Bridge Maintenance, Safety, Management, Digitalization and Sustainability – Jensen, Frangopol & Schmidt (eds)

Data-derived train classification and monitoring of masonry bridge

J. Cheng, A. Chen & M.J. DeJong
University of California, Berkeley, Berkeley, USA

S. Cocking
University of Cambridge, Cambridge, UK

ABSTRACT: CFM-5 bridge in Barkston Ash, UK presents an asset monitoring and maintenance challenge. Since its construction in 1868, the masonry arch bridge has suffered damage and undergone a variety of repairs. The uncertainty of the repairs' effectiveness motivates the need for automated monitoring. As such, Fiber Bragg Grating fiber optic sensors have been installed along the underside of the masonry arch and dynamic strain data have been collected during train passages over the last 2 years. Each FBG measures average strain over its gauge length, and 2 lines of 10 grating rosettes were installed along the arch soffit. These provide distributed strain data over the extent of the arch, yielding more global information compared to typical point sensors. In this context, this paper presents a framework for post-processing and interpreting the monitoring data from bridge CFM-5. The objectives are to investigate long-term and future structural performance. To achieve this, data must first be normalized against train type; using features extracted from the dynamic response to a train passing overhead, the train type is classified via XGBoost, a framework for gradient-boosted decision trees. Then, the dynamic responses are analyzed to evaluate the sensitivity to different train types by observing peak dynamic strains and their correlation with temperature.

1 INTRODUCTION

Maintenance and management of infrastructure assets continues to be a challenge for asset managers. Masonry arch bridges play a significant role in the infrastructure of transportation networks across Europe, making up roughly 40-50% of rail and road bridge spans in the UK (Augusthus-Nelson et al. 2016). The number of masonry railway and road bridges across Europe is estimated to be around 300,000 (Brencich & Morbiducci 2007).

These structures are aging, with many experiencing 100 or more years of continuous operation. Inevitably, interventions are needed to repair damage suffered during operation. However, these repairs can unfavorably alter the stiffness profile of a bridge, resulting in further damage.

The complexities associated with managing masonry arch bridges motivate the need for comprehensive monitoring solutions. Automated data capture and analysis can significantly reduce the time, effort, and cost required to monitor and assess such bridges. Many sensing technologies have been applied to bridge monitoring. Deformations and changes to bridge geometry can be detected by LiDAR (Ye et al. 2018). Computer vision algorithms can be used to automatically detect cracks in images of masonry structures (Chaiyasarn et al. 2018). Dynamic deformations can be monitored with digital image correlation (DIC) techniques (Acikgoz et al. 2018).

Fiber Bragg grating (FBG) fiber optic sensors offer dynamic strain sensing for long term monitoring of structural degradation and have been installed at multiple masonry railway bridges throughout the UK (Alexakis et al. 2018). From these FBG deployments, preliminary analysis of changes in dynamic strain have been conducted, and monitoring of force flow through the masonry rail bridges has been carried out. In addition, statistical shape analysis (SSA) has been applied to FBG data to track changes in dynamic strain profiles over time (Alexakis et al. 2018).

DOI: 10.1201/9781003483755-399

These studies either only analyze behavior with regard to a single train type or rely on operator-provided train classification. In addition, some train classes are not easily distinguishable from one another through manual interpretation of the strain record. Thus, an automated train classification workflow is necessary to streamline the monitoring process when labelled events are not available to evaluate the effect of different train classes on bridge degradation.

Following recent research regarding the use of FBG sensing on masonry arch bridges, this paper examines the use of gradient-boosted decision trees for automated train classification. Long-term analysis of bridge responses against train class and temperature are also presented in this paper.

1.1 *Bridge CFM-5*

The bridge evaluated in this paper is the Barkston Road Bridge, designated as CFM-5 by the asset owner, Network Rail. Built in 1868, bridge CFM-5 is a single arch masonry bridge, carrying a variety of passenger and freight rail traffic over a local road at a skew angle of 26°. CFM-5 supports two trackways, a northern track for eastbound trains and a southern track for westbound trains.

