

Identifikation inhomogener Materialeigenschaften von Flächentragwerken mit Physics Informed Neural Networks

Lukas Lippold, Niklas Rödiger, Thomas Most, Carsten Könke

Institut für Strukturmechanik, Bauhaus-Universität Weimar

Zusammenfassung: Theorie-basierte Neuronale Netze (NN) beziehen physikalisches Vorwissen in Form von Differentialgleichungen und Randbedingungen in das Training ein und können verwendet werden, um simple ein- und zweidimensionale Probleme der Statik zu lösen. Für einen Zugstab unter Eigengewicht wird zunächst der Dehnungsverlauf berechnet. Weiterhin wird für eine Scheibe eine analytische Fourierlösung mit der Ausgabe eines optimierten NN verglichen. Zuletzt wird ein NN verwendet, um einen Riss in einer Scheibe mithilfe virtueller Messwerte zu identifizieren.

1 Physics Informed Neural Networks

Wie sein biologisches Vorbild besteht ein künstliches neuronales Netz aus Neuronen, welche eingehende Signale verbundener Neuronen aufnehmen und zu einem Ausgangssignal transformieren. Über die Stärke der Neuronenantwort entscheidet eine stetige, nicht-lineare Aktivierungsfunktion, welche als Argument eine linear gewichtete Summe der Eingangssignale einliest. Mithilfe eines Optimierungsalgorithmus können die Gewichtungen schrittweise so angepasst werden, dass ein beliebiger Zusammenhang zwischen Eingangs- und Ausgangsgrößen angenähert wird. Die Genauigkeit dieser Näherung hängt von der Wahl der Netzwerkstruktur und -parameter ab.

Im klassischen datengetriebenen Ansatz werden die NN mithilfe großer Datensätze von Eingangs-Ausgangswert-Paaren trainiert. Die zu minimierende Verlustfunktion \mathcal{L} ist ein Maß dafür, wie gut die Vorhersagen des NN die Trainingsdaten abbilden. Neben den häufig verwendeten Varianten wie dem Feedforward NN oder Convolutional NN wurden in [1] die Physics Informed Neural Networks (PINN) eingeführt. Diese können verwendet werden, um Funktionen $u(x)$ zu approximieren, welche bspw. die Lösung einer Differentialgleichung (DGL) darstellen, und verzichten dabei gänzlich auf präparierte Trainingsdaten. Stattdessen

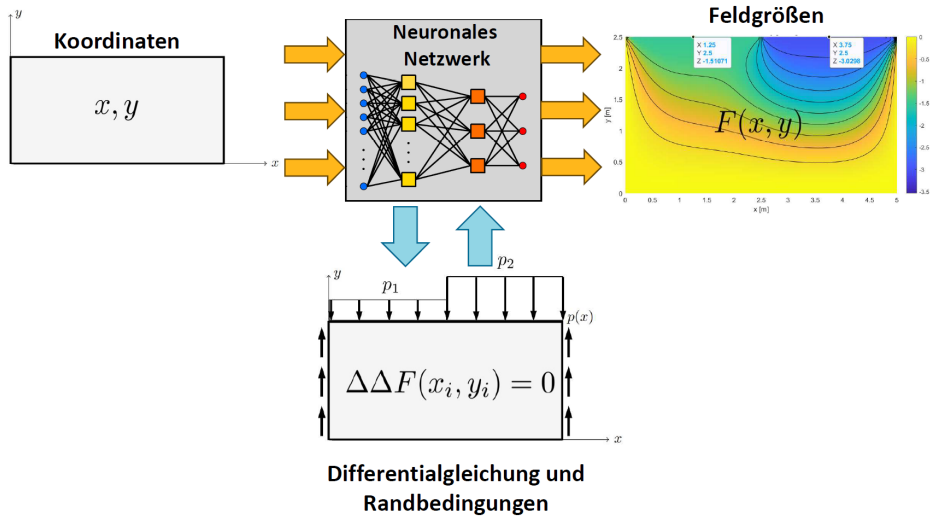


Abbildung 1: Physics Informed Neural Network - Approximation von mechanischen Feldgrößen unter Berücksichtigung von Differentialgleichungen und Randbedingungen

