical option in resource-constrained environments is the use of basic sealants, such as silicone. These are easy to apply, low-cost, and require minimal technical expertise.

The use of baffles and multi-pass airflow design is recommended for enhancing airflow uniformity. These two techniques promote uniform air distribution and increase air residence time in the drying chamber, resulting in better drying performance. Automated control is the most effective method for improving airflow distribution; however, its complexity and high cost make it unsuitable for resource-constrained environments.

The selected interventions demonstrate that significant thermal improvements can be achieved in solar dryers without sophisticated simulations or experimental studies. Overall, the framework provides a simple, low-cost, and scalable approach to enhancing solar drying in real-world setups.

6. Validation of the framework

The existing literature on thermal losses in solar dryers provided the basis for developing this framework. Conventional solar dryers operate at low thermal efficiency due to significant heat losses, which can be as high as 50% (Murthi et al., 2025). The integration of locally available insulation materials and sealants can substantially reduce these losses. Applying the proposed heat recovery mechanisms can improve the overall thermal efficiency of solar dryers.
There is a need to implement these solutions in locally fabricated dryers. Simulations can also be performed to evaluate the impact of all interventions on thermal performance. The framework enables significant reduction of heat losses without requiring expert knowledge or high capital investment. In this way, the thermal performance of solar dryers can be enhanced using simple, cost-effective mechanisms without the need for sophisticated technologies.

7. Practical implications

The proposed framework is applicable to both small-scale and industrial solar drying systems. In resource-constrained environments, substantial improvements in thermal performance can be achieved using simple and low-cost interventions, such as local insulation materials, sealants, partial air recirculation, and basic multi-pass airflow configurations. These interventions require minimal technical expertise and are highly cost-effective.

Although primarily designed for low-resource settings, the approach can also be applied to larger-scale drying systems. The key heat loss pathways are similar across different dryers, while the choice of solutions may vary depending on economic conditions. Industrial-scale implementations can incorporate additional technologies, such as thermal storage, heat exchangers, high-performance insulation, and airtight construction.

Conclusion

This paper presents a process-based framework for integrating heat recovery in solar dryers, with emphasis on practical strategies tailored for resource-constrained environments. The framework enhances thermal performance by linking major heat loss pathways—exhaust air, structural losses, air leakage, and non-uniform airflow—to the most suitable, locally applicable interventions.
Cost-effective strategies identified in this study include partial air recirculation, low-cost insulation coupled with double glazing, basic sealants, and baffles combined with multi-pass airflow design. The framework bridges the gap between theoretical research and real-world applications, enabling small-scale fabricators to improve drying performance without complex and resource-intensive numerical or experimental studies.

Future studies should focus on validating these interventions in field conditions and providing additional recommendations to further enhance the performance and scalability of solar dryers. Overall, this approach contributes significantly to sustainable post-harvest management in resource-constrained environments.

References

[1] Aacharya, A., Rissler, C., Baral, B., Lhendup, T., Andersson, M., & Davidsson, H. (2024). Development of a novel solar dryer with an incorporated heat exchanger. *Solar Energy*, *269*, 112327.

[2] Amer, B. M. (2022). Solar Drying: Principles and Applications. *Thermal Food Engineering Operations*, 179-196.

[3] Atalay, H., Turhan Çoban, M., & Kıncay, O. (2017). Modeling of the drying process of apple slices: Application with a solar dryer and the thermal energy storage system. Energy, 134, 382–391. https://doi.org/10.1016/j.energy.2017.06.030

[4] Bansal, P., Mohabir, A., & Miller, W. (2016). A novel method to determine air leakage in heat pump clothes dryers. *Energy*, *96*, 1-7.

[5] Efendi, M. (2024). Influence of glazing type on the drying kinetics and thermal performance of indirect solar dryer for jelly candy. *Renewable Energy*, *231*, 120950.

[6] Goel, V., Bhattacharyya, S., Kumar, R., Pathak, S. K., Tyagi, V. V., & Saini, R. P. (2023). Identification of barriers and drivers to implementation of solar drying technologies: V. Goel et al. *Journal of Thermal Analysis and Calorimetry*, *148*(7), 2977-3000.

[7] Kidane, H., Farkas, I., & Buzás, J. (2025). Enhancing the drying uniformity in solar drying systems: computational and experimental study. *International Journal of Thermofluids*, 101408.

[8] Li, T., Li, C., Li, B., Li, C., Fang, Z., Zeng, Z., ... & Huang, J. (2020). Characteristic analysis of heat loss in multistage counter-flow paddy drying process. *Energy Reports*, *6*, 2153-2166.

[9] Murthi, M. K., Munimathan, A., Kareemullah, M., & Khan, T. Y. (2025). Experimental analysis and optimization of an indirect solar dryer integrated with heat exchanger for enhanced exhaust air utilization. *Case Studies in Thermal Engineering*, 107101.

[10] Nyamai, M., & Anuri, P. (2025). Post-harvest loss reduction strategies and their contribution to household food security and nutrition in Kitui County, Kenya. African Journal of Food, Agriculture, Nutrition and Development, 25(09), 27726–27740. https://doi.org/10.18697/ajfand.146.26285

[11] Tiwari, A. (2016). A review on solar drying of agricultural produce. *Journal of Food Processing & Technology*, *7*(9), 1-12.

[12] Venkatesh, R., Singh, P. K., Karthik, K., Vinayagam, M., Verma, A., Prabagaran, S., ... & Alharbi, S. A. (2025). Integration of phase change material for enriching the solar collector featured with dryer configuration enhanced via alumina/titanium dioxide nanoparticle: performance study. *Journal of Thermal Analysis and Calorimetry*, *150*(12), 9535-9548.

[13] Wang, D., Zhang, M., Ju, R., Mujumdar, A. S., & Yu, D. (2023). Novel drying techniques for controlling microbial contamination in fresh food: A review. *Drying Technology*, *41*(2), 172-189.

Machine Learning for Vibration Signal Analysis: A Literature Review

Tamim DEEB[1], László KÁTAI[2]
[1]Doctoral Program of Mechanical Engineering, MATE, Gödöllő
[2]Department of Machine ConstructionInstitute of Technology, MATE, Gödöllő

Abstract

Vibration signals carry diagnostic information about mechanical and structural health that is now extracted primarily by machine learning rather than by hand-crafted signal processing alone. This review synthesises 100 papers covering the two major application domains of vibration-plus-ML, rotating machinery fault diagnosis and structural health monitoring, and the three method families that span them: classical machine learning on engineered features, deep learning architectures consuming time-frequency images, and hybrid pipelines combining signal processing, physics-informed objectives, and explainable AI. Three observations recur. First, no single method dominates across applications: classical ML retains advantages on small, well-engineered problems and at the edge, while deep learning takes over when labelled data scales and signals carry non-stationary content best captured by 2D time-frequency representations. Second, the boundary between signal processing and machine learning has dissolved, with continuous wavelet transform scalograms now the dominant input to CNN-based diagnostic pipelines. Third, the remaining open problems are not transducer-physics or architecture problems but integration challenges: reproducibility, cross-machine generalisation, edge deployment, multi-modal fusion, and operator trust. The review covers vibration-signal representations, classical and deep-learning approaches, applications in rotating machinery and civil infrastructure, hybrid and explainable-AI methods, comparative method-selection criteria, and the open challenges that will shape the next phase of progress.

Keywords

vibration signal analysis, machine learning, fault diagnosis, convolutional neural network, deep learning, explainable AI.

1. Introduction

Mechanical vibration carries a remarkable amount of diagnostic information. A bearing developing a spall, a gear losing a tooth, a bridge cable losing tension, a wind-turbine blade developing a crack — each produces a vibration signature.

Detecting and classifying those signatures has been the work of condition monitoring and structural health monitoring for decades. What has changed in the last few years is the toolkit. Machine learning has become the dominant analytical layer between the raw vibration signal and the diagnostic decision, replacing or augmenting hand-engineered signal processing with learned models that can extract diagnostic features directly from time, frequency, or time-frequency-domain representations.

The shift toward ML has reshaped the field along three axes simultaneously. The first is the representation axis. Where classical analysis worked with hand-engineered statistical features, the modern pipeline more often converts the signal into a 2D time-frequency image — a Continuous Wavelet Transform scalogram, a Short-Time Fourier Transform spectrogram, a Synchrosqueezing Transform map — and feeds the image to a Convolutional Neural Network designed for image input (Hamdaoui et al., 2023; Verstraete et al., 2017; Łuczak, 2024). The second is the architecture axis. CNNs dominate the deep-learning side of the literature, but hybrid CNN-LSTM combinations are now standard when temporal structure matters, and recent papers are introducing transformers, graph neural networks, and physics-informed neural networks (Mousavi et al., 2025; Tama et al., 2022). The third is the interpretability axis. Explainable AI tools — SHAP, Grad-CAM, LIME, LRP-Z — have moved from research curiosities to standard practice, because plant engineers and SHM practitioners need a basis for trusting the model's output before they act on it (Brusa et al., 2023; Mey et al., 2022; Chen & Lee, 2020).

The story is not, however, one of deep learning displacing everything that came before. Classical ML SVM, random forest, KNN, decision trees on engineered features, remains highly competitive when the labelled dataset is small, when feature engineering is well-matched to the physics, and when the deployment platform is compute-constrained (Atmaja et al., 2022; Shandookh et al., 2024; Kafeel et al., 2021). Operational Modal Analysis remains the working baseline for civil SHM, on top of which ML pipelines build (Saidin et al., 2023; Wang et al., 2022). Several recent papers explicitly choose interpretable classical models over higher-accuracy black-box models for the deployment context, on the grounds that an unaccountable 99% accuracy is worse than an auditable 95% (Khalil & Rostam, 2024).

Figure 1 sketches the canonical pipeline that recurs throughout this literature: a vibration sensor produces a raw signal, a signal-processing layer converts it into a representation the model can consume, an ML or DL model performs the classification or regression, and an interpretability layer optionally audits the output. The remainder of this review unpacks each stage of that pipeline and the methods used at each.

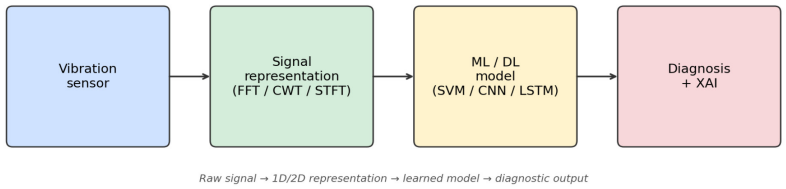


Figure 1. Canonical vibration-plus-ML pipeline

This review covers the methods, applications, and open problems in vibration-plus-ML across two major application domains (rotating machinery and structural health monitoring) and three method families (classical ML, deep learning, and hybrid/physics-informed/XAI approaches.

2. Vibration signals: representations and preprocessing

Before any machine-learning algorithm consumes a vibration signal, the signal has to be put into a form the algorithm can work with. The choice of representation is at least as consequential as the choice of model, because most of the diagnostic information lives in how the signal is structured in time and frequency, not in the raw samples themselves. Three representations dominate the literature in this collection: time-domain statistical features, frequency-domain spectra, and time-frequency images. Each carries different inductive biases that suit different problems.

2.1. Time-domain features

The simplest representation is a vector of statistical descriptors computed directly on the time-series signal: RMS, peak-to-peak, kurtosis, skewness, crest factor, and zero-crossing rate. Papers throughout this collection use small numbers of these features to feed classical ML classifiers. Atmaja et al. (2022) reached 99.75% weighted accuracy on lab-scale unbalance, misalignment, and bearing faults with an SVM trained on seven such features (Atmaja et al., 2022). Shandookh et al. (2024) hit 99% on belt-drive faults using statistical descriptors including variance, MAD, and signal energy. The advantage is computational economy: the feature vector is small, the classifier is fast, and the result is interpretable. The disadvantage is that statistical features lose the temporal localisation and frequency content that more sophisticated representations preserve.

2.2. Frequency-domain representations

The Fast Fourier Transform (FFT) and its derivatives — Power Spectral Density (PSD), cepstrum, envelope spectrum — convert the time signal into a frequency-axis representation that exposes harmonic structure, resonance peaks, and characteristic defect frequencies. FFT-based inputs remain a workhorse for both classical ML and deep learning. Janssens et al. (2016) trained a 2D-CNN directly on the FFT spectrum, treating it as a 1D image (Janssens et al., 2016). Wang et al. (2020) reshaped the FFT into a 2D matrix for their MDIAN domain-adaptation network (Wang et al., 2020). Han et al. compared FFT-and-PCA features against multiple classifiers and found uniform 100% state classification on precision-machining data (Han et al., 2021). The unresolved limitation is stationarity: FFT assumes the signal is stationary across the window it analyses, which mechanical fault signals often violate.

2.3. Time-frequency representations

Time-frequency methods preserve both temporal and spectral information. The Short-Time Fourier Transform (STFT) divides the signal into overlapping windows and computes the FFT of each. The Continuous Wavelet Transform (CWT) uses scaled wavelet basis functions to capture features at multiple resolutions simultaneously. The Discrete Wavelet Transform (DWT) and Wavelet Packet Transform (WPT) provide computationally cheaper alternatives. Each produces a 2D image that can be fed directly to a CNN designed for image input.

The empirical evidence in this collection strongly favours CWT scalograms over STFT spectrograms when high temporal-frequency resolution matters. Hamdaoui et al. (2023) reported 96.09%–99.70% accuracy with CWT scalograms feeding a custom three-layer CNN on rolling-element bearings under varying speed, well above their SVM, DNN, and raw-LSTM baselines (Hamdaoui et al., 2023). Łuczak (2024) stacked six CWT scalograms (one per IMU axis) into a single RGB image, processing each fault classification in under 50 ms (Łuczak, 2024). The Chen and Lee (2020) paper provided an STFT-based alternative on CWRU bearing data with Grad-CAM verifying that the network attended to the high-frequency resonance bands (Chen & Lee, 2020). Rahim et al. (2024) compared time-series, FFT, and CWT representations directly on electric-motor mechanical faults and concluded that CWT magnitude scalograms preserved diagnostic information that FFT alone could not (Rahim et al., 2024).

Figure 2 illustrates why this representation matters. A synthetic healthy-bearing signal carries low-amplitude harmonic content with broadband noise; a faulty bearing introduces periodic high-frequency impulses tied to the characteristic defect frequency. In the time domain these impulses can be hard to separate from background noise, but in the time-frequency representation the periodic structure shows up clearly as repeating high-frequency bursts. A CNN trained on

these images can pick up the defect signature without any hand-engineered feature extraction.

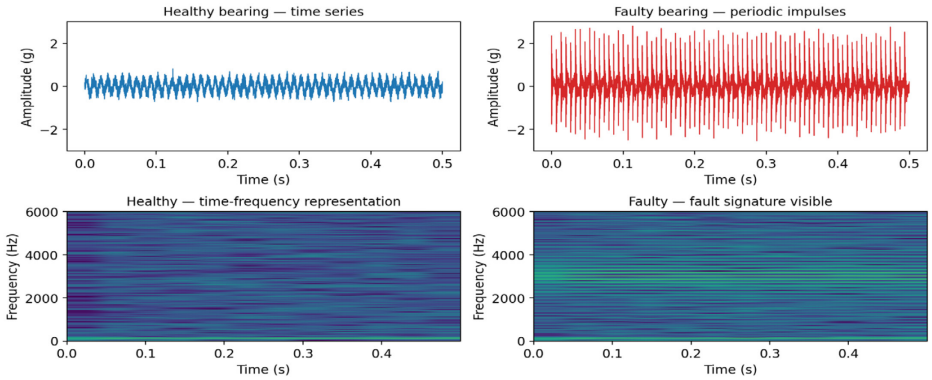


Figure 2. Time-series and time-frequency representations of synthetic healthy versus faulty bearing signals. Periodic impulses at the bearing's characteristic defect frequency are difficult to isolate in the time domain (top) but appear as distinct high-frequency bursts in the time-frequency map (bottom). Schematic generated by the author for illustration.

A small but growing subset of papers uses more exotic time-frequency representations. Radicioni et al. (2025) applied Synchrosqueezing Transform (SST) spectrograms for unsupervised anomaly detection with convolutional autoencoders, finding that the sharper time-frequency localisation of SST helped reconstruction-based anomaly detection (Radicioni et al., 2025). Mey et al. (2022) used Order Analysis with Fourier-transformed unbalance data, achieving 99.66% on rotating machinery (Mey et al., 2022).

2.4. Decomposition methods as preprocessing

Empirical Mode Decomposition (EMD) and its descendants — Ensemble EMD (EEMD), Variational Mode Decomposition (VMD), Empirical Wavelet Transform (EWT) — decompose a non-stationary signal into a small set of intrinsic mode functions (IMFs) without assuming a basis. The decomposition is a preprocessing step that can be paired with any downstream classifier. Mousavi et al. (2025) reconstructed time-series signals via VMD before feeding them into a 1D-CNN+LSTM hybrid for offshore wind-turbine monopile structures, pushing accuracy from 94.72% on raw data to 98.05% on VMD-reconstructed data (Mousavi et al., 2025). Kafeel et al. (2021) used EMD-derived IMFs to build a hybrid 13-feature time/frequency vector for SVM classification on induction-motor faults, reaching 98.2% accuracy (Kafeel et al., 2021). The consensus is that decomposition methods improve performance when the signal carries multiple superimposed periodicities such as gear meshing alongside

bearing defects, that conventional time or frequency domain methods cannot cleanly separate.

2.5. Denoising and missing-data handling

Industrial signals are noisy, and operational data has gaps. Jang et al. (2022) used a denoising autoencoder (DAE) as a preprocessing layer for one-class SVM-based anomaly detection on urban-railway electric motors, producing F1-scores of 0.967 on both 11-kW and 55-kW units (Jang et al., 2022). The DAE preserved the signal's statistical distribution better than wavelet thresholding, which is the alternative approach to denoising. Martins et al. (2024) addressed the missing-data problem in industrial sensor logs with a DNN-MLP imputation step before classification (Martins et al., 2024). Both interventions are pre-classification preprocessing rather than algorithm choices in the classical ML sense.

2.6. Synthesis

Table 1 summarises which representations pair well with which classifier families, based on the patterns seen across this corpus.

Table 1. Pairing of vibration-signal representations with classifier families, summarised across the papers reviewed in Sections 3 and 4.

Representation	Best-fit classifier	Strengths	Weaknesses
Statistical descriptors (RMS, kurtosis, skewness)	SVM, KNN, random forest	Compact, interpretable, low compute	Loses temporal and frequency content
FFT spectrum / PSD	SVM, 1D-CNN	Exposes harmonics and defect frequencies	Assumes stationarity
STFT spectrogram	2D-CNN	Time-frequency localisation; standard image input	Fixed window trades time/frequency resolution
CWT scalogram	2D-CNN	Multi-resolution; sharp localisation of single-frequency components	Higher compute; CNN-specific
EMD / VMD intrinsic mode functions	SVM, 1D-CNN	Adaptive to non-stationary signals	Hyperparameter-sensitive, mode mixing
Denoising autoencoder output	One-class SVM, classical ML	Removes noise without distribution skew	Needs training data; adds complexity

The literature is consistent on three points. First, statistical features remain a strong baseline when the dataset is small and the physics is well-understood. Second, CWT scalograms have become a common default time-frequency input

to CNN-based diagnostics; in head-to-head comparisons with STFT spectrograms the two often reach comparable accuracy, with CWT preferred when sharper localisation of single-frequency components matters. Third, decomposition methods (VMD, EMD) are the most reliable way to handle non-stationary content, but they add a hyperparameter-tuning step that classical FFT does not require. The choice of representation is therefore not orthogonal to the choice of classifier, a CWT scalogram is a good input for a CNN but a poor one for an SVM, and a small statistical feature vector is the inverse. Sections 3 and 4 elaborate on how these representations have been combined with classical and deep-learning models respectively.

3. Classical machine learning on engineered vibration features

Before the deep-learning shift discussed in Section 4, the workhorse of vibration analysis was a hand-engineered feature vector — statistical descriptors, FFT bin energies, wavelet coefficients — fed to a classical classifier such as SVM, KNN, random forest, or linear discriminant analysis. That pipeline is far from obsolete. The five papers in this section show that classical approaches still hold their own on bearings, motors, and machining applications, and in some cases beat deep networks when datasets are small or feature engineering is well-matched to the physics.

3.1. Feature engineering: which transform feeds which classifier

The choice of signal representation is at least as important as the choice of classifier. Han et al. (2021) used FFT followed by PCA to compress raw audio and vibration signals, then evaluated KNN, CNN, and SVM for state classification and Neural Network Regression (NNR) and Support Vector Regression (SVR) for continuous-parameter prediction on precision machining data. All three classifiers reached 100% state classification accuracy on PCA-reduced features, but NNR consistently outperformed SVR for regression. The headline contribution is that fusing audio and vibration channels improved regression accuracy beyond either modality alone (Han et al., 2021). The headline limitation is data volume, the method was tested on a limited number of signal points and would need scaling for production deployment.

Altaf et al. (2022) explored a less commonly used transform: statistical features extracted not from the raw signal but from the second derivative of the time, FFT, and PSD domain signals. They evaluated KNN, SVM, and Kernel LDA on roller-bearing faults (outer race, inner race, and ball faults). Kernel LDA on PSD-derivative statistical features reached 99.13% accuracy while compressing the feature vector by over 95% (from 40,000 samples to an 8×228 array), and outperformed traditional EMD-based pipelines on the same problem (Altaf et al., 2022). The second-derivative trick is the contribution: it separates fault classes that overlap heavily in the raw and first-derivative feature spaces.