wird für die Verlustberechnung an Kollokationspunkten p_i ausgewertet, wie genau die Lösung des Modells die zugrundeliegende DGL, beschrieben durch den Differentialoperator $\mathcal{N}[u] = 0$, löst. Zusätzlich können Rand- und Anfangsbedingungen (RB) an entsprechenden Punkten \tilde{p}_i berücksichtigt werden. Die Verlustfunktion lässt sich mit den mittleren quadratischen Fehlern schreiben zu

$$\begin{aligned} \mathcal{L}_\theta &= \mathcal{L}_{\text{DGL}} + \mathcal{L}_{\text{RB}} \\ &= \frac{1}{N_c} \sum_i^{N_c} \{\mathcal{N}[u_\theta(p_i)]\}^2 + \frac{1}{N_r} \sum_i^{N_r} \{u_\theta(\tilde{p}_i) - u(\tilde{p}_i)\}^2. \end{aligned} \quad (1)$$

N_c und N_r stehen jeweils für die Anzahl an Kollokations- und Randpunkten. Der Vektor θ beinhaltet alle trainierbaren Netzwerkparameter, welche das Modell definieren. Das Prinzip der PINNs ist in Abb. 1 am Beispiel einer belasteten Scheibe veranschaulicht.

2 Lösung von Randwertproblemen

2.1 Zugstab unter Eigengewicht

Zunächst wird anhand eines einfachen eindimensionalen Problems die Funktionsweise der PINNs demonstriert. Das Verschiebungsfeld eines Zugstabes soll unter Berücksichtigung der Differentialgleichung sowie von Randbedingungen approximiert werden. Dabei wird die Verschiebung und Dehnung in Abhängigkeit der Position x formuliert. Es wurde in allen Rechnungen dieser Arbeit ein linear elastisches Materialverhalten angenommen. Für einen Zugstab der Dichte ρ , mit konstanten Werten für Querschnittfläche A und Elastizitätsmodul E ,

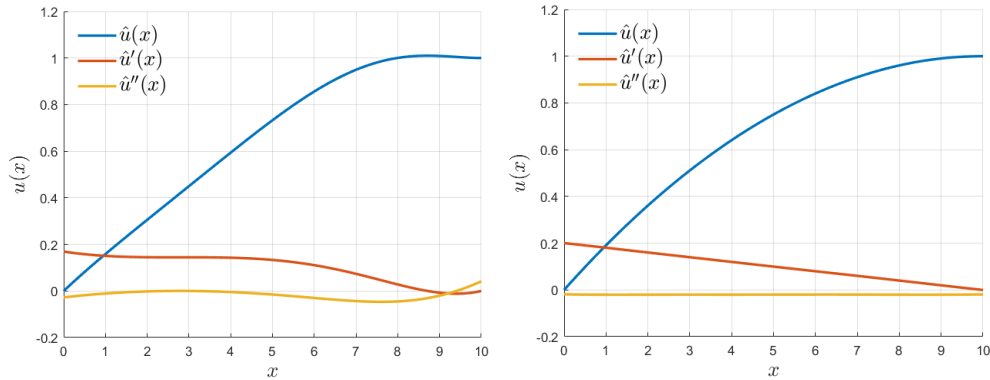


Abbildung 2: Approximation der Verschiebungsfunktion eines Zugstabes unter Eigengewicht: links unter Verwendung einer geringen Wichtung für die Einhaltung der Differentialgleichung ($\alpha = 0.01$) und rechts mit sehr hoher Wichtung ($\alpha = 10.0$)

der einer in x -Richtung wirkenden Gravitation g ausgesetzt ist gilt

$$u''(x) = -\frac{\rho \cdot g}{E} = konst. \quad (2)$$

Zur Lösung dieser Differentialgleichung wurde ein Netzwerk mit der Position x als Eingang und der Verschiebung $u(x)$ als Ausgang definiert. Bei diesem eindimensionalen Problem ist ein versteckter Layer ausreichend. Als Randbedingungen wurde der Stab bei $x = 0$ eingespannt und am anderen freien Ende eine Längenänderung Δl vorgegeben

$$u(0) = u'(l) = 0, \quad u(l) = \Delta l. \quad (3)$$

In der Verlustfunktion für das Netzwerktraining wurden die Abweichungen von den Randbedingungen und die Erfüllung der Differentialgleichung in Glg. 2 wie folgt berücksichtigt

$$L = (\hat{u}(0))^2 + (\Delta l - \hat{u}(l))^2 + (\hat{u}'(0))^2 + \alpha \cdot \sigma^2(\hat{u}''(x)) \rightarrow min, \quad (4)$$

wobei der Fehler in der Differentialgleichung als Varianz der zweiten Verschiebungsableitung definiert wurde, welchen an den Kollationspunkten ausgewertet wird. Der Faktor α ist ein Wichtungsfaktor, der die beiden Terme, Erfüllung der Randbedingungen sowie der Differentialgleichung, in der Verlustfunktion entsprechend gegeneinander wichten kann.

Für eine vorgegebene Längenänderung $\Delta l = 1.0$ m sowie eine Stablänge von $l = 10.0$ m wurde ein Netzwerk für verschiedene Wichtungsfaktoren mit der Verlustfunktion in Glg. 4 trainiert. Dabei wurden 10 Neuronen im versteckten Layer mit Tangens hyperbolicus Aktivierungsfunktionen verwendet. In Abb. 2 sind der approximierte Verschiebungsverlauf sowie die erste und zweite Ableitung dargestellt. In der Abbildung ist ersichtlich, dass bei zu geringer Wichtung der Differentialgleichung die Randbedingungen zwar eingehalten sind, die Verschiebungsfunktion jedoch nicht korrekt abgebildet wird. Wird dieser Anteil höher gewichtet, wird der mechanisch korrekte, quadratische Verlauf der Verschiebungsfunktion sehr gut repräsentiert. Die annähernd konstante zweite Ableitung der Approximationsfunktion $\hat{u}''(x)$

ergibt dabei automatisch den rechten Teil der Differentialgleichung in Glg. 2 und kann direkt zur Bestimmung des als konstant angenommenen Elastizitätsmoduls verwendet werden.

2.2 Scheibe unter Rechtecklast

2.2.1 Analytische Beschreibung

Im nächsten Schritt wird sich auf das einfache Beispiel einer Scheibe bezogen, ein Tragwerk dessen Dicke in z-Richtung klein gegenüber den anderen Raumrichtungen ist und welches nur Belastungen innerhalb der x-y-Ebene ausgesetzt ist. Der resultierende ebene Spannungszustand lässt sich vollständig über die Verschiebungen u und v in x- und y-Richtung beschreiben. Die Gleichgewichts- und Verträglichkeitsbedingungen für die drei am Scheibenelement wirkenden Spannungsgrößen σ_x , σ_y und $\tau_{xy} = \tau_{yx}$ lauten

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + X = 0, \quad (5)$$

$$\frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{xy}}{\partial x} + Y = 0, \quad (6)$$

$$\frac{\partial^2 \varepsilon_x}{\partial y^2} + \frac{\partial^2 \varepsilon_y}{\partial x^2} = \frac{\partial^2 \gamma_{xy}}{\partial x \partial x}. \quad (7)$$

X und Y sind die Volumenkräfte in x- und y-Richtung.