Rahim et al. (2024) make the case for skipping classification entirely when the visual signature is unambiguous. Comparing time-series, FFT, and CWT scalogram representations of electric-motor mechanical faults (bearing damage, looseness of bearing mountings), they showed that CWT magnitude scalograms preserved both frequency content and time localisation in a way that FFT alone could not (Rahim et al., 2024). The CWT view enabled visual inspection-based diagnosis without requiring an ML classifier, a useful baseline for understanding what an automated system would otherwise need to extract.

3.2. EMD-based features and SVM variants

Empirical mode decomposition (EMD) remains a popular preprocessing step for non-stationary mechanical signals. Kafeel et al. (2021) decomposed the vibration signal of a 45-kW three-phase induction motor into intrinsic mode functions (IMFs), built a hybrid 13-feature time/frequency vector from those IMFs, and tested SVM, KNN, LDA, and decision trees. SVM with a quadratic kernel reached 98.2% accuracy, 96.6% sensitivity, and 100% specificity on real industrial-scale data from centrifugal water pumps, validated on 220 multi-axis signals (Kafeel et al., 2021). The contribution is that the validation used real industrial data rather than the lab-simulated CWRU or MFPT datasets that dominate the literature. The limitation is binary classification: the current implementation distinguishes healthy from faulty without isolating fault types.

3.3. Denoising autoencoders as a preprocessing layer

A practical mid-point between classical ML and full deep learning is to use a small neural network as a preprocessing layer. Jang et al. (2022) trained a denoising autoencoder (DAE) to clean vibration signals from electric motors in urban railway station equipment, then fed time-domain statistical features (mean, RMS, skewness, kurtosis, and standard deviation) into a one-class SVM for anomaly detection. The mean feature achieved an F1-score of 0.967 on the 11-kW motor and the skewness feature reached the same 0.967 on the 55-kW motor, while standard deviation — the most informative feature overall — produced an F1-score of 0.96 on both motor types (Jang et al., 2022). The advantage of the DAE over a wavelet-based denoiser is that it removes noise without the data-loss and distribution-skewing that wavelet thresholding tends to introduce. The limitation is that the study only verified performance on a single fault type and needs extension to bearing and shaft faults.

4. Deep learning architectures for vibration signals

Deep learning has reshaped vibration analysis in the past five years. The dominant pattern is no longer a hand-engineered feature vector followed by a shallow classifier, but a learned representation extracted directly from the raw

signal or from a 2D time-frequency image of it. The eight papers in this section illustrate the architectural choices that recur across applications: 2D CNNs on time-frequency images, hybrid 1D-CNN/LSTM combinations on temporal signals, transfer learning across sensor types, and unsupervised autoencoders for anomaly detection.

4.1. CNNs on time-frequency images

The most common recipe is to convert a 1D vibration signal into a 2D time-frequency image — a scalogram, spectrogram, or related representation — and feed it to a convolutional network designed for image input. Chen and Lee (2020) implemented this pipeline with STFT spectrograms and a 2D-CNN for rolling-bearing fault classification on the CWRU dataset, then used Grad-CAM to verify that the network was looking at the structurally meaningful high-frequency resonance bands and not at spurious features (Chen & Lee, 2020). Łuczak (2024) extended the idea to multi-axis fan-blade imbalance detection by stacking six continuous wavelet transform (CWT) scalograms — one per axis of a 6-DOF IMU — into a single three-channel RGB image. The combined CWTx6-CNN reached fault classification in under 50 ms per sample, with the CWT computation itself taking less than 35 ms, and the network was further audited with Grad-CAM, LIME, and occlusion analysis to confirm interpretability (Łuczak, 2024).

Gómez Muñoz et al. (2025) used the same scalogram-plus-CNN template but added a transfer-learning twist. They trained GoogLeNet on CWT scalograms generated from full-field laser Doppler vibrometer (LDV) measurements of a metallic plate with non-linear boundary impacts, then fine-tuned on sparse accelerometer data. GoogLeNet reached up to 100% classification accuracy on specific frequency-band categories, substantially outperforming ResNet (49.25%) and SqueezeNet (50%) on the same task (Gómez Muñoz et al., 2025). The contribution matters because it inverts the usual cost structure: the expensive instrument is used to teach the model, after which the cheaper accelerometer can be deployed at scale.

4.2. Hybrid 1D-CNN and recurrent architectures

When the temporal structure of the signal matters as much as its frequency content, hybrid 1D-CNN and recurrent designs perform better than pure 2D-CNN-on-image approaches. Mousavi et al. (2025) used variational mode decomposition (VMD) to reconstruct the raw time-series signal, then fed the reconstructed signal directly into a 1D-CNN+LSTM hybrid for damage detection in monopile offshore wind-turbine structures with soil interaction. The VMD-reconstructed input pushed accuracy from 94.72% on raw data to 98.05% on the reconstructed series, and the architecture eliminates manual feature engineering entirely (Mousavi et al., 2025).

Rihi et al. (2024) took a different hybrid approach: an LSTM forecasts future vibration sequences, those forecasted sequences are converted to MFCC spectrograms, and a CNN performs the final classification. The two-stage design reached 99.95% accuracy and an AUC-ROC of 1.0 on a real-world mining-grinding-mill dataset, and 99.96% accuracy across 16 fault classes on the CWRU bearing dataset. The advantage over static diagnostic models is proactive intervention — the LSTM provides lead time before the fault actually manifests (Rihi et al., 2024).

4.3. Unsupervised autoencoders for anomaly detection

When labelled fault data is scarce, supervised classifiers cannot be trained. Radicioni et al. (2025) addressed this with unsupervised convolutional autoencoders (CAEs) and variational autoencoders (VAEs) trained on Synchrosqueezing Transform (SST) spectrograms for predictive maintenance in industrial machinery. The CAEs with spatial latent spaces outperformed those with standard dense latent spaces on both reconstruction quality and anomaly-detection sensitivity. The paper introduces "ReLU residuals" that amplify the pixel-level visibility of anomalies, and reports honestly that VAEs did not perform as well because the standard reconstruction-based metrics they minimise are not inherently suited to anomaly detection (Radicioni et al., 2025).
Martins et al. (2024) took an unsupervised route that is more pragmatic and less architecturally exotic. Their pipeline imputes missing values from a drying-press-roll sensor using a DNN-MLP, applies PCA for dimensionality reduction, then uses K-means and a Hidden Markov Model for state classification. The contribution is operational rather than algorithmic: the system handles the spaced or missing time-series observations that are typical of industrial sensor logs, without requiring labelled training data (Martins et al., 2024).

4.4. What the literature confirms

Abdullah et al. (2025) provide a useful synthesis. Their five-year review of ML and DL methods for fault diagnosis and prediction in rotating machinery (bearings, rotors, gears, shafts) finds that deep learning approaches consistently outperform classical ones on the same data. They report 1D-DCNN at 99% accuracy and ResNet-3N at 99.23% on rotating-machinery faults, against KNN at 66.61% and random forest at 71.45% on the same problems (Abdullah et al., 2025). The size of the gap is consistent with what the empirical papers in this section show. The qualifier is that the gap exists when there is enough labelled data to train the deep network; for low-data settings, the classical methods covered in Section 3 remain competitive.

Table 2 collects reported accuracies from across the corpus to make the comparison concrete. The deep-learning column is consistently above the classical column on the same benchmarks, but the classical column never drops below the threshold a practitioner would care about when interpretability or compute cost dominates the decision

Method family	Architecture / approach	Reported accuracy	Source
Classical	KNN on CWRU	66.61%	Abdullah et al. (2025)
Classical	Random forest on CWRU	71.45%	Abdullah et al. (2025)
Classical	Decision tree on engineered features	95.8%	de las Morenas et al. (2023)
Classical	Random forest with 6 features (CWRU wind turbine)	98.58%	Section 5.3
Classical	SVM-Q on 13-feature EMD vector	98.2%	Kafeel et al. (2021)
Classical	SVM on 9 spectrum features	99.75%	Atmaja et al. (2022)
Deep	1D-DCNN on rotating machinery	99.0%	Abdullah et al. (2025)
Deep	ResNet-3N on rotating machinery	99.23%	Abdullah et al. (2025)
Deep	CWT scalogram + custom 3-layer CNN	96.09–99.70%	Hamdaoui et al. (2023)
Hybrid	LSTM forecast + MFCC + CNN	99.96%	Rihi et al. (2024)
Hybrid	MDIAN (cross-condition transfer)	99.06%	Wang et al. (2020)
Hybrid	VMD + 1D-CNN-LSTM	98.05%	Mousavi et al. (2025)

5. Rotating machinnery and bearing fault diagnosis

Rotating machinery is the single largest application domain for vibration-based machine learning. The 37 papers in this section span every combination of signal representation and classifier, but they cluster into a handful of recurring patterns. The dominant pipeline today combines a time-frequency representation (scalogram or spectrogram) with a CNN or CNN-LSTM hybrid; classical ML methods remain competitive when the feature set is well-matched to the physics; and recent work is increasingly attentive to transfer learning, interpretability, and reproducibility constraints.

5.1. Time-frequency representations with CNNs

The single most common recipe in this collection is to convert a vibration signal into a 2D time-frequency image and feed it to a convolutional network. Janssens et al. (2016) provided one of the early demonstrations, training a 2D-CNN directly on the scaled DFT of bearing signals and reaching 93.61% accuracy and 94.06% F1-score, significantly above a feature-engineered random forest baseline (Janssens et al., 2016). Subsequent papers refined the inputs and the architectures. Hamdaoui et al. (2023) combined variational mode decomposition and wavelet thresholding with CWT scalograms and a custom three-layer CNN, delivering accuracies between 96.09% and 99.70% across inner race, outer race, and ball faults under varying speed (Hamdaoui et al., 2023). Other recent work converts variable-speed gearbox signals to CWT and DWT scalogram images and applies augmentation (flipping, Gaussian noise) to address the perennial scarcity of labelled fault data.

Wang et al. (2020) attacked one of the harder generalisation problems with their MDIAN architecture, a modified ResNet-50 with multi-scale feature extractors and Conditional Maximum Mean Discrepancy (CMMD) for domain alignment. MDIAN reached 99.06% mean testing accuracy across transfer tasks between bearing operating conditions, well above standard CNNs and SVMs that struggle when the test distribution shifts (Wang et al., 2020). A 2025 transfer-learning paper standardised heterogeneous datasets (CWRU, HUST, PADERBORN) into fixed-size spectrograms and compared ImageNet-initialized DenseNet121 against a domain-specific "VibNet". Full fine-tuning reached 97.58% accuracy on CWRU; the domain-specific network reached 97.39% with lower variance, suggesting that for well-funded labs with relevant pretraining data, domain-specific transfer is the more reliable choice.

5.2. Hybrid CNN-LSTM and recurrent architectures

When the temporal structure of the signal carries diagnostic content that a 2D image flattens, hybrid architectures consistently outperform pure CNNs. A 2025 paper on industrial motors with inner-race and ball defects reported that a CNN-LSTM hybrid reached 94.6%, against 88.2% for SVM, 91.5% for CNN, and 92.7% for LSTM in isolation. A 2025 study on rotating-machinery defects reached 97% average fault-detection accuracy with a CNN under industrial noise, after applying SMOTE to recover a 15% gain on minority-class recall. Across the corpus, CNN-LSTM pairs sit slightly above pure CNN on most rotating-machinery benchmarks, with the differential largest under variable speed.

5.3. When classical ML still wins

Despite the deep-learning shift, several papers in this collection show that classical methods remain competitive when feature engineering matches the

physics. Atmaja et al. (2022) reached 99.75% weighted accuracy on lab-scale unbalance, misalignment, and bearing-fault data using SVM on nine spectrum-domain statistical features. They also released the dataset as a free open-access CSV, which is a notable contribution to a literature otherwise dominated by closed benchmarks (Atmaja et al., 2022). Shandookh et al. (2024) achieved 99% accuracy on belt-drive system faults with a 100-tree random forest using statistical features including variance, MAD, zero-crossing rate, autocorrelation, and energy. The takeaway is not that deep learning is unnecessary but that for problems with clean signals and well-understood physics, a small classical pipeline can match much heavier models.

A 2025 wind-turbine-bearing study on the CWRU dataset reinforces this with a nuance: a random forest classifier reached 98.58% accuracy with six features and 97.52% with only four, while the comparison SVC could reach up to 100% in clean baseline tests. The random forest's value lay in its greater robustness under noisy conditions rather than in dominating peak accuracy. Industrial reality reinforces the same trend: a 2023 study by de las Morenas et al. on broken rotor bars and bearings under cyber-physical-system constraints found that simple decision trees reached 95.8% accuracy and 96.8% recall, outperforming heavier random forests (86.6%) and SVMs (71.4%) on their data, and the model was deployed on an affordable Arduino edge platform.

5.4. Industrial deployment and edge constraints

Several papers explicitly trade peak accuracy for deployability. Khalil and Rostam (2024) built a semi-automated maintenance system for cement-plant rotating machinery using FFT energy features and an MLP-plus-SVC stacking ensemble that reached 93.24% accuracy, well above the 86% baseline single SVC, with the explicit constraint that the pipeline had to remain interpretable for plant engineers (Khalil & Rostam, 2024). Wang et al. (2022) implemented a real-time vibration analyzer on a Raspberry Pi with a frequency-domain integrator that extracts velocity and displacement from a single acceleration sensor (Wang et al., 2022).

5.5. Explainability for rotating-machinery diagnosis

Two papers in this section apply XAI explicitly to rotating-machinery classifiers. Brusa et al. (2023) integrated SHAP with SVM and KNN for rolling-element bearings, using SHAP to identify five top features common to both classifiers (p3, p6, p10, p18, p19) and demonstrating that reduced-feature models retained equivalent performance to the full feature set. Mey et al. (2022) ran LIME, Grad-CAM, and LRP-Z on a CNN trained on Fourier-transformed unbalance data, hitting 99.66% accuracy and using the XAI methods to evaluate which explanation tool produced the most physically coherent attributions on cyclic time-series. Their conclusion that Grad-CAM and LRP-Z are individually

flawed but complementary is consistent with what Section 8 finds across the broader XAI literature.

5.6. What the systematic reviews in this folder say

Several review papers in this notebook bracket the empirical work. Lv et al. (2022) surveyed dozens of models for early fault prognosis, highlighting that deep learning's appeal is its powerful adaptive feature extraction and end-to-end characteristics, with transfer learning and GAN-based data generation specifically addressing the lack of labelled fault data (Lv et al., 2022). Tama et al. (2022)'s PRISMA-guided systematic review reports that CNNs (39%) and stacked autoencoders (28.8%) are the dominant deep architectures for rotating-machinery fault diagnosis, with graph neural networks (20.3%) emerging as the third most common (Tama et al., 2022). Kriston (2021) provided a precise scoring system that ranked STFT (3.6) and cepstrum (3.5) as the most suitable transforms for fault detection across industrial machines, gearboxes, bearings, and diesel engines (Kriston, 2021).

6. Structural health monitoring of civil infrastructure

Structural Health Monitoring (SHM) — bridges, buildings, dams, towers, offshore platforms — is the second major application domain in this collection, with 45 papers covering applications and methods. The patterns differ from those in rotating-machinery diagnosis. Operational Modal Analysis (OMA) remains a central tool because civil structures are typically driven by ambient excitation rather than by controlled inputs, and the ML literature builds on top of OMA outputs as much as on raw signals. Environmental and operational variability — temperature, humidity, traffic load — confound damage detection in ways that bearing diagnostics rarely face. The recent trajectory of the field has been to layer CNN, LSTM, and hybrid architectures on top of classical modal analysis, then add unsupervised and Bayesian methods to handle environmental confounders.

6.1. Operational Modal Analysis as the foundation

OMA — Frequency Domain Decomposition (FDD), Stochastic Subspace Identification (SSI-PC, SSI-DATA), and their variants — remains the working baseline. Saidin et al. (2023) reported that FDD and SSI methods produce closely matched modal estimates on bridge SHM data, with Modal Assurance Criterion values approaching 1, and that the correlation between their predicted stiffness and theoretical stiffness reached $R^2 = 0.9996$ with under 10% discrepancy (Saidin et al., 2023). Cardoni et al. (2025) evaluated automated FDD on bridges (Yonghe, PolyU) and cultural-heritage towers (Moletta) and found that an α=0.2 threshold worked across structures with very different

dynamics, with AFDD generally outperforming traditional cov-SSI although both methods struggle with closely spaced torsional modes (Cardoni et al., 2025). These OMA-derived parameters — natural frequencies, mode shapes, damping ratios — are what most ML pipelines actually consume.

6.2. CNN and CNN-LSTM hybrids for damage and anomaly classification

The dominant deep-learning recipe for SHM mirrors the one used in rotating machinery: convert a vibration signal or its OMA derivatives into a 2D representation, then apply a CNN or CNN-LSTM hybrid. Zhang et al. (2023) classified four classes of bridge-SHM data anomalies (outlier, square, missing, minor) using a CNN-LSTM whose input was a two-dimensional matrix of the residual signal and the PSD sequence (raw data was replaced by these extracted features to reduce input size), reaching 97.87% validation accuracy with 99.8% precision on the missing-pattern class, and ran faster than a comparable bidirectional LSTM (Zhang et al., 2023). A 2025 paper on civil-engineering damage detection reached 96.2% accuracy and an F1-score of 0.95 with a CNN-RNN hybrid on benchmark datasets, sustaining 93.4% under noisy conditions. A 2024 multi-class anomaly study on bridges (trend, drift, minor, missing) hit 93.3% overall test accuracy but flagged a real weakness: the drift class only reached 56.8% precision, suggesting that not all anomaly types are equally separable in the 2D image representation.

6.3. Environmental and operational confounding

A distinctive SHM challenge is separating damage signatures from environmental effects (temperature, humidity, wind, traffic). Wang et al. (2022) reviewed the methods used to do this — modal frequency normalisation, PCA, ICA, Mahalanobis distance — and the ML architectures that complement them (SVR, BPNN, RF, LSTM, BiLSTM, K-means, DBSCAN) for long-span bridges and tall buildings (Wang et al., 2022). The headline limitation they identify is that many methods assume environmental measurements are available, which is often not the case. Daneshvar and Sarmadi (2022) addressed the same problem with an information-based unsupervised anomaly-detection method that combines local density, a one-class nearest-neighbour rule, and a probabilistic threshold derived from semi-parametric extreme value theory, validated on large-scale civil structures including the Z-24 benchmark bridge (Daneshvar & Sarmadi, 2022). The Z-24 bridge dataset recurs throughout this literature as the closest thing the field has to CWRU — a structure with documented damage states usable for ML validation.

6.4. Unsupervised and generative approaches

Labelled damage data is rarer in civil SHM than in rotating-machinery research, because civil structures cannot easily be deliberately damaged for training data. This drives an unusual prevalence of unsupervised and generative methods.

Vailati et al. (2024) and Deng et al. (2023) both contribute as review papers that synthesise the SHM landscape: Vailati et al. (2024) summarises work by Yang and Liu showing that natural-frequency sensitivity excludes false positives caused by noise more effectively than MREU or algebraic baselines on cable-stayed bridges and steel frames, while Deng et al. (2023) reviews the CNN-based semantic recovery framework of Jiang et al. and other applications of GANs and BPNNs to Bridge Health Monitoring (Vailati et al., 2024; Deng et al., 2023). The trend is consistent: SHM applications are pushing harder on unsupervised, autoencoder-based, and generative methods than rotating-machinery diagnosis is.

6.5. Wind-turbine blades and beams

Several papers in this folder address wind-turbine blade SHM specifically. Ogaili et al. (2022) reached 97% classification accuracy on wind-turbine blade crack-tip, mid-span, and root damage using KNN combined with ReliefF feature ranking, well above the ~79.9% achieved by Chi-square-ranked features with SVM, while Information Gain ranking yielded the lowest precision of the feature-selection methods tested (Ogaili et al., 2022). Dang and Nguyen (2023) introduced an mFF-SHM framework combining graph learning, triplet loss, and AdaBoost on a 2D steel frame structure with semi-rigid connections. The model reached 95% accuracy and 94% F1-score, ran twice as fast as a raw-data LSTM, and retained >90% F1-score with 10% of the data missing thanks to a triplet-loss self-supervised pretext task (Dang & Nguyen, 2023).

6.6. Transfer learning and emerging directions

Transfer learning is the single most impactful recent direction in SHM. Zar et al. (2023) reviewed work by Teng et al. (2023) reporting that 1D-CNN transfer learning boosted damage-detection accuracy by up to 47% while approximately halving training convergence time (Zar et al., 2023). The cross-cutting finding is consistent with what Section 4 discusses for the deep-learning literature more broadly: when labelled data is scarce in the target domain, transferring features from a richer source domain is often more impactful than tuning the architecture. Several papers in this folder explore newer architectures specifically for SHM. Li et al. (2023) combined K-means with SSI and SSA for automated parameter identification on masonry towers, high arch dams, and bridges, with explicit attention to Bayesian model updating in the presence of model errors. Earlier work by Tao et al. (2020) cited in this folder's 2024 reviews outlines hybrid STFT + GAN methods for unsupervised fault diagnosis in rolling bearings, with separate studies addressing TBM main beams and wind-turbine planetary gears under different methods.

6.7. The reviews of the SHM field

Three substantial reviews bracket the empirical work. Avci et al. (2021) tracked the transition from ARX/ARMA parametric methods to autoencoders and 1D/2D CNNs, highlighting CNN success on concrete box-girder bridges, a steel grandstand simulator, and mass-spring/beam systems (Avci et al., 2021). Zhang et al. (2022) reviewed both signal-processing transforms (DFT, FRF, SFRF, FDD, MUSIC, CWT) and ML classifiers (SVM, OC-SVM, RF, KNN, decision trees, ANN, 1D/2D-CNN, autoencoders), demonstrating that frequency-domain CNNs reduce computational time compared with time-domain networks (Zhang et al., 2022). Toh and Park (2020) survey two complementary results from the literature they review: a Wu et al. 1D-CNN+SVM method handling 358 samples/sec for rotating machinery, and a Khatir et al. nMSEDI damage index that computes in 0.28 seconds versus 60 seconds for the LFCR baseline on civil structures (Toh & Park, 2020).