Es wird eine rechteckige Scheibe mit den Abmessungen 5 m x 2.5 m x 0.15 m konstruiert, welche durch eine un stetige Randlast an der Oberkante belastet wird. Es handelt sich hierbei um eine konstante Flächenlast von 20 kN/m², die zwischen den Viertelpunkten angreift. Die Scheibe ist an ihren Rändern gegen vertikale Verschiebungen gehalten.

Der sich einstellende Verlauf der Spannungsgrößen unter den gegebenen Randbedingungen wurde zunächst mit einem Fourieransatz analytisch gelöst und anschließend in Kap. 2.2.3 mit dem Ergebnis des PINNs aus Kap. 2.2.2 verglichen.

Für das Aufstellen einer analytischen Lösung werden Glg. 5- 7 mithilfe der Airyschen Spannungsfunktion in einer biharmonischen Differentialgleichung zusammengefasst [2]. Die un stetige Randlast wurde in einer Fourierreihe entwickelt und für jeden Reihenterm mithilfe biharmonischer Ansatzfunktionen eine Einzellösung berechnet, die zu einer Gesamtlösung superponiert wurden. Mittels FEM-Software [3] wurde unter Nutzung linearer 2D-Elemente zusätzlich eine numerische Vergleichslösung erstellt, welche in Abb. 3 zusammen mit den Randbedingungen dargestellt ist.

2.2.2 PINN Lösung

Es wurde ein PINN zur Lösung der Scheibengleichung mit den in Kap. 2.2.1 definierten Randbedingungen konstruiert. Statt die Airyfunktion als Hilfsmittel zu verwenden, können direkt die Verschiebungskomponenten u und v durch zwei separate neuronale Netzwerke modelliert werden. Alle Netzwerke verwenden eine Tangens hyperbolicus Aktivierungsfunktion und wurden zunächst mit ADAM, einem weiterentwickelten Gradientenabstiegsverfahren [4]

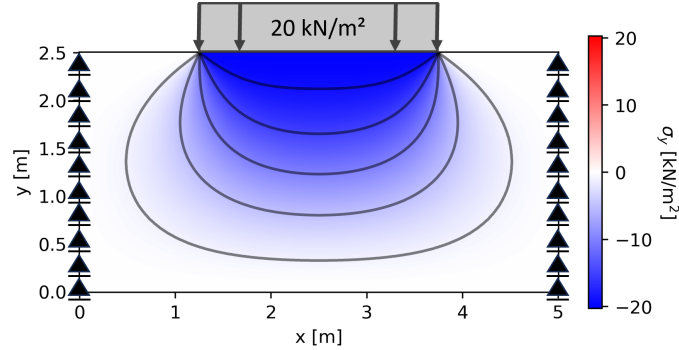


Abbildung 3: FEM Lösung der Scheibengleichung für die Spannungs Komponente σ_y einer mittels Schub eingespannten Scheibe mit Rechtecklast auf einem Gitter von 501 x 251 Knoten. Die Isolinienwerte betragen [-19,-15,-10,-5,-1] kN/m²

und anschließend mit einem Quasi-Newton Verfahren, dem Broyden–Fletcher–Goldfarb–Shanno (BFGS) Algorithmus [5], optimiert. Zum Erstellen und Trainieren der PINNs wird das Ten-sorflow Framework [6] und die Keras Bibliothek [7] verwendet.

Für den Operator $\mathcal{N}(u,v)$ der Glg. 1 wurden die Gleichgewichtsbedingungen 5 und 6 durch Einsetzen des linearen Elastizitätsgesetzes über die Verschiebungsgrößen u und v ausgedrückt. Die Volumenkräfte X und Y wurden zu Null angenommen

$$\mathcal{N}_1(u,v) = \frac{E}{1-\mu^2} \left(\frac{\partial^2 u}{\partial x^2} + \mu \frac{\partial^2 v}{\partial x \partial y} \right) + \frac{E}{2(1+\mu)} \left(\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 v}{\partial x \partial y} \right) = 0, \quad (8)$$

$$\mathcal{N}_2(u,v) = \frac{E}{1-\mu^2} \left(\frac{\partial^2 v}{\partial y^2} + \mu \frac{\partial^2 u}{\partial x \partial y} \right) + \frac{E}{2(1+\mu)} \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 u}{\partial x \partial y} \right) = 0. \quad (9)$$

Die zwei Anteile \mathcal{N}_1 und \mathcal{N}_2 gehen separat quadriert in die Verlustfunktion ein. Alle Spannungsgrößen lassen sich mithilfe des Materialgesetzes aus den Ableitungen der Verschiebungen berechnen. Es wurde untersucht, wie sich die Gewichtung der Verlustterme, die Dichte der Kollokationspunkte und die Netzwerkstruktur auf die Lösungsgenauigkeit des PINNs auswirken. Zur Bewertung dient die Wurzel des mittleren quadratischen Fehlers (RMSE) gegenüber der FEM Lösung aus Kap. 2.2.1 für σ_y .

Die Bedeutung der geeigneten Gewichtung in der Verlustfunktion wurde bereits in Kap. 2.1 demonstriert. Für die Scheibe ergibt sich ein Optimum bei einem Faktor für den Verlustterm der Randbedingungen von $\alpha = 3.0$. Für zu kleine Werte konnte der un stetige Lastverlauf nicht abgebildet werden. Für zu große Werte wird die Lösung innerhalb der Scheibe zunehmend unphysikalisch.

Die Anordnung der Kollokationspunkte erfolgte in beiden Raumrichtungen mit einer konstanten Anzahl an Punkten pro Meter (ppm). Die Punktdichte wirkte sich auf die darstellbaren Details aus. Höhere Punktdichten führen allgemein zu geringeren Fehlern und einer stabileren Konvergenz der Optimierung. Zuverlässige Ergebnisse waren ab einer Dichte von etwa 7 ppm zu erkennen.

Auch die Anzahl der Neuronen entscheidet darüber, welches Maß an Detail das NN darstellen kann. In [8] wird für ein PINN eine Mindestanzahl von drei versteckten Layern mit jeweils