7. Reviews, surveys and comparative synthesis from the field

A handful of recent review papers provide the most useful global view of vibration-plus-ML, and they converge on a small number of recurring observations. Five reviews appear in this collection, covering structural dynamics broadly, machine condition monitoring, belt conveyor idlers specifically, rotating-machinery diagnosis end-to-end, and the underlying sensor hardware. Together they sketch the state of the field with a consistency that the empirical papers individually cannot.

7.1. Structural dynamics and vibroacoustics

Cunha et al. (2023) provide the broadest scope, surveying ML applications across structural dynamics and vibroacoustic problems. Their central observation is that ML is most valuable where mechanistic models are computationally expensive or unknown, since the algorithm can model the underlying physical phenomenon from sampled data alone. They make a precise version of the small-data-versus-big-data distinction that recurs throughout this literature: traditional ML suits small datasets where domain experts can hand-craft features, while deep learning is the better fit for large-scale data because it automates feature extraction (Cunha et al., 2023). They identify the limited availability of labelled data and the influence of environmental and operational variability as the two main bottlenecks, and point to Physics-Guided Machine Learning (PGML) and hybrid models as the most active responses. The contribution is a roadmap for selecting an ML approach across SHM, active noise/vibration control, and surrogate modelling for product design.

7.2. Machine condition monitoring

Mohd Ghazali and Rahiman (2021) review the full pipeline of machine condition monitoring: data acquisition instruments, feature extraction, and AI-based fault recognition. They report that computer-based analyzers built around DSP or FPGA components paired with MEMS accelerometers are becoming the cost-effective alternative to standalone analyzers (Mohd Ghazali & Rahiman, 2021). On the algorithm side, SVM remains the most frequently used method for fault classification because of its accuracy and low computational cost, but DNNs and CNNs are gaining traction for their automated feature extraction. The reviewers single out generalisation as the major unsolved problem: models trained on historical data in controlled environments often fail when applied to new datasets or industrial settings with different operating conditions. The contribution is the integrated coverage of hardware and algorithms, which most reviews treat separately.

7.3. Belt conveyor idler fault detection

Alharbi et al. (2023) narrow the scope to a single but commercially important domain: belt conveyor idlers. The review confirms that vibration and acoustic signals are highly effective for early fault detection in idlers and that vibration-based methods dominate the published literature. Deep learning approaches — CNNs and autoencoders — improve detection accuracy by bypassing manual feature extraction and modelling the highly non-linear patterns produced by mechanical degradation under load (Alharbi et al., 2023). The two persistent challenges are the noisy operational environment of conveyors and the scarcity of large labelled public datasets containing idler fault signatures. The contribution is the first comprehensive review dedicated to this specific application class.

7.4. Rotating-machinery pipeline audit

Bagri et al. (2024) deliver the most uncomfortable finding of the five. Their systematic examination of the entire vibration analysis pipeline for rotating machinery — preprocessing, transformation, classification, prognosis — produces three concrete results. Time-scale decomposition methods (EMD and VMD) coupled with time-frequency analysis are essential for handling the non-stationary content of mechanical signals. SVM, CNN, LSTM, KNN, and random forest are the five most frequently solicited AI algorithms for fault detection (Bagri et al., 2024). And, critically, only around 10% of recent studies provide real-time experimental validation, and many do not disclose enough about their preprocessing parameters to be reproducible. The contribution is an honest audit of reproducibility, and the result raises a serious flag about how to read the high accuracies reported throughout the rest of this literature.

7.5. Sensor hardware as the underlying constraint

Ma et al. (2025) review the sensor side rather than the algorithm side. They classify vibration sensors into seven main physical mechanisms — piezoelectric, electromagnetic, piezoresistive, triboelectric, photonic, capacitive, and related types — and note that hybrid sensors combining two transduction principles are increasingly used to acquire diverse spectra and multi-dimensional fault characteristics simultaneously (Ma et al., 2025). Their discussion of integration with AI covers early warning systems and digital-twin frameworks for predictive maintenance among other directions, without singling them out as more active than the rest.

8. Hybrid models, physics-infromed learning and explainable AI

Three threads in the recent literature address shortcomings of pure data-driven ML for vibration analysis. Hybrid models combine classical signal processing with deep learning to retain the inductive biases of mechanical understanding. Physics-informed neural networks (PINNs) embed governing equations directly into the learning objective. Explainable AI (XAI) makes black-box models interpretable to engineers who need to act on their outputs. None of these threads is mature yet, but together they answer the criticism that pure CNN pipelines are unaccountable systems whose 99% accuracy on a benchmark dataset gives no reason to trust them on a new machine.

8.1. Hybrid signal-processing and deep-learning pipelines

The most common kind of hybrid model in this collection is a pipeline where a classical transform (CWT, EMD, VMD) preprocesses the signal and a CNN or CNN-LSTM performs the classification. Mousavi et al. (2025) ran VMD reconstruction before a 1D-CNN+LSTM for monopile offshore-wind-turbine damage, gaining 3.3 percentage points (94.72% → 98.05%) over raw data (Mousavi et al., 2025). Hamdaoui et al. (2023) combined variational mode decomposition and wavelet thresholding with CWT scalograms feeding a custom three-layer CNN for rolling bearings under varying speed, reaching 96.09%–99.70% accuracy (Hamdaoui et al., 2023). The pattern is that the classical transform supplies the inductive bias the network would otherwise have to learn from scratch, and the network supplies the discriminative power the transform alone cannot.

A more interesting recent hybrid combines forecasting and diagnosis. Rihi et al. (2024) ran an LSTM forecaster on the raw vibration signal, converted the forecasted sequences to MFCC spectrograms, and classified those with a CNN. The system reached 99.95% accuracy and an AUC-ROC of 1.0 on a real mining-mill dataset (Rihi et al., 2024). The proactive intervention advantage is that

diagnostic conclusions can be reached before a fault has fully developed in the present signal.

8.2. Physics-informed neural networks

A 2025 bibliometric review identified PINNs as one of the active frontiers for predictive maintenance of critical industrial assets including bearings, compressors, hydraulic pumps, and conveyor belts. The integration challenge with PINNs is the weighting between physics-based and data-driven loss terms, and the absence of standardised benchmarks for comparing PINN-based methods against pure data-driven baselines. Cunha et al. (2023)'s broader review framed Physics-Guided Machine Learning (PGML) and hybrid models as the most active response to the interpretability and data-scarcity problems that limit pure ML for structural dynamics (Cunha et al., 2023). The empirical evidence in this collection on PINNs is thinner than on data-driven CNNs, which means PINN performance claims should be read as preliminary rather than established.

8.3. Explainable AI for vibration analysis

XAI is appearing in a growing subset of vibration-ML papers, although the broader literature still treats deep networks as black boxes. The most commonly used tools in this collection are SHAP, Grad-CAM, LIME, and LRP-Z, each of which produces a different kind of explanation.

Brusa et al. (2023) integrated SHAP with SVM and KNN classifiers for rolling-element-bearing fault diagnosis, identifying five top features common to both classifiers (p3, p6, p10, p18, p19) and showing that reduced-feature models retained equivalent performance to the full feature set (Brusa et al., 2023). The contribution is operational: SHAP-based feature reduction simplifies the deployed model without sacrificing accuracy.

Łuczak (2024) applied Grad-CAM, LIME, and occlusion together to a CWT-CNN classifier for fan-blade imbalance, demonstrating that the three methods provide complementary rather than redundant information about the network's decision-making (Łuczak, 2024). Mey et al. (2022) ran LIME, Grad-CAM, and LRP-Z on a CNN trained on Fourier-transformed unbalance data, reaching 99.66% accuracy and concluding that the three methods were individually flawed but together produced a coherent explanation of the network's attention on cyclic time-series (Mey et al., 2022). The cross-cutting observation from these papers is that no single XAI method is sufficient — combinations are needed because each method's failure mode is different.

8.4. Why XAI matters specifically for vibration ML

The need for interpretability in vibration analysis is sharper than in many other ML applications. Plant engineers and structural-health-monitoring practitioners need to know why a model flagged a fault before they take action: shutting down a turbine costs millions, and a model that produces high accuracy on a

benchmark but no auditable rationale offers no basis for decision. Cunha et al. (2023) flagged the limited availability of labelled data and environmental and operational variability as the top two open problems in structural-dynamics ML, with interpretability also identified as a recurring drawback of black-box models (Cunha et al., 2023). The 2023 paper on predictive maintenance with random forests over LSTMs made the same point operationally: it argued that transparent ML approaches like RF yield better Decision Support Systems than black-box DL, even when accuracy is slightly lower.

8.5. Self-supervised and unsupervised hybrid approaches

A related thread treats data scarcity itself as a learning signal. Dang and Nguyen (2023) used self-supervised learning (triplet loss) within an mFF-SHM framework that combined graph learning and AdaBoost, retaining >90% F1-score even with 10% of the data missing on 2D steel-frame structures (Dang & Nguyen, 2023). Radicioni et al. (2025) trained convolutional and variational autoencoders on SST spectrograms for anomaly detection without labelled fault data (Radicioni et al., 2025). The unifying idea is that the hybrid model — classical preprocessing plus unsupervised learning — avoids the labelled-data bottleneck without sacrificing the rigor of classical analysis.

9. Open challenges and future directions

The field is healthy by most measures. Sensors are cheaper, datasets are larger, deep-learning architectures routinely reach 99% on canonical benchmarks, and explainability tools have moved from research add-ons to standard practice. The remaining problems are not transducer-physics or algorithm-architecture problems. They are reproducibility, deployment, transfer, and trust. This section identifies the six that recur most consistently in the empirical and review literature in this collection.

9.1. Reproducibility and reporting standards

Bagri et al. (2024) audited the rotating-machinery vibration-analysis pipeline and found that only about 10% of recent studies provide real-time experimental validation, and many do not disclose enough about preprocessing parameters to be reproducible (Bagri et al., 2024). The headline accuracies in this literature should be read against that constraint. Two consequences follow. First, the same methods reported with different preprocessing pipelines may produce materially different results on the same data, which makes head-to-head comparison across papers difficult. Second, the absence of standardised reporting means that even when a paper reports 99% accuracy on a dataset, a practitioner cannot reliably reproduce the result without contacting the authors. The fix is cultural rather than technical, but computer vision solved a similar problem partly through

community pressure on benchmarks and code release, vibration ML does not yet have the equivalent.

9.2. Benchmark saturation and cross-machine generalisation

The CWRU bearing dataset and Z-24 bridge dataset have become so dominant that accuracies on them have effectively saturated: multiple methods reach 99% or 100% (Abdullah et al., 2025; Hamdaoui et al., 2023; Section 4). Further accuracy improvements on these datasets are no longer informative. What the field needs is harder evaluation: cross-machine generalisation, transfer between operating conditions, and accuracy under realistic noise and missing-data conditions. The transfer-learning results of Wang et al. (2020) on cross-condition bearing tasks (Wang et al., 2020) and Zar et al. (2023) on SHM transfer (Zar et al., 2023) point in this direction, but cross-machine evaluation is still rare in the published literature. Mohd Ghazali and Rahiman (2021) called generalisation out explicitly as the major unsolved problem in machine condition monitoring (Mohd Ghazali & Rahiman, 2021).

9.3. Labelled-data scarcity and unsupervised methods

Labelled fault data is hard to acquire — civil structures cannot easily be deliberately damaged for training, and industrial machinery typically has very few documented fault states relative to the variety of failure modes that can occur. The two responses in this collection are transfer learning (covered above) and unsupervised methods (autoencoders, GANs, clustering, anomaly detection). Radicioni et al. (2025) trained convolutional autoencoders on SST spectrograms for unsupervised anomaly detection (Radicioni et al., 2025). Daneshvar and Sarmadi (2022) used an information-based anomaly-detection method combining local density, a one-class nearest-neighbour rule, and a probabilistic threshold derived from semi-parametric extreme value theory, validated on the Z-24 benchmark bridge (Daneshvar & Sarmadi, 2022). These approaches work but produce less precise diagnostic outputs than supervised classifiers; they identify anomalies rather than naming fault classes.

9.4. Edge AI and on-device inference

Most of the deep-learning pipelines reviewed here assume server-class compute for both training and inference. For deployment at scale across distributed industrial assets — fleets of motors, networks of bridge sensors, agricultural and mining equipment — that assumption is uneconomic. Łuczak (2024) demonstrated a CWT-CNN classifier processing each fault classification in under 50 ms with Grad-CAM interpretability layered on top (Łuczak, 2024). A 2024 IoT-driven framework reached 92% accuracy with 150 ms latency over LoRaWAN. The gap between these and the heavy-compute baseline papers is where the practical research lies: quantised CNNs, hardware accelerators for

FFT and wavelet computation, lightweight transformer variants, and on-device inference frameworks. The literature is only partially addressing this.

9.5. Multi-modal and sensor fusion

Many real applications carry multiple complementary signals — vibration, acoustic, current, temperature — and fusing them often beats any single modality. Han et al. (2021) demonstrated this for CNC tool-wear regression, where fused audio + vibration outperformed either alone (Han et al., 2021). Ma et al. (2025) noted that hybrid sensors combining piezoelectric and triboelectric mechanisms simultaneously acquire diverse spectra and multi-dimensional fault characteristics (Ma et al., 2025). The unresolved problem is synchronisation and calibration across heterogeneous sensors, particularly at the temporal granularity required for high-frequency mechanical signals. The architecture also matters: most published fusion approaches concatenate features at a late stage, but earlier-stage fusion (joint embedding) is under-explored.

9.6. Operator trust and explainability

A subtler but increasingly important challenge is whether plant operators trust the outputs of black-box deep-learning models. Mey et al. (2022) found that no single XAI method is sufficient and that complementary methods together produce more credible explanations (Mey et al., 2022). Cunha et al. (2023) framed interpretability as one of the top two open problems in structural-dynamics ML (Cunha et al., 2023). The next generation of pipelines will likely combine post-hoc XAI with physics-informed objectives that make the model's reasoning more closely aligned with mechanical understanding from the outset.

9.7. Standardisation

A final challenge is the absence of standard performance reporting and standard test rigs. Bagri et al. (2024) call for it indirectly through their reproducibility audit (Bagri et al., 2024). the lack of a common evaluation protocol limits the conclusions that can be drawn from the cumulative literature (Tama et al., 2022; Avci et al., 2021; Zhang et al., 2022). The field would benefit from a small set of standard rigs, standardised cross-condition test splits, and uncertainty-budget reporting analogous to what exists in metrology.

10. Conclusion

Machine learning for vibration analysis has matured from a niche research direction into a tightly coupled stack of signal representations, classical and deep-learning models, and application-specific pipelines. The papers reviewed here collectively support three conclusions.

First, the dominant pipeline in modern practice is signal → time-frequency representation (most often a CWT scalogram) → CNN or CNN-LSTM hybrid → optional XAI for interpretation. This pattern reaches ≥99% accuracy on canonical bearing datasets routinely enough that benchmark accuracy is no longer the differentiating metric. Within rotating-machinery diagnosis it is established as the default. Within SHM it now coexists with Operational Modal Analysis, with modern deep-learning pipelines increasingly bypassing manual modal identification altogether and mapping raw vibration data to damage states end-to-end.

Second, classical machine learning has not been displaced. SVM, random forest, and decision trees still produce competitive accuracy when the dataset is small, the features are well-matched to the physics, and the deployment platform is compute-constrained. The choice between classical and deep approaches is not a question of which is "better" in general but of which fits the labelling regime, the compute budget, and the interpretability requirements of the operating context.

Third, the remaining open problems are no longer about transducer physics or model architecture. They are about reproducibility, cross-machine generalisation, edge deployment, multi-modal sensor fusion, and operator trust. Bagri et al. (2024)'s audit of preprocessing transparency, Wang et al. (2020)'s domain-adaptation work, Łuczak (2024)'s sub-50-ms edge deployment, and Mey et al. (2022)'s multi-method XAI comparisons are exemplars rather than routine practice.

The field has acquired the sensors and algorithms it needs. The next phase of progress will come from the discipline of integrating them at scale: standardised benchmarks that test cross-source generalisation, reporting practices that make results reproducible, deployment pipelines that fit edge hardware, and explanation tools that operators trust.

References

[1] Abdullah, A. T., Hussein, H. M., Sabeeh, R. S., et al. (2025). A Comprehensive Review of Machine Learning Algorithms for Fault Diagnosis and Prediction in Rotating Machinery. Journal's University of Babylon for Engineering Sciences (JUBES), 33(4), 111-142.

[2] Alharbi, F., Luo, S., Zhang, H., et al. (2023). A Brief Review of Acoustic and Vibration Signal-Based Fault Detection for Belt Conveyor Idlers Using Machine Learning Models. Sensors, 23(4), 1902. https://doi.org/10.3390/s23041902

[3] Ali, M. S., Mohite, D. B., Gandhe, G. R., et al. (2025). Deep learning-based fault detection in civil engineering structures: A computational mathematics approach using vibration data. International Journal of Applied Mathematics, 38(3s), 153-168.

[4] Ali, F., Donyaparastlivari, L., Alghamaz, M., et al. (2025). Integrating Piezoelectric Sensors for Enhanced Failure Prediction of Residential Buildings in Hurricanes. Sensors, 25(6), 2269. 10.3390/s25062269

[5] Altaf, M., Akram, T., Khan, M. A., et al. (2022). A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals. Sensors, 22(5), 2012. 10.3390/s22052012

[6] Atmaja, B. T., Ihsannur, H., Suyanto, et al. (2022). Lab-scale Vibration Analysis Dataset. arXiv preprint. 10.48550/arXiv.2212.14732

[7] Au, S. K., Brownjohn, J. M., Li, B., et al. (2021). Understanding and managing identification uncertainty of close modes in operational modal analysis. Mechanical Systems and Signal Processing, 147(1), 107018. 10.1007/s13349-021-00499-4

[8] Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. Mechanical Systems and Signal Processing, 147, 107077. https://doi.org/10.1016/j.ymssp.2020.107077

[9] Bagri, I., Tahiry, K., Hraiba, A., et al. (2024). Vibration Signal Analysis for Intelligent Rotating Machinery Diagnosis and Prognosis: A Comprehensive Systematic Literature Review. Vibration, 7(4), 1013-1062. https://doi.org/10.3390/vibration7040054

[10] Brusa, E., Cibrario, L., Delprete, C., et al. (2023). Explainable AI for Machine Fault Diagnosis: Understanding Features' Contribution in Machine Learning Models for Industrial Condition Monitoring. Applied Sciences, 13, 2038. 10.3390/app13042038

[11] Cardoni, A., Elahi, A. R., & Cimellaro, G. P. (2025). The role of reinforced concrete roofs in the seismic performance of masonry buildings. Engineering Structures, 323(1), 119210. 10.1016/j.engstruct.2024.119210

[12] Chen, H. Y., & Lee, C. H. (2020). Vibration Signals Analysis by Explainable Artificial Intelligence (XAI) Approach: Application on Bearing

Faults Diagnosis. IEEE Access, 8, 134246-134256. https://doi.org/10.1109/ACCESS.2020.3006491

[13] Cunha, B. Z., Droz, C., Zine, A.-M., et al. (2023). A Review of Machine Learning Methods Applied to Structural Dynamics and Vibroacoustic. Mechanical Systems and Signal Processing.

[14] Daneshvar, M. H., & Sarmadi, H. (2022). Unsupervised learning-based structural health monitoring by Mahalanobis distance metrics. Engineering Structures, 1(1), 1-18. 10.1016/j.engstruct.2022.114102

[15] Dang, V. H., & Nguyen, T. T. (2023). Robust Vibration Output-only Structural Health Monitoring Framework Based on Multi-modal Feature Fusion and Self-learning. Periodica Polytechnica Civil Engineering, 67(2), 416-430. 10.3311/PPci.21756

[16] Deng, Z., Huang, M., Wan, N., et al. (2023). The Current Development of Structural Health Monitoring for Bridges: A Review. Buildings, 13(6), 1360. 10.3390/buildings13061360

[17] Gómez Muñoz, C. Q., et al. (2025). Structural health monitoring through deep learning: a study on scalogram-based vibration signal. Journal of Low Frequency Noise, Vibration and Active Control, 45(1), 405-419.

[18] Hamdaoui, H., Ngiejungbwen, L. A., Gu, J., et al. (2023). Improved signal processing for bearing fault diagnosis in noisy environments using signal denoising, time–frequency transform, and deep learning. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 45, 576. 10.1007/s40430-023-04471-9

[19] Han, S., Mannan, N., Stein, D. C., et al. (2021). Classification and Regression Models of Audio and Vibration Signals for Machine State Monitoring in Precision Machining Systems. Journal of Manufacturing Systems.

[20] Hassan, I. U., Panduru, K., & Walsh, J. (2024). An In-Depth Study of Vibration Sensors for Condition Monitoring. Sensors, 24, 740. 10.3390/s24030740

[21] He, M., & He, D. (2020). A new hybrid deep signal processing approach for bearing fault diagnosis using vibration signals. Neurocomputing, 396, 542–555. 10.1016/j.neucom.2018.12.088

[22] Jang, J.-G., Noh, C.-M., Kim, S.-S., et al. (2022). Vibration data feature extraction and deep learning-based preprocessing method for highly accurate motor fault diagnosis. Journal of Computational Design and Engineering, 10(1), 204. 10.1093/jcde/qwac128

[23] Janssens, O., Slavkovikj, V., Vervisch, B., et al. (2016). Convolutional Neural Network Based Fault Detection for Rotating Machinery. Journal of Sound and Vibration, 377, 331–345. 10.1016/j.jsv.2016.05.027

[24] Kafeel, A., Aziz, S., Awais, M., et al. (2021). An Expert System for Rotating Machine Fault Detection Using Vibration Signal Analysis. Sensors, 21(22), 7587. 10.3390/s21227587

[25] Khalil, A. F., & Rostam, S. (2024). Machine Learning-based Predictive Maintenance for Fault Detection in Rotating Machinery: A Case Study.