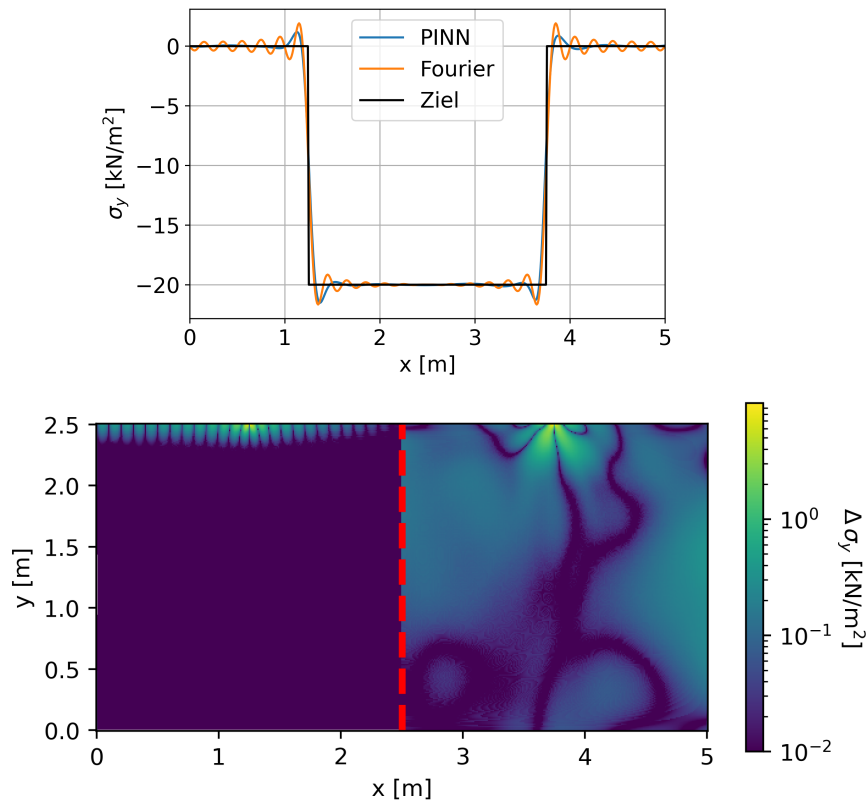


Abbildung 4: Vergleich der Lösungen für σ_y der gelagerten Scheibe berechnet mit analytischem Fourierreihenansatz aus 50 Reihentermen und PINN mit zwei versteckten Layern à 40 Neuronen.

Oben: Absolutwerte am oberen Scheibenrand zusammen mit angesetzter Last.

Unten: RMSE zur FEM Referenz. Auf der linken Hälfte die Fourier- auf der rechten Hälfte die PINN Lösung.

128 Neuronen empfohlen. Die Ergebnisse dieser Parameterstudie ergeben jedoch keine deutlichen Verbesserungen für Netzwerke mit mehr als zwei versteckten Layern à 20 Neuronen.

2.2.3 Methodenvergleich

Für σ_y werden die Ausgabe des PINNs und eine Fourierlösung mit 50 Reihentermen gegenübergestellt. Das PINN mit dem niedrigste RMSE ergibt sich für eine Verlustwichtung von $\alpha = 3$, einer Punktdichte von 7 ppm und ein Netzwerk mit zwei versteckten Layern à 40 Neuronen. Das obere Diagramm in Abb. 4 vergleicht die beiden Lösungen von $\sigma_y(x, y = 2.5)$ mit der angesetzten Randlast. Die Anzahl der Fourierterme wurde so gewählt, dass die halbe Wellenlänge des letzten Terms einem Gitterabstand von 10 ppm entspricht. Für beide Methoden zeigt sich nahe der Unstetigkeit eine Oszillation um den realen Verlauf. Diese klingt in

der PINN Lösung schnell nach außen hin ab, während sie sich in der Fourierlösung über die gesamte Breite deutlich fortsetzt.

In der unteren Darstellung von Abb. 5 wird ersichtlich, dass die PINN Lösung in der restlichen Scheibe deutlich größere Abweichungen aufweist. Der RMSE beläuft sich auf $1 \cdot 10^{-1}$ und ist am größten im Bereich der Unstetigkeitsstelle der Last sowie teilweise am linken und rechten Rand. Die Fourierlösung liefert eine deutlich geringere RMS Abweichung von $1.3 \cdot 10^{-2}$, die sich auf den oberen Rand der Scheibe konzentriert und auf einer Höhe von etwa 20 cm stark abnimmt.

3 Parameteridentifikation

3.1 Inverse Probleme

Die bisherigen Beispiele als auch ein Großteil der Veröffentlichungen zu PINNs befassen sich mit der Lösung von Vorwärtsproblemen, d.h. zu gegebenen Randbedingungen ist die resultierende Strukturantwort gesucht. Bei inversen Problemen wird hingegen von der Systemreaktionen auf unbekannte Systemeigenschaften wie etwa Materialparameter geschlossen.

PINNs sind ein flexibles Werkzeug zur Lösung inverser Probleme. Daten über bekannte Systemreaktionen werden in einem zusätzlichen Verlustterm $\mathcal{L}_{\text{Daten}}$ berücksichtigt. Unbekannte Parameterfelder können der Ausgabe des NNs hinzugefügt werden und geben nach erfolgreichem Training Lösungen, die mit den gegebenen Randbedingungen und Systemreaktionen verträglich sind.

In [9] wurde bereits der E-Modul und die Poissonzahl eines homogenen Stabes und einer Platte aus Verformungsdaten bestimmt. Im Folgenden wird eine PINN hingegen auf eine inhomogene Scheibe angewendet, um eine simulierte Schädigung zu identifizieren.

3.2 Scheibe mit Riss

Durch eine lokale Absenkung der Steifigkeit wird eine Schädigung in die belastete Scheibe aus Kap. 2.2.1 eingebaut. Zur Schädigungsidentifikation soll aus dem Verformungsbild der nun ortsabhängige Elastizitätsmodul $E(x,y)$ abgeleitet werden. Die Gleichgewichtsbedingungen für die Verlustfunktion müssen um zusätzliche Terme erweitert werden

$$\mathcal{N}_1(u,v,E) = \mathcal{N}_1(u,v) + \frac{\partial E}{\partial x} \left[\frac{\partial u / \partial x + \mu \partial v / \partial y}{1 - \mu^2} \right] + \frac{\partial E}{\partial y} \left[\frac{\partial u / \partial y + \partial v / \partial x}{2(1 + \mu)} \right] \quad (10)$$

$$\mathcal{N}_2(u,v,E) = \mathcal{N}_2(u,v) + \frac{\partial E}{\partial y} \left[\frac{\partial v / \partial y + \mu \partial u / \partial x}{1 - \mu^2} \right] + \frac{\partial E}{\partial x} \left[\frac{\partial u / \partial y + \partial v / \partial x}{2(1 + \mu)} \right]. \quad (11)$$

Bei identischen Lagerbedingungen wird statt einer Rechtecklast eine konstante Flächenlast von 20 kN/m^2 an der Oberkante aufgebracht. In der ursprünglich homogenen Scheibe mit einem E-Modul von 1 N/m^2 wird in einem in der Mitte der Unterkante anliegenden 4 cm breiten und 75 cm hohen Streifen der E-Modul auf 10 % des Ursprungswertes gesetzt, um

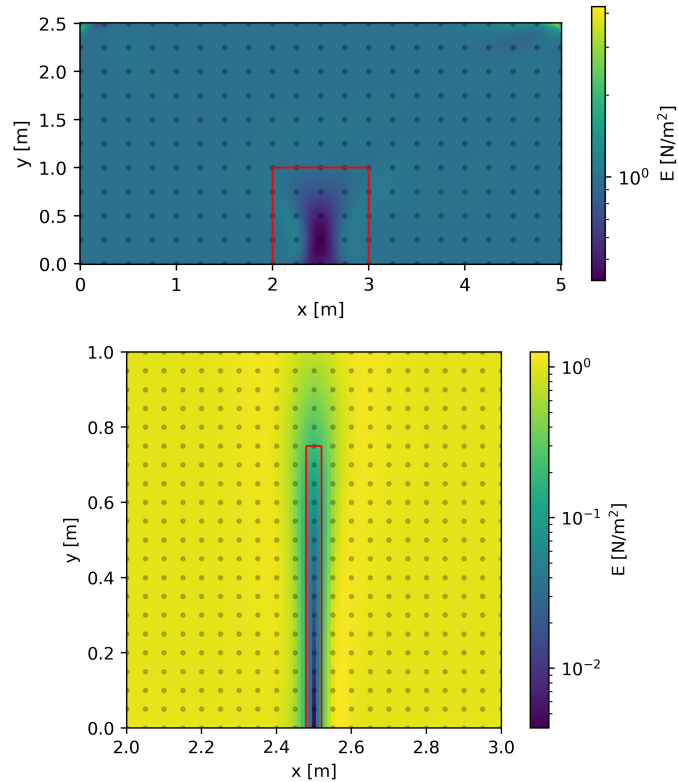


Abbildung 5: Mittels PINN identifizierter E-Modul der geschädigten Scheibe.

Oben: Virtuelle Messdaten wurden mit einer Dichte von 4 ppm eingelesen. Im umrahmten Bereich wurde eine Schädigung festgestellt.