Engineering, Technology & Applied Science Research, 14(2), 13181-13189. 10.48084/etasr.6813

[26] Kriston, B. J., & Jálics, K. (2021). Application of vibro-acoustic methods in failure diagnostics. Journal of Physics: Conference Series, 1935, 012002. 10.1088/1742-6596/1935/1/012002

[27] Lv, Y., Zhao, W., Zhao, Z., et al. (2022). Vibration signal-based early fault prognosis: Status quo and applications. Advanced Engineering Informatics, 52, 101609. 10.1016/j.aei.2022.101609

[28] Ma, L., Li, Z., Yang, S., et al. (2025). A Review on Vibration Sensor: Key Parameters, Fundamental Principles, and Recent Progress on Industrial Monitoring Applications. Vibration, 8(4), 56. https://doi.org/10.3390/vibration8040056

[29] Martins, A., Fonseca, I., Farinha, J. T., et al. (2024). Prediction maintenance based on vibration analysis and deep learning — A case study of a drying press supported on a Hidden Markov Model. Applied Soft Computing, 163, 111885. https://doi.org/10.1016/j.asoc.2024.111885

[30] Mey, O., & Neufeld, D. (2022). Explainable AI Algorithms for Vibration Data-Based Fault Detection: Use Case-Adadpted Methods and Critical Evaluation. Sensors, 22, 9037. 10.3390/s22239037

[31] Mohd Ghazali, M. H., & Rahiman, W. (2021). Vibration Analysis for Machine Monitoring and Diagnosis: A Systematic Review. Shock and Vibration, 2021, 9469318. https://doi.org/10.1155/2021/9469318

[32] Mousavi, Z., Feng, R. M., Farhadi, M., et al. (2025). Damage detection based on proposed deep CNN-LSTM network. Applied Ocean Research, 165, 104856. https://doi.org/10.1016/j.apor.2025.104856

[33] Ogaili, A. A. F., Jaber, A. A., & Hamzah, M. N. (2022). An intelligent fault diagnosis approach based on machine learning. Open Engineering, 1(1), 1-16. 10.1515/cls-2022-0214

[34] Radicioni, L., Bono, F. M., & Cinquemani, S. (2025). Vibration-Based Anomaly Detection in Industrial Machines: A Comparison of Autoencoders and Latent Spaces. Machines, 13(2), 139. https://doi.org/10.3390/machines13020139

[35] Rahim, S., Najmi, A., Samin, R., et al. (2024). The Effects of Signal Processing Techniques in Damage Detection and Structural Health Monitoring. Journal of Physics: Conference Series, 2721(1), 012022. 10.1088/1742-6596/2721/1/012022

[36] Rihi, A., Baïna, S., Mhada, F. Z., et al. (2024). Innovative predictive maintenance for mining grindingmills: from LSTM-based vibration forecasting to pixel-based MFCC image and CNN. The International Journal of Advanced Manufacturing Technology, 135, 1271-1289. https://doi.org/10.1007/s00170-024-14588-3

[37] Saidin, S. S., Kudus, S. A., Anuar, M. A., et al. (2023). Vibration-based structural damage identification. Case Studies in Construction Materials, 18(1), e01752. 10.1016/j.cscm.2023.e01752

[38] Shandookh, A. A., Ogaili, A. A. F., & Al-Haddad, L. A. (2024). Failure analysis in predictive maintenance: Belt drive diagnostics with expert systems and Taguchi method for unconventional vibration features. Heliyon, 10, e34202. 10.1016/j.heliyon.2024.e34202

[39] Tama, B. A., et al. (2022). Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. Artificial Intelligence Review. 10.1007/s10462-022-10293-3

[40] Toh, G., & Park, J. (2020). Review of Vibration-Based Structural Health Monitoring Using Deep Learning. Applied Sciences, 10(5), 1680. 10.3390/app10051680

[41] Wang, X., Shen, C., Xia, M., et al. (2020). Fault diagnosis approach based on transfer learning. Reliability Engineering and System Safety, 202, 107050. 10.1016/j.ress.2020.107050

[42] Wang, K., Xu, Z.-J., Gong, Y., et al. (2022). Mechanical Fault Prognosis through Spectral Analysis of Vibration Signals. Algorithms, 15(3), 94. 10.3390/a15030094

[43] Wang, Z., Yang, D. H., Yi, T. H., et al. (2022). Eliminating environmental and operational effects on structural modal frequency: A comprehensive review. Structural Control and Health Monitoring, 29(11), e3073. 10.1002/stc.3073

[44] Zar, A., Li, S., Hussain, Z., et al. (2023). Towards vibration-based damage detection of civil engineering structures: overview. International Journal of Mechanics and Materials in Design, 19(1), 634-662. 10.1007/s10999-023-09692-3

[45] Zhang, C., Mousavi, A. A., Masri, S. F., et al. (2022). Vibration feature extraction using signal processing techniques for structural health monitoring: A review. Mechanical Systems and Signal Processing, 177(1), 109175. 10.1016/j.ymssp.2022.109175

[46] Zhang, J., Zhang, J., & Wu, Z. (2023). Structural Vibration Data Anomaly Detection Based on Multiple Data Patterns. Structural Control and Health Monitoring, 2023(1), 3906180. 10.1155/2023/3906180

[47] de las Morenas, J., Moya-Fernández, F., & López-Gómez, J. A. (2023). The Edge Application of Machine Learning Techniques for Fault Diagnosis in Electrical Machines. Sensors, 23, 2649. 10.3390/s23052649

[48] Łuczak, D. (2024). Machine Fault Diagnosis through Vibration Analysis: Continuous Wavelet Transform with Complex Morlet Wavelet and Time–Frequency RGB Image Recognition via Convolutional Neural Network. Electronics, 13(2), 452. https://doi.org/10.3390/electronics13020452

Integrating circular economy principles into quality management systems across global supply chains: A systematic literature review, gap analysis, and conceptual framework

Emile NASEEM[1], Viktor MEDINA[2]
[1]Doctoral Program of Mechanical Engineering, MATE, Gödöllő
[2]Department of Engineering Management, Institute of Technology, MATE, Gödöllő

Abstract

The circular economy is a new way to achieve sustainable development, but it faces challenges in managing quality across supply chains. I present a review of 341 studies from three areas of research. These areas are quality management, circular economy, and supply chain management. This review aims to examine how quality management systems can be changed for supply chains in a circular economy. I used a search protocol to find studies from 2000 to 2025 on Scopus, Web of Science, Google Scholar, and ProQuest Dissertations. The review found a major gap in research. There are no complete frameworks to integrate circular economy principles into quality management systems across supply chains. Six areas where research is lacking were identified. These areas are frameworks for quality management in circular supply chains, coordination of quality across supply chains, metrics to measure quality, guidance on implementing quality management, testing of quality management systems, and standards for quality management. I developed a framework that shows how quality management systems can be adapted for circular supply chains, considering the context, capabilities of organizations, and institutions. This framework addresses organizational, supply chain, and institutional levels, and it provides a basis for further research. It has implications for theory, practice, and standards development in quality management and circular economy implementation. I also outline a plan for future research in this area.

Keywords:

circular economy; quality management systems; circular supply chain management; ISO 9001; remanufacturing quality; recycled material quality; supply chain quality coordination; systematic literature review

1. Introduction

The concept of the circular economy (CE) has witnessed unprecedented recognition as a viable alternative to the conventional linear 'take, make, dispose' approach and has been recognized as a potential solution for decoupling economic growth from the depletion of resources and environmental degradation. Global initiatives such as the European Union's 2015 and 2020 Circular Economy Action Plans have accelerated the development and implementation of circular economy practices such as remanufacturing, recycling, refurbishing, and 'product as a service' business models. Despite the unprecedented policy push and market demand for the development and implementation of circular economy practices, the transition has been hindered by issues related to quality management that have been insufficiently addressed by conventional frameworks. Quality management is a cross-cutting issue that affects every aspect of the development and implementation of the circular economy. For instance, products received for remanufacturing are received in a highly uncertain and variable condition that affects the operations and economics of remanufacturing businesses. Recycled materials have inherent properties that affect manufacturing processes. For instance, recycled materials have a batch-to-batch variation of 20-50% in their mechanical properties. Quality management has also been hindered by customer perceptions related to products developed under the circular economy. For instance, despite objective quality being equal or even greater than that of new products, products developed under the circular economy have been associated with a 20-30% discount. Quality management theory and standards such as ISO 9001:2015 were developed with the linear economy concept and were based on a unidirectional supply chain with virgin materials and forward supply chain configurations.

Three research streams are directly relevant to this problem. First, the quality management literature offers strong theoretical foundations based on Deming's System of Profound Knowledge, Juran's Quality Trilogy, and ISO 9001 principles but shows limited engagement with the principles of a circular economy (Foster, 2008; Psomas et al., 2018). Second, the circular economy literature offers a deep understanding of the principles and barriers related to a CE approach but superficially addresses the related concept of quality management while ignoring the strong theoretical foundations of the quality management literature (Kirchherr et al., 2023; Geissdoerfer et al., 2017). Third, the supply chain management literature offers relevant insights into supply chain coordination mechanisms and complexity management but does not address the concept of quality management in the context of a circular supply chain approach (Batista et al., 2018; Farooque et al., 2019).

Most critically, the intersection between the above three research streams is very limited and underdeveloped. For example, while individual research studies have addressed specific issues related to quality management in remanufacturing

(Bansal et al., 2024; Guide & Van Wassenhove, 2009), recycled material quality management issues due to variability (Ragaert et al., 2017; Jiang & Bateer, 2025), and reverse logistics inspection issues (Ferguson et al., 2009), no such framework has been developed that incorporates such isolated research findings into a holistic approach for managing quality issues related to a supply chain approach. This is a problem that must be addressed as a result of the inherent need for coordination and quality management between multiple firms that must exist as a result of a successful circular economy approach (Batista et al., 2018; Garcia-Buendia et al., 2024).

The study's research contributions are fourfold. First, this study presents a holistic, integrative synthesis of the traditionally separate research literatures on circular economy, supply chain management, and quality management, examining points of convergence and divergence that have not been previously examined together. Second, this study articulates six specific research gaps that are interconnected, along with the supporting literature. Third, this study presents a multi-theoretical conceptual framework based on systems theory, the resource-based view, stakeholder theory, institutional theory, and transaction cost economics. Fourth, this study formulates research questions that offer a future research agenda.

The rest of the paper is organized as follows. Section 2 presents the methodology for the systematic literature review. Section 3 presents the thematic synthesis of the literature from the four research domains. Section 4 presents the multi-level conceptual framework. Section 5 presents the implications, limitations, and theoretical contributions of the study. Section 6 presents the research questions that offer the future research agenda, while Section 7 concludes the paper.

2. Methodology

2.1 Research Design

The research design for this study is based on the systematic literature review methodology in accordance with established guidelines for management research (Tranfield et al., 2003; Denyer & Tranfield, 2009). A systematic literature review design was chosen for its ability to produce comprehensive and transparent results in all three of the contributing research domains. This systematic literature review design is based on PRISMA-informed literature review design guidelines for management research, where clinical design requirements are not suitable for management disciplines due to the wide range of theoretical and empirical traditions in quality management, circular economy, and supply chain management (Webster & Watson, 2002).

2.2 Search Strategy

The systematic literature review search for relevant literature in all four of the contributing research domains was conducted through four academic databases: Scopus, Web of Science Core Collection, Google Scholar, and ProQuest Dissertations & Theses Global. The choice of academic databases for systematic literature review in operations management and sustainability research literature review design is in accordance with Tranfield et al. (2003) and Tranfield et al. (2003) recommendations for systematic literature review in operations management and sustainability research literature review design. Three different search strategies were used in combination for literature search in all four of the contributing research domains. Firstly, searches were conducted for literature in each of the contributing research domains individually using keywords such as "quality management systems," "ISO 9001," "supply chain quality management," "circular economy," "closed loop supply chains," "reverse logistics," "remanufacturing," and "sustainable supply chain management." Secondly, literature from all four of the contributing research domains was searched for using Boolean operators and keywords such as ("quality management" OR "quality assurance" OR "ISO 9001") AND ("circular economy" OR "circularity" OR "closed loop" OR "reverse logistics" OR "remanufacturing" OR "product recovery") AND ("supply chain" OR "supply network" OR "value chain"). Thirdly, forward and backward citation tracking of relevant literature from all four of the contributing research domains were conducted for literature not identified through keyword searches (Webster & Watson, 2002).

2.3 Inclusion and Exclusion Criteria

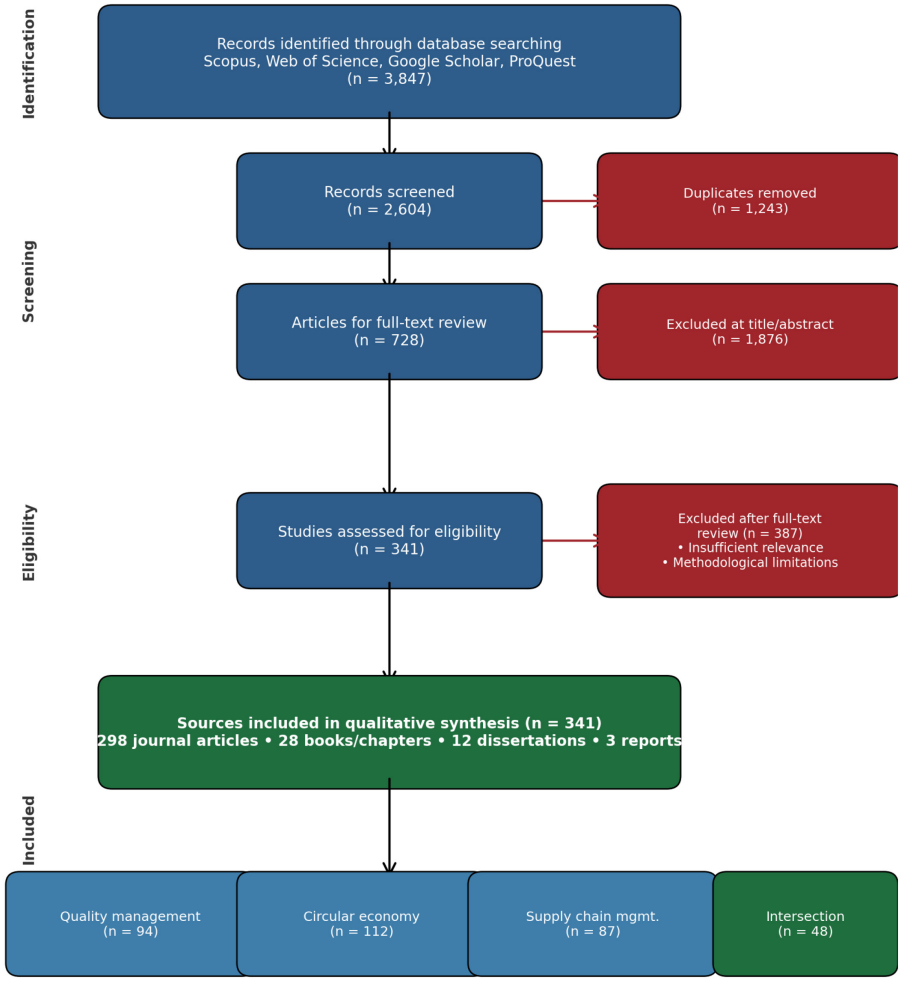
Inclusion criteria for studies were that they were published as a journal article or a doctoral dissertation by an academic publisher or an academic institution recognized by a reputable accreditation body; were written in the English language; focused on the topics of quality management, circular economy, supply chain management, and their intersections; and were theoretically or empirically relevant and rigorous. Exclusion criteria were that the study did not fit any of the inclusion criteria or that it lacked academic rigor and relevance. Seminal works that were published prior to the year 2000 and were foundational to the development of the relevant research streams were included. For example, seminal works by Deming (1994) and Juran (1988) were included as they formed the foundation for the development of the relevant research streams.

2.4 Data Analysis Approach

An integrative synthesis methodology was used for the analysis, which is suitable for an interdisciplinary literature review (Webster & Watson, 2002). Articles were synthesized based on thematic analysis of the domains of the

sources, methodology used, key findings, theories, and gaps in knowledge. In the analysis of the domains, a concept-centric methodology was used as opposed to an author-centric methodology. In the analysis of the domains, a concept-centric methodology was used as opposed to an author-centric methodology. Gap analysis was conducted using a systematic comparison of existing knowledge and knowledge required for quality management in a circular economy. This led to the identification of the six gaps as presented in Section 3.5.

Figure 1. PRISMA-Informed Flow of the Systematic Literature Search



Note. Search period: 2000–2025. Database searches conducted across Scopus, Web of Science Core Collection, Google Scholar, and ProQuest Dissertations & Theses Global, supplemented by backward and forward citation tracking.

Figure 1. PRISMA-informed flow of the systematic literature search.
Note. Search period: 2000–2025. Database searches were conducted across Scopus, Web of Science Core Collection, Google Scholar, and ProQuest Dissertations & Theses Global, supplemented by backward and forward citation tracking.

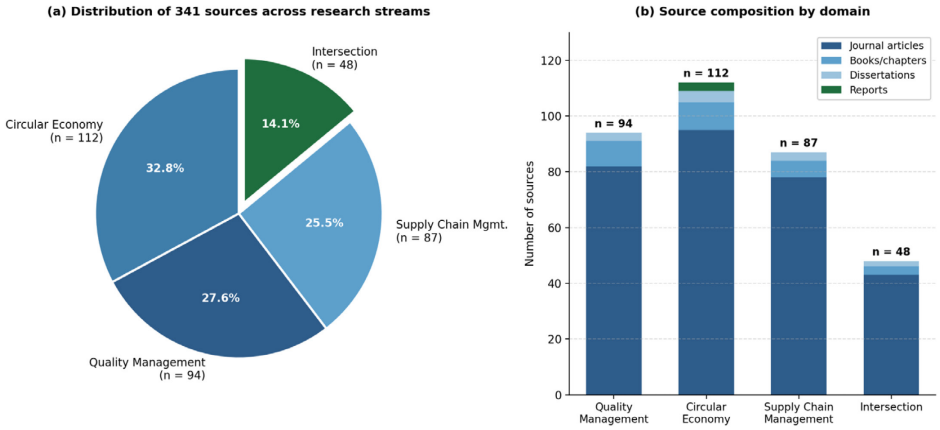


Figure 2. Distribution of sources across research streams.
Note. Panel (a) shows the share of each research stream within the final corpus (N = 341). Panel (b) decomposes each stream into journal articles, books/chapters, doctoral dissertations, and practitioner reports.

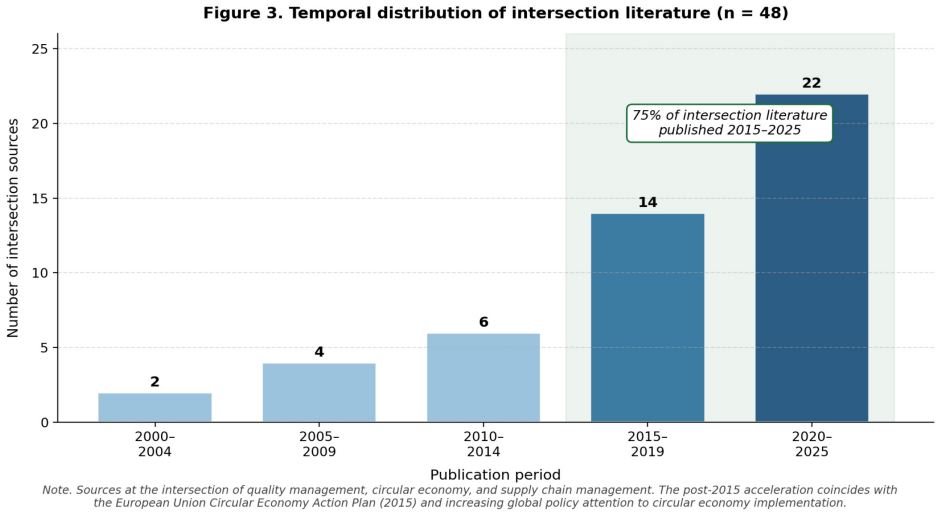


Figure 3. Temporal distribution of intersection literature (n = 48).
Note. The shaded region marks 2015–2025. The post-2015 acceleration coincides with the European Union Circular Economy Action Plan (2015) and increasing global policy attention to circular economy implementation.

3. Literature Review and Theoretical Synthesis

3.1 Quality Management: Foundations and Limitations

Quality management has developed through five different eras: inspection (1920s-1940s), statistical quality control (1930s-1960s), quality assurance (1960s-1980s), total quality management (1970s-1990s), and finally modern quality management systems (1990s-present) (Garvin, 1988; Dale et al., 2016). The basic philosophies of Deming (1994), Juran (1988), Crosby (1979), Ishikawa (1985), and Taguchi (1986) have formed the principles of quality management that are still followed in modern quality management systems. ISO 9001:2015 represents the most modern and widely accepted standard for quality management systems. It comprises seven basic principles of quality management: customer focus, leadership, engagement of people, process approach, improvement, evidence-based decision making, and relationship management (ISO, 2015). ISO 9001 comprises ten clauses and represents a comprehensive framework for organizational context, leadership, planning, support, operation, performance evaluation, and improvement. ISO 9001 is the most widely used standard for quality management systems with more than 1.1 million certifications worldwide and represents the most dominant form of quality management and coordination in supply chains (ISO, 2023).

Supply chain quality management (SCQM) represents the extension of quality management systems beyond organizational boundaries. Research has confirmed that supply chain quality management includes supplier quality management, inter-organizational quality coordination, and multi-tier quality visibility (Foster, 2008; Kaynak & Hartley, 2008). Quality metrics include process-level metrics (defect rates and process capability), product-level metrics (warranty claims and field failures), and supply chain-level metrics (supplier performance and customer satisfaction and cost of quality) (Nair, 2006).

Critical Limitations for Circular Supply Chain Contexts. Despite its inclusiveness, the quality management literature reveals some fundamental limitations when applying its principles and practices to circular supply chain contexts. For instance, while developing the ISO 9001:2015 standard, it was based on some implicit assumptions about unidirectional material flows, the need for virgin material inputs, a single cycle of product usage, and a conventional forward supply chain configuration. These assumptions are incompatible with the principles and practices of a circular economy in a number of critical dimensions. First, while the standard's process approach emphasizes a unidirectional and sequential supply chain configuration, it does not accommodate reverse material flows and associated coordination. Second, while the standard's approach to supplier management emphasizes a constant and controllable level of quality from suppliers, this assumption is violated by the inherent variability of recycled material suppliers. Third, while the standard's performance monitoring and measurement approach emphasizes a focus on forward-flow quality indicators, it does not accommodate multi-cycle quality

tracking and material degradation. Finally, while the standard's focus on a single organizational context does not accommodate multi-organizational coordination and cooperation that are inherent in a circular supply chain context where multiple actors must coordinate their efforts to achieve a desired level of quality.

3.2 Circular Economy: Principles, Models, and Quality Challenges

The notion of the circular economy has crystallized around three key principles defined by the Ellen MacArthur Foundation (2013): "designing out waste and pollution," "keeping resources in use," and "regenerating natural systems." An analysis of 114 definitions of the circular economy was made by Kirchherr et al. (2017), who concluded that the emphasis of circularity was placed on "resource efficiency," "closed-loop," and "systemic value retention." A hierarchy of circular strategies was defined by the 9Rs, from refuse, rethink, reduce, reuse, repair, refurbish, remanufacture, repurpose, recycle, recover, etc., showing decreasing levels of resource efficiency from the center outwards (Potting et al., 2017). An important point to make here is that every R-strategy poses specific quality management issues: the inner loops of the circular economy, i.e., reuse and repair, require evaluation of continued fitness-for-purpose, while the middle loops of the circular economy, i.e., refurbish and remanufacture, require restoration of products to a certain level of quality despite variable input conditions, and finally, the outer loops of the circular economy, i.e., recycle, require managing the level of material degradation despite the level of quality (Kirchherr et al., 2023).

Figure 4. The 9R Framework and Quality Management Implications

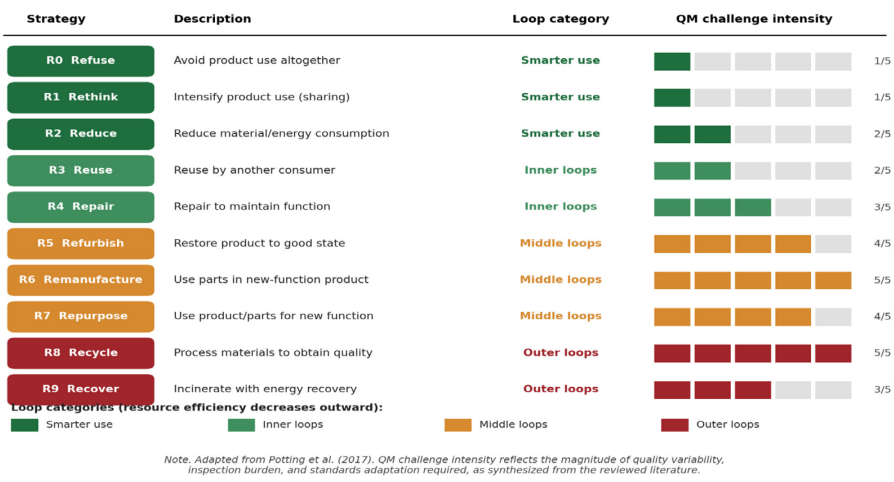


Figure 4. The 9R framework and quality management implications.
Note. Adapted from Potting et al. (2017). Quality management challenge intensity reflects the

magnitude of quality variability, inspection burden, and standards adaptation required, as synthesized from the reviewed literature.

New circular business models have appeared, such as the circular supply chain based on recycled materials, resource recovery, product life extension based on product durability and maintenance, product sharing based on intensive product utilization, and product as a service where manufacturers retain ownership (Bocken et al., 2016). Each of these models presents unique quality management requirements. In the product as a service model, for example, the quality management requirements change significantly since the manufacturers must maintain the quality of the products for multiple use cycles, which is not the case in traditional product sales (Bansal et al., 2024).

Quality has been recognized as a major barrier in the implementation of the circular economy in the literature, but the analysis of the issue has remained superficial (Kirchherr et al., 2023). In the literature, the issue of quality has been recognized as one of the top five barriers to the implementation of the circular economy (Kirchherr et al., 2023). In the systematic analysis of the literature, Govindan and Hasanagic (2018) noted the presence of quality-related challenges in the technical, economic, and market barrier categories. Nevertheless, the literature analysis of the issue of quality in the context of the circular economy has remained superficial, failing to apply quality management theory in the analysis of the issue of quality in the circular economy and suggesting quality management solutions—a problem that the current paper aims to solve.

3.3 Supply Chain Management in Circular Contexts

Supply chain management has evolved from a logistics-driven discipline to a strategic concept that includes internal integration, supplier integration, and customer integration (Mentzer et al., 2001; Flynn et al., 2010). The development of sustainable supply chain management (SSCM) has led to the extension of SCM to the integration of environmental and social aspects in addition to the economic ones (Seuring & Müller, 2008; Carter & Rogers, 2008). Most recently, the development of circular supply chain management (CSCM) as a new paradigm in the field of SCM requires the fundamental transformation of the existing supply chains, as opposed to the incremental improvement of the existing ones (Farooque et al., 2019; Garcia-Buendia et al., 2024).

Circular supply chains add to traditional supply chain complexity in multiple ways. Horizontal complexity is enhanced by the introduction of different types of partners, such as collectors, sorters, and recovery operations. Vertical complexity is heightened by the extension of the supply chain tiers, thereby bridging forward and reverse supply chains. Spatial complexity is also enhanced by the extension of reverse logistics operations beyond geographical areas where forward supply chains operate. Dynamic complexity is heightened by the uncertain volume and time of product returns (Batista et al., 2018; Farooque et

al., 2019). The management of multiple tiers is a challenge, as visibility is lost beyond the first tier of the supply chain, yet visibility is essential in the context of circular supply chains, where operations such as collection, sorting, and recovery may be several tiers away from the firm of primary concern.

Supply chain governance and collaboration mechanisms, such as trust development, information sharing, collaborative planning, and contracting, are vital for the performance of the circular supply chain, yet they face unique challenges. According to transaction cost economics, the high asset specificity, uncertainty, and transaction frequency of exchange in the context of the circular supply chain make a hybrid governance structure more appropriate, consisting of both contractual and relational elements (Williamson, 1985). However, research on the effectiveness of such hybrid structures designed for quality coordination in the context of the circular supply chain is severely limited (Batista et al., 2018; Garcia-Buendia et al., 2024).

3.4 The Quality-Circular Economy-Supply Chain Nexus: A Critical Synthesis

The systematic intersection search resulted in 48 sources that explicitly deal with the theme of quality management within the context of supply chains for a circular economy. Moreover, 75% of the literature has been published between 2015 and 2025. This underpins the emerging nature of this research area. The intersection literature can be sub-divided into five sub-areas: remanufacturing quality management with 16 sources, recycled material quality management with 12 sources, reverse logistics quality management with 10 sources, quality perceptions with 8 sources, and finally standards and certification with 2 sources.

Remanufacturing quality issues. The remanufacturing literature extensively discusses the quality uncertainty associated with the variability of core condition. The products received for remanufacturing are in a variety of conditions, from new to severely used, causing fundamental uncertainty in remanufacturing planning, processing, and pricing (Guide & Van Wassenhove, 2009; Bansal et al., 2024). The mathematical models used to optimize inspection and grading methods (Ferguson et al., 2009), and the more recent developments in quality control models for remanufacturing processes such as disassembly, cleaning, inspection, repair, and reassembly (Liu et al., 2023, 2025), all show that, while technical quality parity between remanufactured and new products can be achieved in most product categories, there are quality discounts averaging 20-30% due to customer perceptions, even where objective quality measures indicate equivalence (Abbey et al., 2017; Hazen et al., 2017). However, the above-mentioned research focuses primarily on single-firm remanufacturing operations and does not extend to the supply chain.

Recycled material quality variability. The material science literature extensively discusses the degradation mechanisms and the quality variability of materials, including plastics, metals, and paper, where plastics are subject to

chain scission, oxidation, and contamination, restricting recycling to 5-7 cycles, with a 20-50% property variation from the same recycler (Ragaert et al., 2017; Jiang & Bateer, 2025), metals accumulating tramp elements from recycling requiring periodic dilution with virgin materials to maintain alloy specifications (Allwood et al., 2011), and paper fibers shortening from recycling requiring cascading to progressively lower value uses (Villanueva & Wenzel, 2007), all of which cause manufacturing issues including increased defect rates 2-4 times higher than virgin material processing, reduced process capability, and the risk of equipment damage (Vilaplana & Karlsson, 2008), and most importantly, the conventional approach to supplier qualification assumes constant and controllable quality, which is fundamentally violated by the recycled material suppliers who are subject to the inherent feedstock variability beyond their control.

Reverse logistics quality management. Gatekeeping reflects the essential quality management activity in reverse logistics, deciding upon the quality of products that are to be accepted into the system of recovery, as opposed to rejection (Ferguson et al., 2009). Research has established that inspection strategies are influenced by inspection costs, processing costs, value distribution of recoveries, and uncertainty characteristics. However, inspection procedures, grading schemes, and disposition systems are still in their infancy stages of development. In addition, information asymmetry poses a fundamental problem because information on usage of products prior to collection is generally unavailable, and consumers may misrepresent the condition of products, making inspection necessary to overcome the information gap (Vlachos et al., 2007). Collection operation quality, i.e., the means of collection, handling, and transportation of products, plays a significant role in determining the quality of reverse logistics but remains a completely unexplored area of research.

Quality standards and certification gaps. Inconsistencies of ISO 9001 in circular system environments are generally accepted, but adaptation recommendations are still unavailable. Industry-specific remanufacturing standards are available in automotive and aerospace industries, but overall industry-wide quality standards for circular products are still in their development stages (Parker et al., 2015; Matsumoto & Komatsu, 2015). In addition, certification schemes that provide assurance of quality of circular products are still unavailable. Cradle to Cradle provides quality assurance at the design stage but does not extend to the performance stage of the product life cycle. Environmental certifications are also unavailable because they fail to measure quality dimensions.

Quality metrics limitations. Although traditional quality metrics are deemed necessary but insufficient, they prove to be so in a circular supply chain setting. As traditional quality metrics focus only on forward flow, a single cycle, and virgin material, they do not address multi-cycle quality, material degradation, recovery operations, and integrated quality/circularity. Although some circular quality metrics, like remanufacturing yields, quality parity indices, and material

circularity indices, have been proposed, they remain conceptual and need to be validated, standardized, and implemented (Linder et al., 2017; Matsumoto & Komatsu, 2015). Quality metrics for a supply chain, which evaluate coordination effectiveness among circular supply chain partners, remain particularly missing (Batista et al., 2018).

Table 1 Theme to Evidence Matrix: Intersection Literature Summary

Theme	Key Evidence	Status	Gap Severity
Remanufacturing quality uncertainty	Variable core conditions extensively documented; mathematical models for inspection/grading	Partially addressed at firm level	High for SC-level
Recycled material variability	Degradation mechanisms characterized; 20–50% property variation documented	Material science advanced; QM implications absent	Very high
Quality perceptions barrier	20–30% quality discount; terminology and warranty effects documented	Well documented	Moderate
Reverse logistics gatekeeping	Optimal inspection models developed; practical procedures absent	Theoretical models only	High
Supply chain quality coordination	Identified as critical need; mechanisms not developed	Conceptual recognition only	Very high
Quality standards/certification	ISO 9001 limitations acknowledged; adaptation absent	Gap recognized, not addressed	Very high
Quality metrics for CE	Proposed metrics conceptual; no validation or standardization	Nascent	High

3.5 Articulation of Research Gaps

Comprehensive synthesis indicates that the primary gap is the lack of frameworks for integrating the principles of CE into quality management systems, particularly in supply chains, where CE is increasingly being adopted worldwide, and quality issues have already emerged as a barrier to the implementation of CE, as noted by several studies. The primary gap is comprised of six interconnected gaps, which are discussed below.

Gap 1: Lack of comprehensive quality management systems. No quality management systems for supply chains adopting CE have been developed,

similar to the comprehensiveness of the ISO 9001 standard in linear supply chains (Matsumoto & Komatsu, 2015; Govindan & Hasanagic, 2018; Nguyen et al., 2023). Although some quality management practices, such as gatekeeping inspection, remanufacturing testing, and recycled material qualification, have been developed, such research has not developed comprehensive quality management systems. As such, organizations adopting CE have to develop quality management systems on an ad hoc basis, as no standard guidelines exist.

Gap 2: Lack of quality management systems in supply chains. The quality management literature has mainly focused on single firms, ignoring the supply chain coordination aspect of quality management (Guide & Van Wassenhove, 2009; Matsumoto & Komatsu, 2015). However, CE implementation is dependent on quality coordination among multiple firms, including those in collection, recycling, and manufacturing activities, as noted by several authors, such as Batista et al. (2018) and Farooque et al. (2019). The quality management systems in supply chains, including quality coordination mechanisms, quality governance, quality information systems, and quality collaboration, have received little attention in the literature.

Gap 3: Lack of quality metrics. Although traditional quality metrics are necessary but insufficient, comprehensive quality metric systems for CE have yet to be developed, as noted by Linder et al. (2017) and Govindan & Hasanagic, 2018). Supply chain quality metrics, such as evaluating quality coordination effectiveness, multi-cycle performance, and quality-circularity outcomes, are totally missing. The lack of quality metrics means that quality management systems lack performance feedback, which is a basic requirement of quality management.

Gap 4: Lack of implementation guidance. The research, in most cases, is highly theoretical in nature, using mathematical modeling, concepts, and descriptive studies (Souza, 2013; Guide & Van Wassenhove, 2009). Procedural guidance, inspection methods, testing methods, qualification of suppliers, and coordination of the supply chains are lacking or scarce. The “how to” questions are not adequately answered by the research.

Gap 5: Limited Empirical Validation. The proposed frameworks, metrics, and approaches are not validated to prove the effectiveness of the approach (Matsumoto & Komatsu, 2015; Liu et al., 2023).

Gap 6: Underdeveloped Standards and Certifications. Quality standards and certifications for circular products are still in the early stages of development (Parker et al., 2015; Matsumoto & Komatsu, 2015). ISO 9001 adaptation guidelines are not available. Quality standards and certification schemes that offer quality assurance for circular products across different industries are still nonexistent.

3.6 Theoretical Foundations

Due to the multi-faceted nature of the concept of quality management in circular supply chains, the study is underpinned by a multi-theory approach. Five different theoretical foundations are incorporated into the conceptual framework, as presented in Section 4.

Systems theory considers circular supply chains as complex adaptive systems, where the emergent properties of the supply chain result from the interaction between the different components, rather than from the actions of the individual organizations (Bertalanffy, 1968). Deming's (1994) System of Profound Knowledge is directly applicable to the context of circular supply chains, where the outcomes of the supply chain are a result of the design of the circular supply chain, including the feedback loops between the different supply chain components, the interaction between the different supply chain actors, and the emergent properties of the supply chain.

The resource-based view (RBV) of the firm provides a theory of quality management capabilities, which act as a source of competitive advantage (Barney, 1991; Teece et al., 1997). Circular quality management capabilities, including variable material processing skills, multi-cycle quality tracking, recovery operation quality, and supply chain quality coordination, qualify for being valuable, rare, inimitable, and non-substitutable given their novelty. Capability heterogeneity among firms is a source of cross-firm variation in implementing circular economy and a basis for informing a framework that emphasizes development of capabilities in a staged fashion.

Stakeholder theory is used to address the changing nature of stakeholder relationships in circular economy supply chains, where relationships with a wider circle of stakeholders are now a key feature (Freeman, 1984). Quality in circular economy supply chains is achieved by coordinating diverse stakeholders, including collectors, sorters, recovery operations, secondary markets, regulators, and environmental groups, with varied interests, powers, and legitimacy. Stakeholder salience analysis indicates that customers have high power and legitimacy in determining product quality, while recovery operations have moderate legitimacy but high operational power. Cooperative governance is necessary in circular economy supply chains, given their voluntary nature.

Institutional theory is used to explain the influence of regulatory, normative, and cultural-cognitive pressures in shaping quality management in circular economy supply chains (DiMaggio & Powell, 1983; Scott, 2008). As the circular economy develops, isomorphic pressures will build, including intensification of regulatory, normative, and cultural-cognitive pressures. As a result, standards for circular economy will emerge, including those for circular economy product quality management. Institution entrepreneurship in developing standards for circular economy product quality management contributes to developing a level of institutional infrastructure in a field. Legitimacy in a market is provided by

certification, a necessary condition for market acceptance of circular economy products.

Transaction Cost Economics (TCE) provides a framework for understanding governance structures in quality coordination transactions in uncertain conditions (Williamson, 1985). Quality transactions in circular supply chains have moderate-to-high asset specificity (testing equipment and knowledge), high uncertainty (variable input qualities and markets), and vary in frequency; hence, these are likely to have hybrid governance structures. Difficulty in measuring qualities in circular transactions creates opportunities for opportunism; hence, governance structures are needed.

These five theoretical perspectives, as a whole, provide an explanation for the nature of quality management challenges in circular supply chains, the requirements for quality management, and how it can be achieved. They form the basis for the conceptual framework that follows.

4. Conceptual Framework

4.1 Framework Overview

This section, based on the synthesis of the literature review and the multi-theoretical foundation, presents a conceptual framework that conceptualizes the quality management of circular supply chains as a multi-level system, including the organizational, supply chain, and institutional levels. Quality management outcomes are theorized to be the result of the interaction between these levels, mediated by contextual characteristics, and enabled by organizational capabilities. The conceptual framework incorporates the foundations of quality management, the requirements of the circular economy, and the mechanisms of supply chain coordination, while considering the institutional context of implementation. It is intended to organize the insights from the literature review into a conceptual framework, to identify the theorized relationships between the constructs, to guide the design of the research, and to communicate the research logic that connects the identified gaps to the research approaches (Maxwell, 2013; Miles & Huberman, 1994). Figure 5 presents the framework graphically; Sections 4.2 and 4.3 elaborate its constructs and theorized relationships.

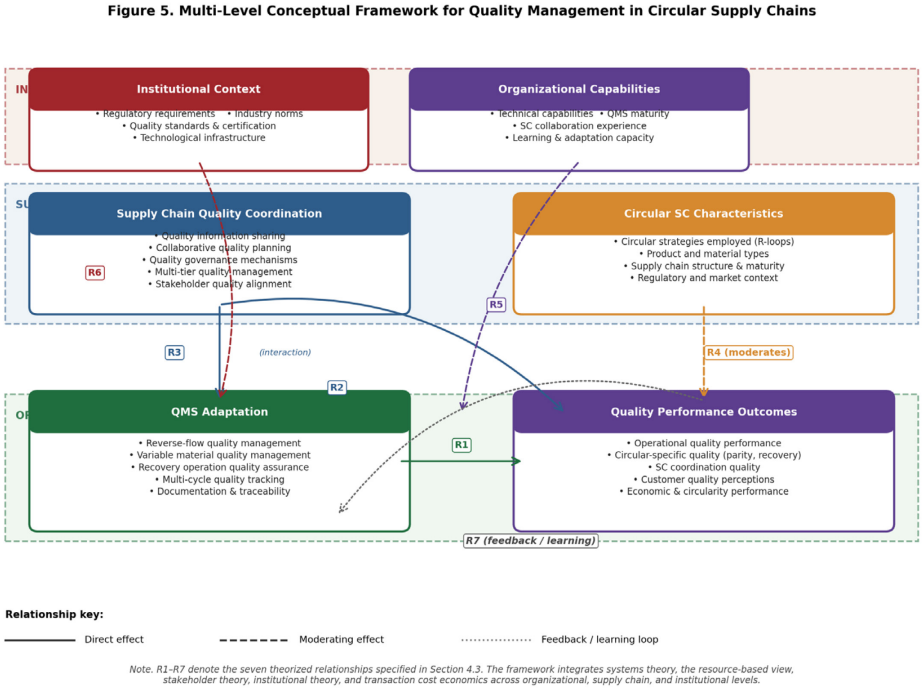


Figure 5. Multi-level conceptual framework for quality management in circular supply chains.

Note. R1–R7 denote the seven theorized relationships specified in Section 4.3. The framework integrates systems theory, the resource-based view, stakeholder theory, institutional theory, and transaction cost economics across organizational, supply chain, and institutional levels.

4.2 Key Constructs

Quality management system adaptation (organizational level): This construct refers to the organizational level of quality management systems that are adapted to the requirements of the circular economy (ISO, 2015; Govindan & Hasanagic, 2018). Quality management system adaptation is composed of five dimensions, namely, reverse flow quality management, variable material quality management, recovery operation quality assurance, multi-cycle quality tracking, and documentation and traceability.