Unten: Ergebnis nach Erhöhung der Messpunktdichte auf 20 ppm im Schädigungsbereich. Der rote Rahmen stellt den wahren Ort der eingebrachten Schädigung dar. Zwischen den Halbwertspunkten des Abfalls in x-Richtung liegen etwa 10 cm. Der mittlere E-Modul innerhalb des Rissbereichs beträgt 9.4 % der ungeschädigten Scheibe.

einen rissähnlichen Schaden zu modellieren. Das resultierende Verschiebungsfeld der Scheibe wurde mit einem FE-Modell berechnet und dient zur Generierung virtueller Messwerte, die zur Berechnung des Verlustterms $\mathcal{L}_{\text{Daten}}$ herangezogen wurden. Die Messdaten wurden zunächst auf einem Gitter konstanter Auflösung eingelesen. In der oberen Grafik von Abb. 5 ist der identifizierte E-Modul der gesamten Scheibe für eine Gitterdichte von 4 ppm dargestellt.

Der Ort des Schadens konnte zunächst auf den rot umrahmten Bereich eingegrenzt werden, in welchem nun die Dichte der Messpunkte auf 20 ppm erhöht wurde. Die neuen Messwerte wurden $\mathcal{L}_{\text{Daten}}$ hinzugefügt. Der damit identifizierte E-Modul ist in der unteren Grafik von Abb. 5 im Bereich des Risses dargestellt. Zum Vergleich ist in rot der Rand des wahren Schadens markiert. Werden die Halbwertsbreiten der E-Modul-Absenkung als Maß für die Rissbreite herangezogen, hat sich diese durch Verdichtung der Messwertpunkte von etwa

40 cm auf 10 cm verringert. Obwohl die Punktdichte in y -Richtung erhöht wurde, ist in beiden E-Modul-Identifikationen das obere Ende des Risses unscharf. Der Grad der Schädigung weicht punktweise z.T. stark ab, der gebildete Mittelwert über den Rissbereich kommt jedoch mit einem Abfall von 9.4 % den angesetzten 10 % bereits sehr nahe.

4 Fazit

Es wurde demonstriert, dass PINNs für einfache ein- und zweidimensionale Problemstellungen in der Baustatik verlässliche Resultate erbringen. Für eine Scheibe mit un stetiger Randlast weist die PINN Lösung im Vergleich zu einer analytischen Fourierlösung größere mittlere Fehler auf. Allerdings lassen sich PINNs ohne Mehraufwand auf komplexere Geometrien übertragen und bieten dort eine robuste Alternative zu fehlenden analytischen Lösungen. Außerdem kann die Genauigkeit der Lösungen weiter verbessert werden etwa durch eine automatisierte Anpassung der Kollokationspunktverteilung, als auch der Verlusttermgewichtung zwischen den Trainingsschritten. Weiterhin konnte erfolgreich eine rissartige Schädigung in der Scheibe identifiziert werden, ohne dass deren Ort dem PINN vorher bekannt war. Hierzu musste schrittweise die Dichte der eingepflegten Messdaten im Schädigungsbereich verfeinert werden. Diese Fähigkeit bietet großes Potenzial für die Anwendung im Bereich des Monitorings.

Literatur

- [1] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. *Journal of Computational Physics*, 2017.
- [2] E. Hake and K. Meskouris. *Statik der Flächentragwerke: Einführung mit vielen durchgerechneten Beispielen*. Springer Berlin Heidelberg, 2001.
- [3] SOFiSTiK AG. *SOFiSTiK FEA 2022*. [Software] <https://www.sofistik.de/>.
- [4] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. 2022.
- [5] D. F. Shanno. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, 1970.
- [6] M. Abadi, P. Barham, J. Chen, et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from [tensorflow.org](https://www.tensorflow.org).
- [7] F. Chollet. Keras. <https://github.com/fchollet/keras>, 2015.
- [8] S. Wang et al. An expert’s guide to training physics-informed neural networks. 2023.
- [9] D. Anton and H. Wessels. Physics-informed neural networks for material model calibration from full-field displacement data. 2022.