Supply chain quality coordination (supply chain level). This construct refers to inter-organizational quality management instruments among partners in a circular supply chain (Batista et al., 2018; Flynn et al., 2010). Its dimensions include: (a) sharing of quality-related information—quality data sharing instruments among partners; (b) collaborative quality planning—approaches to quality standards, specifications, and improvement among partners; (c) quality

governance instruments—quality-related contracts, standards, incentives, and relationship management among partners; (d) multi-tier quality management approaches—quality management instruments for multi-tier supply chains; and (e) stakeholder quality alignment processes—approaches to aligning diverse stakeholders' quality expectations and contributions.

Circular supply chain characteristics (contextual moderators). These factors influence the relationship between quality management approaches and outcomes. They include: circular strategies used (which R-strategies are implemented); product and material types (complexity, degradation); supply chain structure (number of partners, tiers, geographic dispersion); circular maturity level; and contextual factors (regulatory, market).

Quality performance outcomes (dependent variables). Quality performance outcomes in a circular supply chain can be multi-dimensional, including operational quality performance, circular supply chain quality performance, supply chain coordination quality, customer quality perceptions, economic performance, and circular supply chain performance (Linder et al., 2017).

Institutional context and organizational capabilities. The institutional context, including regulatory requirements, quality standards, certification schemes, industry norms, and technological infrastructure, is a source of both opportunities and constraints for supply chain firms (Scott, 2008; DiMaggio & Powell, 1983). Organizational capabilities, including technical, quality management system, supply chain collaboration, learning, and adaptation, and resource capabilities, moderate the effectiveness of supply chain implementation (Barney, 1991; Teece et al., 1997).

4.3 Theorized Relationships

The framework proposes seven important relationships. First, quality management system adaptation has a positive impact on quality performance outcomes (R1). Second, supply chain quality coordination has a positive impact on quality performance outcomes (R2). Third, QMS adaptation and supply chain coordination have a synergistic interaction. Organizations that excel in both areas perform better than those that excel in only one area (R3). Fourth, circular supply chain characteristics mediate the relationships between quality management approaches and outcomes. In more complex contexts, more sophisticated quality management is required (R4). Fifth, organizational capabilities mediate the effectiveness of implementation (R5). Sixth, institutional context facilitates or constrains implementation (R6). Lastly, quality performance outcomes have a feedback effect that influences future quality management practices through a cycle of continuous improvement and learning that leads to improved capabilities (R7).

Table 2 Conceptual Framework: Key Constructs, Levels, and Theoretical Grounding

Construct	Level	Theoretical Basis	Role in Framework
QMS Adaptation	Organizational	Systems theory; QM theory (Deming, 1994)	Independent variable
SC Quality Coordination	Supply chain	Systems theory; Stakeholder theory	Independent variable
Circular SC Characteristics	Contextual	Contingency theory; Systems theory	Moderator
Quality Performance Outcomes	Multi-level	QM theory; RBV	Dependent variable
Institutional Context	Institutional	Institutional theory (DiMaggio & Powell, 1983)	Facilitator/Constraint
Organizational Capabilities	Organizational	RBV (Barney, 1991); Dynamic capabilities	Moderator
Feedback Loops	Systemic	Systems theory; Learning theory	Dynamic mechanism

4.4 Framework Application to Research Design

The conceptual framework is the foundation for a research design consisting of four phases. In Phase 1, Problem Investigation, quality management challenges are examined at organizational and supply chain levels, including capability gaps and institutional constraints. In Phase 2, Framework Development, quality management system adaptation approaches and supply chain coordination mechanisms are developed, based on challenges identification and theoretical foundations. In Phase 3, Pilot Implementation, frameworks are tested through implementation, measuring quality performance outcomes across all dimensions, including feasibility, organizational capabilities, and institutional context. In Phase 4, Synthesis, transferability is examined, frameworks are refined, and standards recommendations are developed, contributing to institutional development.

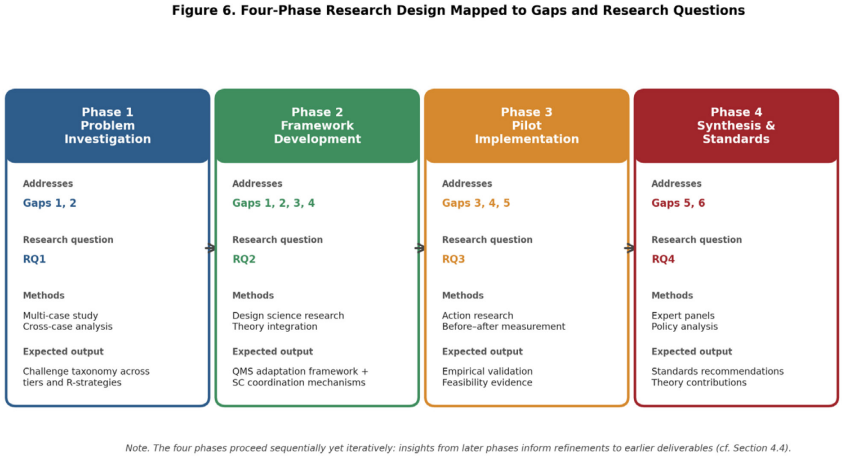


Figure 6. Four-phase research design mapped to gaps and research questions.

Note. The four phases proceed sequentially yet iteratively: insights from later phases inform refinements to earlier deliverables.

5. Discussion

5.1 Interpretation of Findings

This systematic review indicates that three established research domains of quality management, circular economy, and supply chain management have, to a great extent, developed independently of each other, even at the critical point of intersection. Quality management theory has developed highly sophisticated models to support the linear economy, while the circular economy's demands have not been matched by the same level of sophistication. Circular economy research has emerged dramatically, especially after the Ellen MacArthur Foundation's seminal report of 2013, yet continues to identify quality as the number one barrier to implementation, without reference to the established theory of quality management. Supply chain management research is increasingly addressing sustainability, yet largely from the perspective of linear supply chain flow and environmental impact, rather than the circular supply chain.

The intersection research, consisting of just 48 of the total of 341 sources, and concentrated mainly in the last ten years, indicates that this is a growing field, albeit a disjointed one. Knowledge is problem-focused, identifying the issue of quality, yet does not offer a framework for solution-finding. This is, of course, the normal course of events during the early stages of the development

of a new field of management research, where the identification of problems is necessarily the precursor to the development of solution frameworks.

5.2 Theoretical Implications

This systematic review extends the theory of quality management by identifying that the basic assumptions underpinning the ISO 9001 and SCQM models, such as unidirectional supply chain flow, the use of virgin materials, supply chain use, and linear supply chain flow, are violated by the demands of the circular economy. This is not an extension of the theory, as such, but rather a direct challenge to the assumptions that underpin the whole of the quality management theory. The multi-theory framework that this systematic review presents indicates that the challenges to the theory of quality management, as presented, are to be addressed by the integration of systems theory, resource-based view, stakeholder theory, institutional theory, and transaction cost economics, a synthesis that has not been achieved before.

With regard to circular economy theory, the review shows that quality management is not simply one of several hurdles that circular economy strategies face, but rather a cross-cutting challenge that affects all other dimensions of circular economy strategies at the same time. This implies that the importance of quality management integration in circular economy strategies should be elevated from a secondary to a primary consideration (Kirchherr et al., 2023; Govindan & Hasanagic, 2018).

With regard to supply chain management theory, the review shows that the quality coordination challenges of circular supply chains are qualitatively different from those of linear supply chains. This implies that new coordination tools are necessary to address the quality coordination challenges of circular supply chains, which are not covered by existing SCQM theories (Batista et al., 2018; Garcia-Buendia et al., 2024).

5.3 Practical Implications

The review has several implications for the practice of circular economy strategies. First, it implies that organizations adopting circular economy strategies should recognize that their existing ISO 9001-based quality management systems are not simply adequate and require extension, but rather fundamental adaptation to the requirements of the circular economy. Second, it implies that organizations adopting circular economy strategies should anticipate quality coordination challenges in all R-strategies and proactively address them, rather than waiting until quality problems emerge. Finally, it implies that supply chain-level quality coordination should be given at least as much attention as internal quality management because the success of circular economy strategies depends on the quality alignment of multiple organizations, and that organizations adopting circular economy strategies should invest in capability

development in circular quality management because existing quality management capability may be necessary but insufficient.

In the context of standards development organizations, the review demonstrates the value of existing quality standards as a foundation that needs extension to meet the requirements of the circular economy. In particular, the development of standards for circular product quality, recycled material specifications, and remanufactured product certification represent essential infrastructure requirements. The window of opportunity, as the circular economy develops but before path dependence in existing standards becomes entrenched, presents an important opportunity for proactive standards development based on systematic research (Blind & Mangelsdorf, 2016).

5.4 Limitations

Several limitations of the review should be noted. First, the review of English-language literature may have omitted important research published in other languages, especially given the research activity in the context of the circular economy in China, the European Union, and emerging economies. Second, the decision not to review conference literature may have omitted important research that has not yet been published in journal form. Third, the decision to review literature from 2000 to 2025 may not have captured the essential literature published in earlier timeframes, although seminal works from earlier timeframes were included in the review. Fourth, as a literature-based review, the limitations and gaps require further research to validate the significance of the findings. Fifth, the dynamic nature of the circular economy literature means that important literature may have been published within the timeframe of the review (2000-2025) that could influence some of the findings, although the structural limitations and gaps in the literature are well established.

6. Future Research Agenda

Four research questions are formulated based on the six gaps identified, along with a conceptual framework, to provide a research direction for future investigation.

RQ1: What are the quality management challenges that emerge from the incorporation of circular economy concepts into supply chains, and how do they vary between supply chain levels and circular strategies? This research question seeks to address the problem investigation gap by identifying quality management challenges in a systematic manner, including their comparison between supply chain levels and circular strategies, which was not considered in previous literature (Batista et al., 2018; Govindan & Hasanagic, 2018).
Recommendations: Multi-case comparative research methodology.

RQ2: In what manner can quality management systems be adapted and integrated to address the requirements of circular economy concepts in supply

chains at various levels, including the frameworks that support effective implementation? This research question seeks to address the framework development gap by creating a comprehensive quality management framework that incorporates the requirements of circular economy concepts in supply chains, which was the most significant gap identified.
Recommendations: Design science research methodology.

RQ3: What are the effectiveness, feasibility, and transferability of the developed quality management frameworks in real-world supply chain operation? This research question seeks to address the gap of empirical validation of quality management frameworks in real-world supply chains.
Recommendations: Action research methodology.

RQ4: What are the implications of the research on quality management theory, circular economy concepts, supply chain management, and standards development? This research question seeks to address the gap of cross-domain implications of the research.

Recommendations:

Table 3 Future Research Agenda Mapped to Identified Gaps

Gap Addressed	Research Question	Recommended Method	Expected Contribution
Gap 1 (QMS frameworks)	RQ2: Framework development	Design science research	Comprehensive QMS framework for circular SC
Gap 2 (SC-level QM)	RQ1, RQ2: Challenge ID + coordination	Multi-case study; Action research	SC-level coordination mechanisms
Gap 3 (Quality metrics)	RQ2, RQ3: Development + validation	Mixed methods; Pilot testing	Validated circular quality metrics
Gap 4 (Implementation)	RQ2, RQ3: Frameworks + pilots	Design science; Action research	Practical tools, templates, procedures
Gap 5 (Empirical validation)	RQ3: Effectiveness assessment	Before–after implementation study	Evidence of framework effectiveness
Gap 6 (Standards)	RQ4: Standards recommendations	Expert panels; Policy analysis	Evidence-based standards guidance

In addition to the four research questions outlined above, several other research directions can be proposed as a priority. First, research into specific industries and sectors such as the electronics industry, automobile industry, textile industry, and construction industry can help understand the varying nature of

circular quality management needs depending on the complexity and nature of products. Second, research into the role and implications of digital technologies that can support and enable the development and implementation of circular quality management can also be an important research direction. For example, research into digital product passports, blockchain-based material tracing, and the role of artificial intelligence in quality prediction can be important. Third, research into the dynamics and development of circular quality management capabilities can also be important. This can help understand organizational learning dynamics. Fourth, research into international comparisons and how institutional differences affect the development and implementation of circular quality management can also be important. This can help expand the scope and applications of institutional theory.

7. Conclusion

This systematic literature review has provided a comprehensive analysis of the concepts of quality management, circular economy, and supply chain management as separate but not sufficiently integrated research areas, covering a total of 341 literature sources to identify the key gaps in the integration of these three areas of research. It has been established that, while specific quality-related issues in the context of the circular economy have been sufficiently covered, such as quality uncertainty in remanufacturing, material quality in recycling, quality perception barriers, and inspection in reverse logistics, the integration of quality management and circular economy in the context of supply chains lacks sufficient framework.

Six key research gaps in the integration of quality management and the circular economy in the context of supply chains have been established: the lack of comprehensive QMS for circular supply chains, the lack of quality management coordination in the context of the entire supply chain, the lack of quality metrics, the lack of implementation guidance, the lack of sufficient empirical validation, and the lack of standards and certification. These research gaps have key implications for the development of the circular economy, despite the growing pressure of legislation and the market.

A conceptual framework for the integration of the key concepts of the research has been developed based on the principles of system theory, the resource-based view of the firm, stakeholder theory, institutional theory, and transaction cost economics.

The relevance of this study transcends the contribution to knowledge in the academic sphere. The topic of quality management poses a cross-cutting challenge that impacts all the circular economy barriers, including technical, economic, and market barriers, as well as supply chain barriers. Thus, tackling the topic of quality management can lead to the advancement of all the circular economy barriers at the same time. Considering the increased pace at which the

circular economy phenomenon is spreading all over the world, the disruptions in the supply chains that are prompting the need to improve resilience, and the window of time within which standards can be developed as the circular economy evolves, the time for this study is now.

Future studies on the topic should focus on the following areas: (a) the study of the quality management issues in the circular supply chains, (b) the development of a comprehensive and validated quality management framework, (c) the implementation and assessment of the effectiveness of the quality management approach, and (d) the formulation of recommendations for standards development. The area of research on the topic of quality management in the circular economy has the promise of not only advancing the knowledge in the field of management but also providing practitioners the tools and the confidence to adopt the circular economy approach.

References

[1] Abbey, J. D., Kleber, R., Souza, G. C., & Voigt, G. (2017). The role of perceived quality risk in pricing remanufactured products. Production and Operations Management, 26(1), 100–115. https://doi.org/10.1111/poms.12628

[2] Allwood, J. M., Ashby, M. F., Gutowski, T. G., & Worrell, E. (2011). Material efficiency: A white paper. Resources, Conservation and Recycling, 55(3), 362–381. https://doi.org/10.1016/j.resconrec.2010.11.002

[3] Bansal, S., Guide, V. D. R., Jr., & Souza, G. C. (2024). Closed-loop supply chains with product remanufacturing: Challenges and opportunities. Journal of Operations Management, 70(2), 150–189. https://doi.org/10.1002/joom.1298

[4] Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

[5] Batista, L., Bourlakis, M., Smart, P., & Maull, R. (2018). In search of a circular supply chain archetype—A content-analysis-based literature review. Production Planning & Control, 29(6), 438–451. https://doi.org/10.1080/09537287.2017.1343502

[6] Bertalanffy, L. von. (1968). General system theory: Foundations, development, applications. George Braziller.

[7] Blind, K., & Mangelsdorf, A. (2016). Motives to standardize: Empirical evidence from Germany. Research Policy, 45(6), 1285–1296. https://doi.org/10.1016/j.respol.2016.03.007

[8] Bocken, N. M. P., de Pauw, I., Bakker, C., & van der Grinten, B. (2016). Product design and business model strategies for a circular economy. Journal of Industrial and Production Engineering, 33(5), 308–320. https://doi.org/10.1080/21681015.2016.1172124

[9] Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. Educational Researcher, 34(6), 3–15. https://doi.org/10.3102/0013189X034006003

[10] Cao, M., & Zhang, Q. (2011). Supply chain collaboration: Impact on collaborative advantage and firm performance. Journal of Operations Management, 29(3), 163–180. https://doi.org/10.1016/j.jom.2010.12.008

[11] Carter, C. R., & Rogers, D. S. (2008). A framework of sustainable supply chain management: Moving toward new theory. International Journal of Physical Distribution & Logistics Management, 38(5), 360–387. https://doi.org/10.1108/09600030810882816

[12] Crosby, P. B. (1979). Quality is free: The art of making quality certain. McGraw-Hill.

[13] Dale, B. G., van der Wiele, T., & van Iwaarden, J. (2016). Managing quality (6th ed.). Wiley-Blackwell.

[14] Deming, W. E. (1994). The new economics for industry, government, education (2nd ed.). MIT Press.

[15] Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. A. Buchanan & A. Bryman (Eds.), The Sage handbook of organizational research methods (pp. 671–689). Sage.

[16] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160. https://doi.org/10.2307/2095101

[17] Ellen MacArthur Foundation. (2013). Towards the circular economy: Economic and business rationale for an accelerated transition. Ellen MacArthur Foundation.

[18] European Commission. (2020). A new circular economy action plan: For a cleaner and more competitive Europe. European Commission. https://ec.europa.eu/environment/circular-economy/

[19] Farooque, M., Zhang, A., Thürer, M., Qu, T., & Huisingh, D. (2019). Circular supply chain management: A definition and structured literature review. Journal of Cleaner Production, 228, 882–900. https://doi.org/10.1016/j.jclepro.2019.04.303

[20] Ferguson, M. E., Guide, V. D. R., Jr., Koca, E., & Souza, G. C. (2009). The value of quality grading in remanufacturing. Production and Operations Management, 18(3), 300–314. https://doi.org/10.1111/j.1937-5956.2009.01033.x

[21] Flynn, B. B., Huo, B., & Zhao, X. (2010). The impact of supply chain integration on performance: A contingency and configuration approach. Journal of Operations Management, 28(1), 58–71. https://doi.org/10.1016/j.jom.2009.06.001

[22] Foster, S. T. (2008). Towards an understanding of supply chain quality management. Journal of Operations Management, 26(4), 461–467. https://doi.org/10.1016/j.jom.2007.06.003

[23] Freeman, R. E. (1984). Strategic management: A stakeholder approach. Pitman.

[24] Garcia-Buendia, N., Moyano-Fuentes, J., Maqueira-Marín, J. M., & Cobo, M. J. (2024). Squaring circular supply chain management: A comprehensive overview of emerging themes and trends. Business Strategy and the Environment, 33(8), 7890–7912. https://doi.org/10.1002/bse.3932

[25] Garvin, D. A. (1988). Managing quality: The strategic and competitive edge. Free Press.

[26] Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The circular economy—A new sustainability paradigm? Journal of Cleaner Production, 143, 757–768. https://doi.org/10.1016/j.jclepro.2016.12.048

[27] Govindan, K., & Hasanagic, M. (2018). A systematic review on drivers, barriers, and practices towards circular economy: A supply chain perspective. International Journal of Production Research, 56(1–2), 278–311. https://doi.org/10.1080/00207543.2017.1402141

[28] Guide, V. D. R., Jr., & Van Wassenhove, L. N. (2009). The evolution of closed-loop supply chain research. Operations Research, 57(1), 10–18. https://doi.org/10.1287/opre.1080.0628

[29] Hart, C. (1998). Doing a literature review: Releasing the social science research imagination. Sage.

[30] Hazen, B. T., Mollenkopf, D. A., & Wang, Y. (2017). Remanufacturing for the circular economy: An examination of consumer switching behavior. Business Strategy and the Environment, 26(4), 451–464. https://doi.org/10.1002/bse.1929

[31] Ishikawa, K. (1985). What is total quality control? The Japanese way. Prentice-Hall.

[32] ISO. (2015). ISO 9001:2015—Quality management systems: Requirements. International Organization for Standardization.

[33] ISO. (2023). The ISO survey of management system standard certifications 2022. International Organization for Standardization.

[34] Jiang, X., & Bateer, B. (2025). A systematic review of plastic recycling: Technology, environmental impact and economic evaluation. Waste Management & Research, 43(4), 325–342. https://doi.org/10.1177/0734242X241310658

[35] Juran, J. M. (1988). Juran on planning for quality. Free Press.

[36] Kaynak, H., & Hartley, J. L. (2008). A replication and extension of quality management into the supply chain. Journal of Operations Management, 26(4), 468–489. https://doi.org/10.1016/j.jom.2007.06.002

[37] Ketokivi, M., & Choi, T. (2014). Renaissance of case research as a scientific method. Journal of Operations Management, 32(5), 232–240. https://doi.org/10.1016/j.jom.2014.03.004

[38] Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. Resources, Conservation and Recycling, 127, 221–232. https://doi.org/10.1016/j.resconrec.2017.09.005

[39] Kirchherr, J., Yang, N.-H. N., Schulze-Spüntrup, F., Heerink, M. J., & Hartley, K. (2023). Conceptualizing the circular economy (revisited): An analysis of 221 definitions. Resources, Conservation and Recycling, 194, 107001. https://doi.org/10.1016/j.resconrec.2023.107001

[40] Linder, M., Sarasini, S., & van Loon, P. (2017). A metric for quantifying product-level circularity. Journal of Industrial Ecology, 21(3), 545–558. https://doi.org/10.1111/jiec.12552

[41] Liu, C., Chen, J., & Wang, X. (2023). Review and challenges for the remanufacturing assembly quality with uncertainty. Green Manufacturing Open, 1(4), 1–15. https://doi.org/10.20517/gmo.2023.072701

[42] Liu, C., Zhang, C., & Cai, W. (2025). Quality control in remanufacturing process: Theoretical framework, control strategy and application review. International Journal of Advanced Manufacturing Technology, 136, 1–25. https://doi.org/10.1007/s00170-025-16343-8

[43] Matsumoto, M., & Komatsu, S. (2015). Demand forecasting for production planning in remanufacturing. International Journal of Advanced Manufacturing Technology, 79(1–4), 161–175. https://doi.org/10.1007/s00170-015-6787-x

[44] Maxwell, J. A. (2013). Qualitative research design: An interactive approach (3rd ed.). Sage.

[45] Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. Journal of Business Logistics, 22(2), 1–25. https://doi.org/10.1002/j.2158-1592.2001.tb00001.x

[46] Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). Sage.

[47] Nair, A. (2006). Meta-analysis of the relationship between quality management practices and firm performance—Implications for quality management theory development. Journal of Operations Management, 24(6), 948–975. https://doi.org/10.1016/j.jom.2005.11.005

[48] Nguyen, K., Akbari, M., Quang, H. T., McDonald, S., Hoang, T.-H., Yap, T. L., & George, M. (2023). Navigating environmental challenges through supply chain quality management 4.0 in circular economy: A comprehensive review. Sustainability, 15(24), 16720. https://doi.org/10.3390/su152416720

[49] Parker, D., Riley, K., Robinson, S., Symington, H., Tewson, J., Jansson, K., Ramkumar, S., & Peck, D. (2015). Remanufacturing market study. European Remanufacturing Network.

[50] Potting, J., Hekkert, M., Worrell, E., & Hanemaaijer, A. (2017). Circular economy: Measuring innovation in the product chain (PBL Policy Brief No. 2544). PBL Netherlands Environmental Assessment Agency.

[51] Psomas, E., Kafetzopoulos, D., & Gotzamani, K. (2018). Determinants of company innovation and market performance. TQM Journal, 30(1), 54–73. https://doi.org/10.1108/TQM-07-2017-0080

[52] Ragaert, K., Delva, L., & Van Geem, K. (2017). Mechanical and chemical recycling of solid plastic waste. Waste Management, 69, 24–58. https://doi.org/10.1016/j.wasman.2017.07.044

[53] Randolph, J. (2009). A guide to writing the dissertation literature review. Practical Assessment, Research, and Evaluation, 14, Article 13. https://doi.org/10.7275/b0az-8t74

[54] Scott, W. R. (2008). Institutions and organizations: Ideas and interests (3rd ed.). Sage.

[55] Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. Journal of Cleaner Production, 16(15), 1699–1710. https://doi.org/10.1016/j.jclepro.2008.04.020

[56] Souza, G. C. (2013). Closed-loop supply chains: A critical review, and future research. Decision Sciences, 44(1), 7–38. https://doi.org/10.1111/j.1540-5915.2012.00394.x

[57] Taguchi, G. (1986). Introduction to quality engineering: Designing quality into products and processes. Asian Productivity Organization.

[58] Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal, 18(7), 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z

[59] Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3), 207–222. https://doi.org/10.1111/1467-8551.00375

[60] Vilaplana, F., & Karlsson, S. (2008). Quality concepts for the improved use of recycled polymeric materials: A review. Macromolecular Materials and Engineering, 293(4), 274–297. https://doi.org/10.1002/mame.200700393

[61] Villanueva, A., & Wenzel, H. (2007). Paper waste—Recycling, incineration or landfilling? A review of existing life cycle assessments. Waste Management, 27(8), S29–S46. https://doi.org/10.1016/j.wasman.2007.02.019

[62] Vlachos, D., Georgiadis, P., & Iakovou, E. (2007). A system dynamics model for dynamic capacity planning of remanufacturing in closed-loop supply chains. Computers & Operations Research, 34(2), 367–394. https://doi.org/10.1016/j.cor.2005.03.005

[63] Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii–xxiii.

[64] Williamson, O. E. (1985). The economic institutions of capitalism: Firms, markets, relational contracting. Free Press.

[65] Zikopoulos, C., & Tagaras, G. (2015). Reverse supply chains: Effects of collection effort and quality of returns on the manufacturer's profit. European Journal of Operational Research, 246(1), 368–377. https://doi.org/10.1016/j.ejor.2015.04.014

Optical reflectance-based prediction of soil moisture and load-bearing capacity

[1]Nimród FEKETE, [1]Péter KISS, [1]György PILLINGER
[1] *Department of Vehicle Technology, Institute of Technology, Hungarian University of Agriculture and Life Sciences,MATE, Gödöllő*

Abstract

Assessing soil condition, particularly load-bearing capacity, is essential for route planning of autonomous off-road vehicles. Direct in situ measurement is difficult, making remote sensing–based estimation an emerging research focus. Laboratory tests on six soil textures revealed a strong correlation between visible-range reflectance and load-bearing capacity. Bevameter measurements were conducted at increasing moisture levels from dry to saturated states, enabling analysis of moisture effects on mechanical behavior. Reflectance was measured with a portable spectrophotometer, using the dry-state reflectance as a reference to account for inherent soil color. By comparing optical and mechanical data, a predictive equation was developed that incorporates dry-state reflectance constants. Although derived under controlled conditions, the model provides a promising basis for estimating soil load-bearing capacity, with further validation required under field conditions.

Keywords:

Soil color, Remote sensing, Optical reflectance, Soil condition, Load-bearing capacity, Soil moisture content

1. Introduction

Soil reflectance in the visible range depends on moisture and composition, making it a proxy for soil state. Reflectance decreases with moisture due to scattering and absorption (Philpot, 2010; Bowers and Hanks, 1965) and can estimate properties such as moisture, texture, organic content (Whalley and Leeds-Harrison, 1991; Pinheiro et al., 2017), water retention (Bedidi et al., 1992), and is controlled by particle size and composition including carbon and iron (Simon et al., 2020; Yost et al., 2019). Reflectance–moisture relationships are similar across soils, differing mainly in dry-state reflectance (Pillinger et al., 2023), while at the particle level water affects both optical and mechanical behavior (Kézdi, 1977). Soil moisture also governs mechanical properties, reducing strength with saturation (Mahinroosta and Poorjafar, 2017; Yoon and

Go, 2024); approaching the liquid limit sharply lowers load-bearing capacity, affecting vehicle mobility (Pundir and Garg, 2022; Wijayathunga et al., 2023; Jiang et al., 2025). Conventional methods (cone index, CBR, bevameter) are accurate but point-based and labor-intensive (Li et al., 2022), while incorporating moisture improves mobility prediction (Stevens et al., 2017). This study links optical reflectance and mechanical behavior, using established models (Nihal, 2022; Pillinger et al., 2023; Pillinger et al., 2024) to develop a framework predicting soil load-bearing capacity directly from reflectance.

2. Material and method

Properly moistened soil was layered into the compaction cylinder to form a minimum-volume sample capable of supporting its own weight without significant lower-layer compression. The sample was then compacted stepwise using a hydraulic press. At constant moisture content, compaction progressively reduced porosity while proportionally increasing saturation and bulk density.

Soil approached saturation depending on initial moisture and applied pressure, expelling excess water. Color was measured after filling and each compaction stage; the dry state was defined after oven drying at 105 °C for 48 h. Moisture content was calculated from dry mass and added water (w/w %). Full saturation was achieved via a drip method in a ceramic container until droplets were absorbed and surface reflectance stabilized. Reflectance was measured with a Konica Minolta CM-700d spectrophotometer under D65 illumination, 10° observation, SCE mode, using diffuse reflectance from 400–700 nm with 0–175 % range and 8 mm (MAV) aperture. Particle size distribution matched the collection site, removing only >0.5 mm particles. This section describes the experimental setup linking soil reflectance and load-bearing capacity, including controlled sample preparation, bevameter testing, and spectrophotometric measurements.

Soil sample preparation

Eight soil textures were prepared under controlled laboratory conditions to cover a wide range of optical and mechanical properties (Table 1). Samples were air-dried, sieved (<2 mm) to remove aggregates, and stored in sealed containers for uniform initial conditions. Multiple moisture levels were achieved by incrementally adding deionized water to oven-dry soil and mixing thoroughly. After each step, samples equilibrated in sealed containers for ≥12 h, then compacted into cylindrical containers for testing. Moisture content was verified using an HE73 analyzer (1 mg accuracy, 105 °C drying) and oven drying, calculated from mass loss and reported on a dry mass basis (%).

Table 1: Investigated soil textures

Soil texture	Clay [$_m$%]	Silt [$_m$%]	Sand [$_m$%]
Sand	2.20	2.10	95.7
Sandy loam	11.9	16.5	71.6
Loamy sand	5.80	12.1	82.1
Silty Clay	44.6	52.1	3.30
Silty loam	26.2	62.2	11.6
Loam	19.0	49.7	31.3
Clay loam	28.5	40.7	30.8
Clay	55.6	31.3	13.1

Load-bearing capacity measurement using a bevameter

Soil samples were prepared directly in the bevameter bin for consistency, controlling both moisture and initial bulk density. Soil was layered and compacted to ensure uniform density and homogeneity. The bevameter applied vertical load and measured deformation to determine load-bearing capacity under different moisture conditions. Ground pressure varies by vehicle: SUVs/trucks ~0.7–2.1 bar, ATVs/UTVs ~0.2–0.6 bar, tracked vehicles ~0.07–0.35 bar, with a maximum of 3 bar representing typical off-road limits. Laboratory measurements must consider bottom and sidewall effects; adequate soil thickness minimizes these boundary influences. The relationship between pressure plate diameter (d), soil thickness (H), and critical depth (z_0/d) is as follows:

$$\frac{z_0}{d} = \frac{H}{d} - 1 \tag{1}$$

When the soil cone beneath the plate reaches critical depth, bottom effects affect load measurement. To ensure valid results up to z/d = 0.4, minimum soil thicknesses are 21 cm (d = 15 cm) and 28 cm (d = 20 cm); with 30 cm soil, critical depths are 15 cm and 10 cm, giving H/d = 1.5. Typical Hungarian topsoil bulk density ranges 1.1–1.6 g/cm³ (average ~1.2 g/cm³); 1.4 g/cm³ was used. Reflectance was measured immediately before loading (400–700 nm) using a Konica Minolta CM-700d spectrophotometer with SCE mode, internal illumination, and 8 mm aperture (0–175 %), minimizing ambient effects. Dry and saturated states were characterized: dry after oven drying at 105 °C for 48 h

(w/w %), saturated via drip method until full absorption and surface reflectance stabilized.

3. Results

Effect of initial soil color

To evaluate the effect of initial soil color and saturation, reflectance curves in the 600–700 nm range were analyzed at different moisture levels. Figures 1 and 2 show results for loamy sand (0–18%) and sand (0–16%). Due to higher water retention, loamy sand saturates at higher moisture (~35%) than sand (~30%). The upper curve corresponds to dry soil, while lower curves represent increasing moisture, with the lowest indicating full saturation.

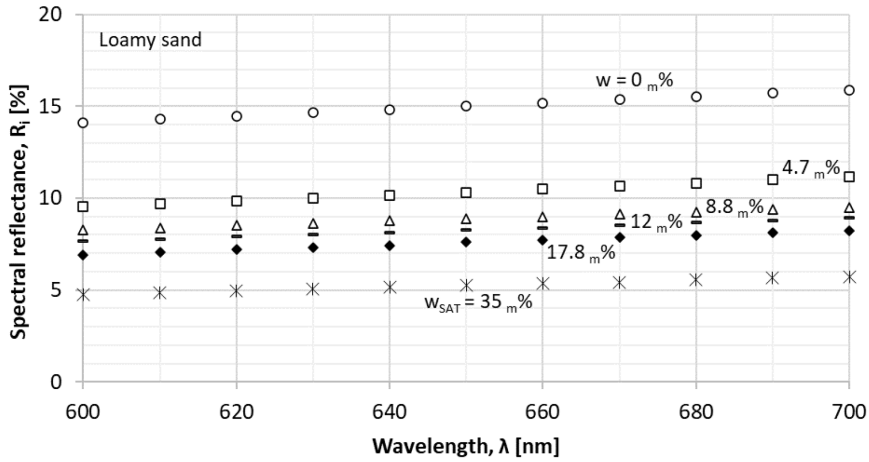


Figure. 1. Wavelength-reflectance relationship of loamy sand soil for different soil moisture contents

It can be seen from the reflectance values that the examined sandy soil starts with a slightly lighter initial value of R700 = 19.5%, and accordingly its wetted values take on higher values compared to the loamy sandy soil, which has a dry R700 value of 15.9%.

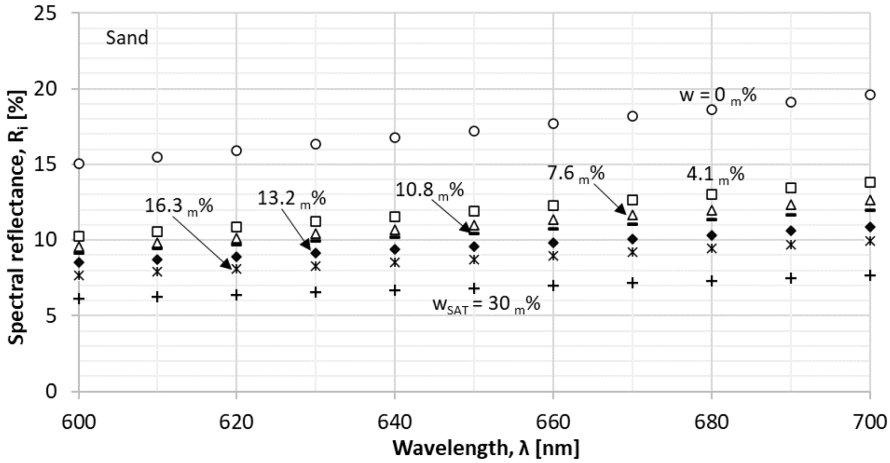


Figure 2. Wavelength-reflectance relationship of sandy soil for different soil moisture contents

Reflectance was quantified using the C-value (Eq. 2), the average over the wavelength range. *C* plotted against moisture (Fig. 3) shows trends consistent with Pillinger et al. (2023), though at lower resolution. Reflectance (log scale) decreases exponentially with moisture; the same relationships hold with different constants (*A, n*).

$$C = \frac{1}{11} \cdot \left(\sum_{600}^{700} R_i \right) \tag{2}$$

Using average constants reduces accuracy. Key points of the curve are the dry-state reflectance (C_0, e.g., 17.3% for sand) and the saturated reflectance (C_{SAT}). Between these, reflectance (C) decreases monotonically with increasing moisture.

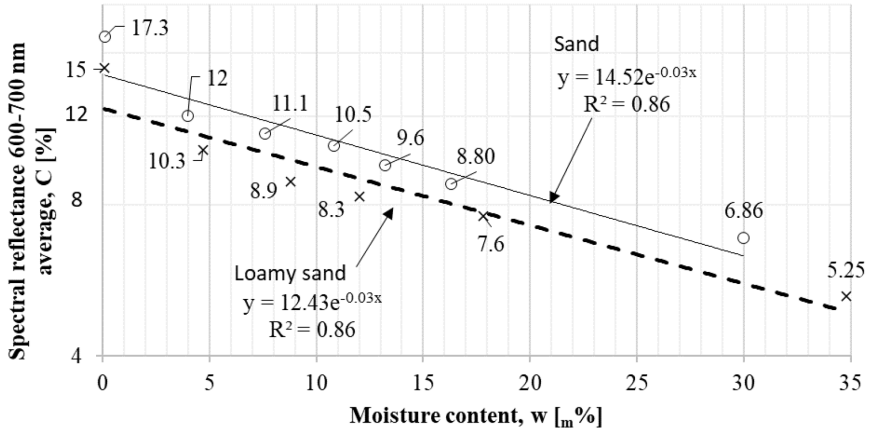


Figure 3. The change of $C_{600\text{-}700}$ as a function of moisture content, on a semi-logarithmic scale

The next step is to calculate the deviation of each "*C*" value from "C_0". The deviations are determined with the relation:

$$\Delta C = C_0 - C \tag{3}$$

Results are plotted on a linear-linear diagram. Both ΔC and moisture increase, though modestly. The moisture at the maximum corresponds to saturation, and $\Delta C = 0$ % for completely dry soil.

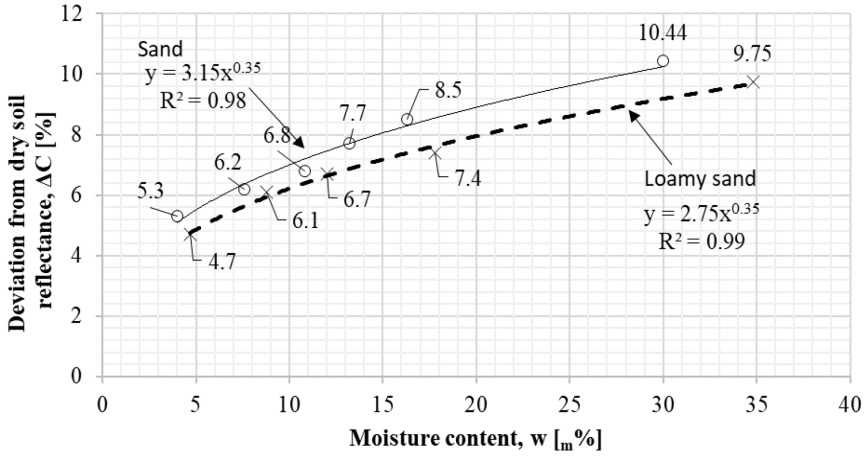


Figure 4. Deviation of the average of the reflectance curves from the maximum value (from the reflectance value of completely dry soil)
Based on Fig. 4 it can be concluded that there is a power function relationship between the moisture content and the ΔC value:

$$\Delta C = A \cdot w^n \tag{4}$$

The equation (5) must be replaced back into the basic relationship "ΔC" (2), so we get the reflectance parameter "C" for the wet state of the soil as a function of the moisture content "w" and the reflectance parameter "C_0" of the dry soil:

$$C = C_0 - A \cdot w^n \tag{5}$$

The relationship has two constants, "A" and the exponent of moisture content "n". For accurate calculations, the values of the constants for the tested soil textures are shown in Table 2.

Table 2: The value of the constants

Soil texture	A [%]	n [-]
Sand	3.15	0.35
Loamy sand	2.75	0.35

Using equation (6), the moisture content of the soil can now also be determined using the dry reflectance parameter "C_0" of the soil and the current reflectance parameter "C" of the soil:

$$w = \sqrt[n]{\frac{C_0 - C}{A}} \tag{6}$$

Reflectance-Based Results and Analysis

Based on previous studies the color number ($\mathcal{C}$) is defined as the average reflectance value measured by the spectrophotometer within the wavelength range of 600–700 nm (Eq. 7).

$$C = \frac{1}{11} \cdot \left(\sum_{600}^{700} R_i \right) \tag{7}$$

Table 3 presents the tested soil textures. These include the optical reflectance number recorded at both the completely dry (C_0) and fully saturated states (C_{SAT}) at the saturation moisture content (ω_{SAT}), and the difference between the two reflectance extremes (ΔC_{max}) representing the measurable range of reflectance variation.

Table 3: Optical reflectance properties of investigated soil textures

	Dry	Fully saturated		Reflectance range
Soil texture	C_0 [%]	C_{SAT} [%]	ω_{SAT} [m%]	ΔC_{max} [%]
Sand	17.5	6.9	30.0	10.6
Sandy loam	34.4	13.4	31.9	21.0
Loamy sand	15.0	5.3	34.8	9.70
Silty Clay	18.4	7.1	36.4	11.3
Silty loam	12.2	5.0	36.6	7.20
Loam	20.5	5.5	42.1	15.0
Clay loam	19.7	7.0	47.2	12.7
Clay	15.8	5.5	53.5	10.3

Figures 5 and 6 show that reflectance varies most at low moisture, with changes slowing toward saturation. At full saturation, surface water prevents reliable colorimeter readings, so the usable reflectance range depends on soil texture.

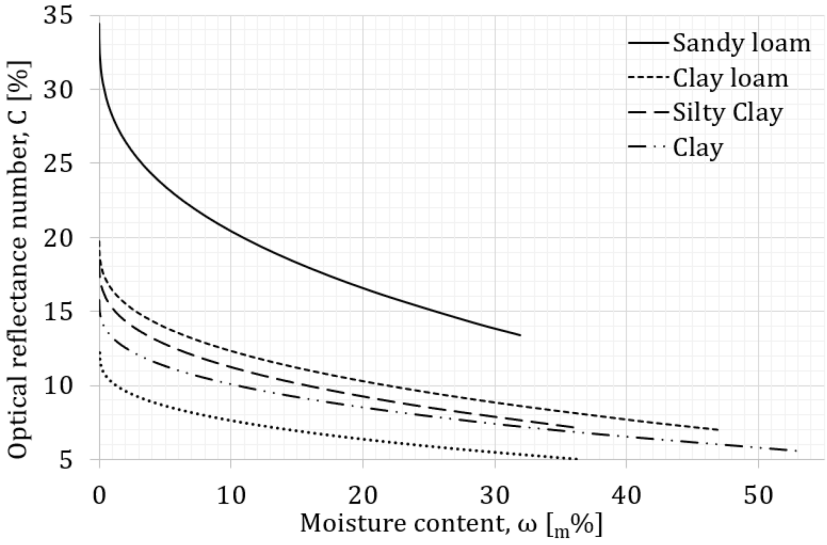


Figure 5. Optical reflectance number as a function of moisture content for different soil textures

The relationship between soil moisture content and the reflectance number is defined by Equation (8).

$$\omega = \sqrt[n_C]{\frac{(C_0 - C)}{A}} \tag{8}$$

This formulation links moisture content (ω) to the decrease in optical reflectance from its dry-state value (C_o). The difference ($C_o - C$) measures reflectance reduction, normalized by A, while the (n_C)-th root introduces a soil-specific nonlinear sensitivity. Higher n_C gives a gradual moisture increase with decreasing reflectance, lower n_C produces a steeper response. Table 7 lists the parameters for Equation 8.

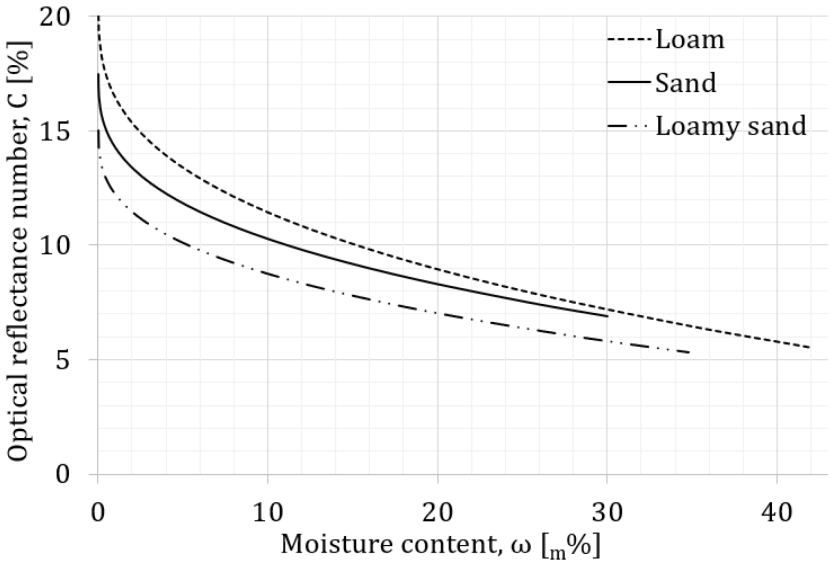


Figure. 6. Optical reflectance number as a function of moisture content for different soil textures

Figures 6 and 7 show a nonlinear decrease in reflectance (C) with increasing moisture (ω) for all soils. Reflectance is highly sensitive in the dry to moderately moist range, with a steep initial decline and flattening near saturation, indicating reduced sensitivity at higher moisture levels. Soil texture influences curve shape: sandy and loamy soils exhibit higher dry reflectance and larger variation, while clay and silty soils show lower initial values and more gradual changes due to finer structure and increased light absorption.

Table 7: Equation 8 constants for various soil textures

Soil texture	Constants	
	A [-]	n_C [-]
Sand	3.2	0.35
Sandy loam	6.3	0.35
Loamy sand	2.8	0.35
Silty Clay	3.2	0.35
Silty loam	2.0	0.35
Loam	4.1	0.35
Clay loam	3.3	0.35
Clay	2.6	0.35

Evaluation of Load-Bearing Behavior

By analyzing the sinkage–pressure curves shown in Figure 8, the load-bearing capacity factor (k) can be determined for any given moisture content. The figure illustrates the relationship between relative depth (z/d) and ground pressure (p) for sandy loam soil, measured using a pressure plate of diameter d=20 cm and a bulk density of ρ=1.4 g/cm^3. The curves represent different moisture contents ranging from 0 % to 25 % by mass and are shown as interpolated trend lines derived from experimental data points.

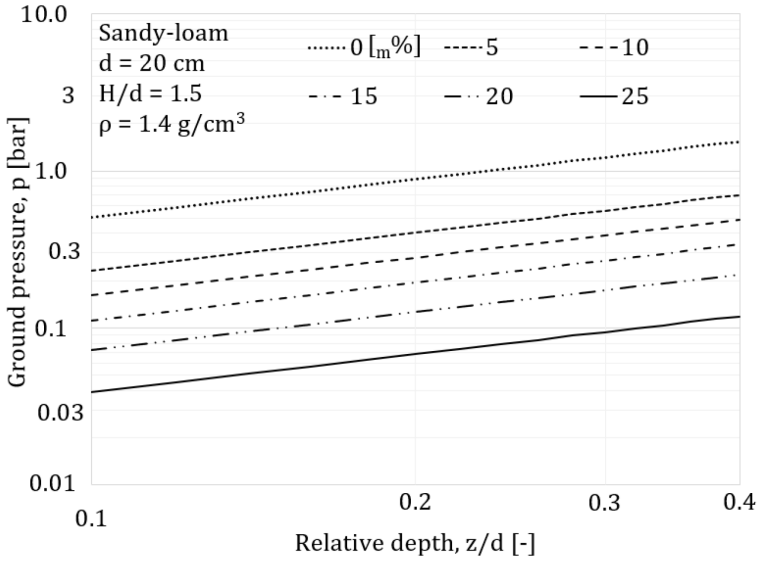
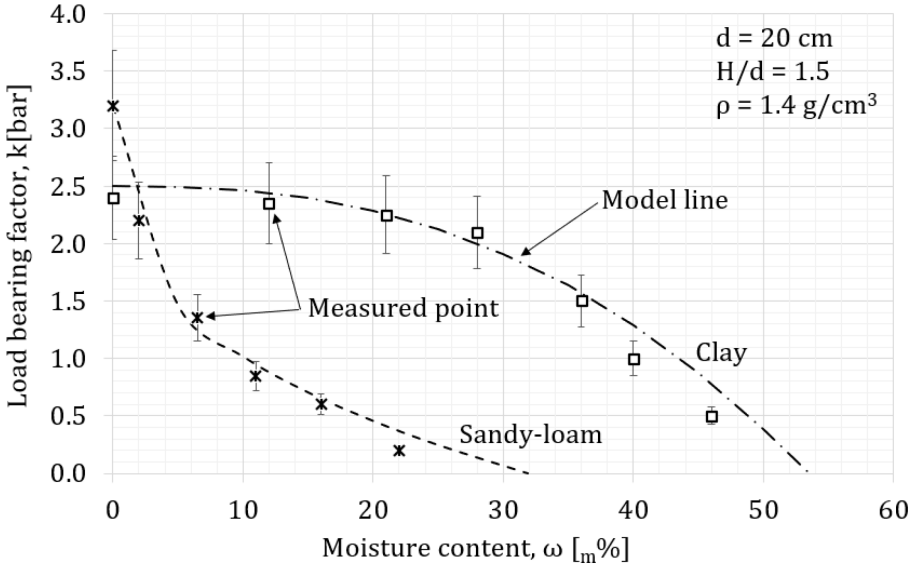


Figure 7. Measured ground pressure as a function of relative sinkage for different moisture contents

The curves show that the ground pressure needed for a given normalized depth decreases with increasing moisture, reflecting water's softening effect by lowering inter-particle friction and cohesion. The largest pressure drop occurs at low moisture (0–10 %), while higher contents show a gradual effect. These curves also enable calculation of the load-bearing capacity factor (k) using the Saakyan formula (Eq. 9), providing a single-value measure of soil strength.

$$k = p \cdot \left(\frac{d}{z}\right)^{-n} \tag{9}$$

Here, p is ground pressure, z is sinkage, d is plate diameter, and n = 0.8 is an empirical exponent describing the pressure–sinkage curve. Figure 9 shows load-bearing factor (k) versus moisture (ω) for sandy loam and clay. Sandy loam decreases rapidly, especially below 15 %, while clay declines more slowly, retaining higher strength up to 20–30 % moisture. This reflects soil structure: clay's fine particles and high surface area maintain cohesion and capillary bonding, whereas sandy loam loses strength quickly as moisture reduces interparticle friction.



Figre 8. Load-bearing capacity factor as a function of moisture content

In figure 8 measured values are indicated by black crosses for sandy loam and black squares for clay, while the dashed lines represent the fitted model curves for each soil type, as described by Equation (10).

$$k = k_0 \cdot \left(1 - \left(\frac{\omega}{\omega_{SAT}}\right)^{n_\omega}\right) \tag{10}$$

The model captures the nonlinear decline of k with moisture, where k_0 is the dry-soil factor and $n\omega$ an empirical exponent. The factor drops sharply at initial wetting, with effects tapering near saturation. Error bars (±15 %) reflect measurement variability and soil heterogeneity.

Correlation between Optical Reflectance and Load-Bearing Capacity
Equation (11) establishes a predictive relationship between soil reflectance and its load-bearing response by integrating optical and mechanical parameters. By combining Equations (8), (9), and (10), we derive a model that predicts ground pressure p based on the soil's optical reflectance characteristics. This model incorporates constants representing the soil's reflectance and strength properties in its dry state.

$$p = k_0 \cdot \left[1 - \frac{(C_0 - C)^m}{\omega_{SAT}{}^{n\omega} \cdot A^m}\right] \cdot \left(\frac{z}{d}\right)^n \tag{11}$$

Here, k_0 is the load-bearing factor and C_0 the optical reflectance of dry soil. The exponent $m = n\omega/nC$ links moisture and reflectance sensitivity (for the studied sandy loam, $m = 0.94$). The term $(z/d)^n$ represents the soil's mechanical response per the Saakyan equation. Equation (11) assumes uniform moisture and density throughout the sample; gradients or density variations may reduce predictive accuracy.

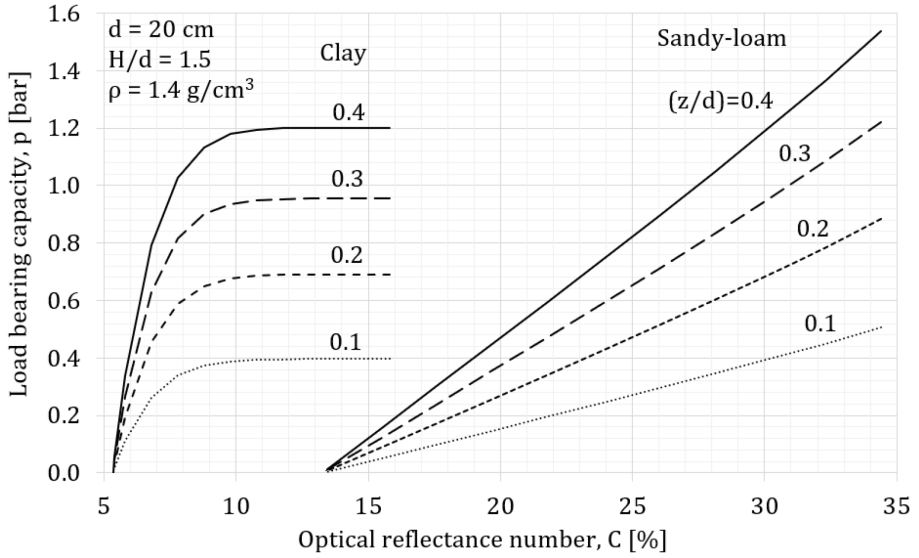


Figure 9. Load-bearing capacity as a function of the optical reflectance number for different z/d values

Figure 9 shows the modeled relationship between optical reflectance (C) and load-bearing capacity (p) for clay and sandy loam at relative depths z/d = 0.1–

0.4, assuming ρ = 1.4 g/cm³ and d = 20 cm (Eq. 11). For sandy loam, p increases almost linearly with C, reflecting proportional strength gains in drier conditions. Clay shows a nonlinear trend: p rises with C up to a threshold, then plateaus, due to high moisture retention and cohesion. Sandy loam exhibits higher reflectance sensitivity at all depths, while clay's response flattens at greater *z/d*, showing reduced dependence on surface reflectance near saturation.

Table 8 Equation 11 constants

	Sandy-loam	Clay
A	6.30	2.60
C_0	34.4	15.8
k_0	3.20	2.50
n	0.80	0.80
n_ω	0.33	2.50
m	0.94	7.10
ω_{SAT}	31.9	53.5

Table 8 lists the parameters of Equation (11) for sandy loam and clay. Dry-state reflectance (C_0) is higher for sandy loam (34.4 %) than clay (15.8 %), reflecting brighter, coarser texture. Dry-state load-bearing factor (k_0) is also higher for sandy loam (3.2 bar vs. 2.5 bar). Saturation moisture (ω_{SAT}) differs: 31.9 % for sandy loam, 53.5 % for clay, while both share n = 0.8. The exponent m, controlling reflectance–strength nonlinearity, is 0.94 for sandy loam (near-linear) and 7.1 for clay (sharp drop with moisture), highlighting texture-specific calibration needs and Equation (11)'s flexibility.

4. Conclusions and recommendations

his study demonstrates that soil surface reflectance in the visible range shows a strong and predictable correlation with load-bearing capacity under controlled conditions. Its applicability is limited near saturation, where reduced moisture sensitivity decreases accuracy.

Using combined optical and mechanical measurements on multiple soil textures (e.g., sandy loam and clay), it was shown that both moisture content and reflectance significantly affect soil strength. The reflectance number (C) and derived models enable estimation of the load-bearing factor (k) and ground pressure (p) as functions of reflectance and relative depth.

A key outcome is the composite predictive equation (Eq. 11), integrating optical and mechanical parameters. This enables practical estimation of soil bearing capacity from reflectance, supporting real-time, non-invasive terrain assessment for autonomous off-road vehicles. The results also highlight the importance of uniform moisture distribution, as model validity depends on consistent soil moisture. Sandy soils wet rapidly due to high permeability, while clay soils absorb water more slowly, often leading to vertical moisture gradients. As a result, surface measurements alone may indicate saturation and low strength, even if deeper layers remain relatively dry and stable.

Soils with similar textures can exhibit significant reflectance variability; therefore, reference values (C_0 and k_0) must be determined individually. For absolute characterization, additional parameters are required. Future work will focus on identifying representative exponent values for sandy and clay soils to improve model generalizability.

Although based on controlled laboratory conditions, the results provide a strong basis for field application. Future work will include validation under real conditions, where surface roughness, vegetation, lighting, and moisture gradients introduce complexity. The integration of remote sensing systems (e.g., UAV-based spectrometers or multispectral cameras) may further enhance the applicability of the approach.

Future work will focus on:

- Refining the model parameters to improve generalizability across soil types,
- Extending the reflectance-based model to predict traction and slip, by integrating optical soil parameters with wheel–soil interaction measurements.
- Validating the method under real-world conditions using UAV-based multispectral sensing, to assess performance across heterogeneous terrain and varying illumination.
- Developing a predictive mobility framework for autonomous and military vehicles, enabling real-time decision support based on reflectance-derived soil parameters.
- Investigating additional soil-related mobility risks, such as mud adhesion, sensor contamination, and mass accumulation effects inferred from surface conditions.

Overall, the methodology and models presented here contribute to a growing body of research aimed at improving terrain-awareness and decision-making for autonomous systems operating in complex, off-road environments.

5. Nomenclature

A	Soil reflectance specific constant	[-]
C	Average reflectance values in the range of 600-700 nm	[%]
C_0	Average reflectance values in the range of 600-700 nm for dry soil	[%]
C_{SAT}	Average reflectance values in the range of 600-700 nm for fully saturated soil	[%]
ω_{SAT}	Saturated moisture content of soil	[$_m$%]
Λ	Visible light wavelength	[nm]
ΔC_{max}	Range of reflectance variation	[%]
d	Diameter of the pressure plate	[cm]
k	Load-bearing capacity factor	[bar]
k_0	Load-bearing factor at dry state	[bar]
m	Composite exponent, $m=n_\omega/n_C$	[–]
n	Shape exponent in the pressure–sinkage relationship (Saakyan model)	[–]
n_C	Reflectance sensitivity exponent	[–]
n_ω	Moisture sensitivity exponent	[–]
p	Ground pressure	[bar]
R_i	Reflectance value at a given wavelength	[%]
ω_0	Moisture content on a dry mass basis	[$_m$%]
z	Vertical sinking	[cm]
z/d	Relative depth	[–]

6. Acknowledgements

Project no. 2022-2.1.1-NL-2022-00012 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the 2022-2.1.1-NL funding scheme.

7. References

[1] Aksakal, E. L., Angin, I., Sari, S. (2021):
[2] Effects of freeze–thaw cycles on consistency limits of soils amended with diatomite. Soil & Tillage Research, 213, 105144.
[3] Hasan, M. F., Abuel-Naga, H. (2024):
[4] Determining Liquid Limit and Plastic Limit of Clay Soils by Electrical Surface Conduction and Diffuse Double Layer Thickness. Minerals, 14(3), 210.
[5] Karumanchi, M., Avula, G., Pangi, R., Sirigiri, S. (2020):

[6] Improvement of consistency limits, specific gravities, and permeability characteristics of soft soil with nanomaterial: Nanoclay. Materials Today: Proceedings.
[7] O'Kelly, B. C. (2024):
[8] Theory of liquid and plastic limits for fine soils, methods of determination and outlook. Geotechnical Research, 11(1), 43–61.
[9] Rehman, H. U., Knadel, M., Kayabali, K., Arthur, E. (2019):
[10] Estimating Atterberg Limits of Fine-Grained Soils by Visible–Near-Infrared Spectroscopy. Vadose Zone Journal.
[11] Bekker, M. G. (1971):
[12] Russian Approach to Terrain-Vehicle Systems (An Exercise in Pragmatism and Continuity). U.S. Army Research Office.
[13] Salman, N. D., Pillinger, G., Kiss, P. (2020):
[14] Soil behaviour under load in case of finite thickness. International Review of Applied Sciences and Engineering.
[15] Hasan, M. F., Abuel-Naga, H. (2024):
[16] Determining Liquid Limit and Plastic Limit of Clay Soils by Electrical Surface Conduction and Diffuse Double Layer Thickness. Minerals, 14(3), 210.
[17] Rehman, H. U., Knadel, M., Kayabali, K., Arthur, E. (2019):
[18] Estimating Atterberg Limits of Fine-Grained Soils by Visible–Near-Infrared Spectroscopy. Vadose Zone Journal, 18(1).
[19] Bedidi, A., B. Cervelle, J. Madeira, and M. Pouget, 1992. Moisture effects on visible spectralcharacteristics of lateritic soils. Soil Science, Vol. 153(2):129-141.
[20] Jenifer L. Yost, Eric E. Roden and Alfred E. Hartemink, 2019. Geochemical Fingerprint and Soil Carbon of Sandy Alfisols, Soil Syst. 2019, 3(3), 59; https://doi.org/10.3390/soilsystems3030059
[21] Kézdi Á., 1977. Talajmechanika I. Tankönyvkiadó, ISBN: 963 17 2256 2
[22] Pillinger Gy., Ahmed E.E.A., Bessenyei K., Kiss P., 2023. Correlation between moisture content and color spectrum of sandy soils, Journal of Terramechanics, Vol. 108, page 39-45. https://doi.org/10.1016/j.jterra.2023.05.002
[23] Reza Mahinroosta, Arash Poorjafar, 2017. Effect of stress state and particle-size distribution on the stress reduction of sandy soils during saturation; Construction and Building Materials Vol. 150; https://doi.org/10.1016/j.conbuildmat.2017.05.177
[24] Seok Yoon, Gyu-Hyun Go, 2024. Numerical analysis study on determination of unfrozen water saturation curve for sandy soil, International Communications in Heat and Mass Transfer, Vol. 154, https://doi.org/10.1016/j.icheatmasstransfer.2024.107418
[25] Thibaut Simon, Yakun Zhang, Alfred E. Hartemink, Jingyi Huang, Christian Walter, Jenifer L. Yost, 2020. Predicting the color of sandy

soils from Wisconsin, USA; Goderma Vol. 361; https://doi.org/10.1016/j.geoderma.2019.114039

[26] William Philpot, 2010. Spectral Reflectance of Wetted Soils, http://dx.doi.org/10.13140/2.1.2306.0169

[27] Bowers, S. A., & Hanks, R. J. (1965). Reflection of radiant energy from soils. Soil Science, 100(2), 130–138. https://doi.org/10.1097/00010694-196508000-00009

[28] Jiang, C., Wei, Y., Chen, Z., & Kuang, X. (2025). The prediction method and application of off-road mobility for ground vehicles: A review. World Electric Vehicle Journal, 16(1), 47. https://doi.org/10.3390/wevj16010047

[29] Li, T., Mu, T., Liu, G., Yang, X., Zhu, G., & Shang, C. (2022). A method of soil moisture content estimation at various soil organic matter conditions based on soil reflectance. Remote Sensing, 14(10), 2411. https://doi.org/10.3390/rs14102411

[30] Pinheiro, É. F. M., Pereira, M. G., & Anjos, L. H. C. (2017). Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the Central Amazon. Remote Sensing, 9(4), 293. https://doi.org/10.3390/rs9040293

[31] Pillinger, G., Ahmed, E. E. A., & Bessenyei, K. (2023). Correlations between moisture content and color spectrum of sandy soils. Journal of Terramechanics, 108, 39–45. https://doi.org/10.1016/j.jterra.2023.05.002

[32] Pillinger, G., Ahmed, E. E. A., & Kiss, P. (2024). Spectral determination of soil moisture content based on the dry color of the soil. In Proceedings of the 21st International and 12th Asia-Pacific Regional Conference of the ISTVS. https://doi.org/10.56884/GV41C6MW

[33] Pundir, S. K., & Garg, R. D. (2022). A comprehensive approach for off-road trafficability evaluation and development of modified equation for estimation of RCI to assess regional soil variation using geospatial technology. Quaternary Science Advances, 5, 100042. https://doi.org/10.1016/j.qsa.2021.100042

[34] Salman, N. D. (2022). Load bearing capacity of soil as a homogeneous finite half-space (Doctoral dissertation, Hungarian University of Agriculture and Life Sciences). https://phd.mater.uni-mate.hu/199/2/Thesis_Nihal_DOI.pdf

[35] Stevens, M. T., McKinley, G. B., & Vahedifard, F. (2017). Full-featured ground vehicle mobility analysis using different soil moisture sources. International Journal of Vehicle Performance, 3(1), 19–35. https://doi.org/10.1504/IJVP.2017.081262

[36] Whalley, R. W., & Leeds-Harrison, B. P. (1991). Estimation of soil moisture status using near infrared reflectance. Hydrological Processes, 5, 321–327. https://doi.org/10.1002/hyp.3360050312

[37] Wijayathunga, L., Rassau, A., & Chai, D. (2023). Challenges and solutions for autonomous ground robot scene understanding and navigation in unstructured outdoor environments: A review. Applied Sciences, 13(17), 9877. https://doi.org/10.3390/app13109877