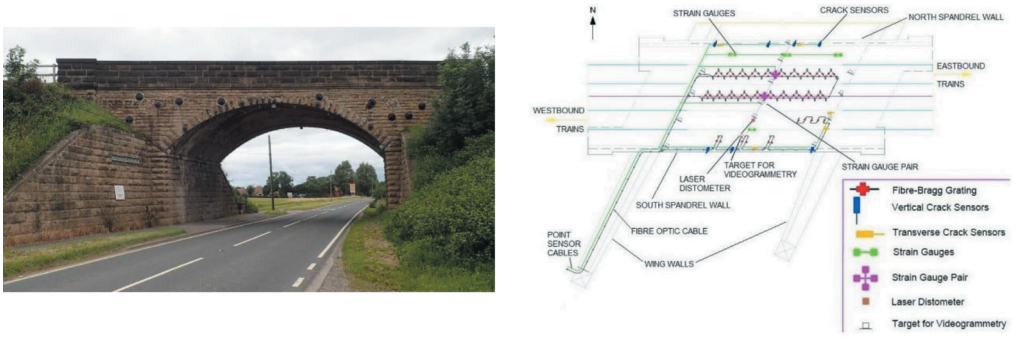

Figure 1. Elevation view of the CFM-5 bridge and its sensor installation plan.

The bridge has suffered various damage throughout its operation. The main signs of damage include separation cracks between the brick arch barrel and the stone spandrel walls and a longitudinal crack in the south-east of the arch barrel. To remedy these damages, various interventions have been implemented at bridge CFM-5, including transverse tie rods, soil anchors, and grout crack stitching. However, the effectiveness of these interventions is uncertain, motivating the need for remote monitoring technologies at bridge CFM-5.

1.2 *Monitoring installation*

Multiple sensing technologies were installed at bridge CFM-5, including FBGs, videogrammetry targets, strain gauges, and laser distometers. This paper focuses on the data obtained from the FBG sensors and ambient temperature sensors.

The FBGs are arranged in two configurations: in pairs to measure in-plane and out-of-plane strains across cracks, or in one of 20 triangular strain rosettes. One line of 10 FBG strain rosettes is installed in the arch soffit beneath the centerline of the northern track, and another line of 10 rosettes is located under the centerline of the entire bridge.

Because trains always travel in opposite directions on opposite tracks, rotational symmetry of the loading and thus response of the bridge has been assumed. Analysis of the strain readings support these assumptions so far (Cocking et al. 2019).

The installation is programmed to automatically record and upload strain readings for each train passing; data are logged during commute times to obtain records with relatively consistent passenger loading (Cocking et al. 2019). The Micron Optics FBG interrogator logs data at a sampling frequency of 250 Hz.

2 TRAIN CLASSIFICATION

2.1 *Traditional approach: Timetable dataset*

The reference dataset used for this classification problem is curated by scraping publicly available records of train departures and arrivals between August 2018 and January 2022, at stations adjacent to CFM-5, into a timetable. This timetable also includes both the direction of travel of the trains across the bridge and the train class. Then, the timestamp and direction of travel of each FBG record is compared with the timetable to match the record with a timetable entry.

Three direction checks are conducted to ascertain the direction of travel, for each FBG record. The first direction check compares the maximum peaks of two FBGs at the arch midspan, one under each track, as the FBG under the northern track will record a higher peak dynamic strain than the FBG under the bridge centerline when a train is passing over the northern track and vice versa. For the remaining two direction checks, the position of maximum dynamic peak of the FBG records on the eastern side of the bridge are compared to those on the western side of the bridge. That is, if a train is travelling eastward on the northern track, a peak is expected first on the western FBG before the eastern FBG. One of these checks is performed per FBG array. If all three direction checks agree, then the direction of the train is denoted accordingly.

Table 1. 2020-2021 dataset train types, carriage numbers, and total counts after cleaning.

Train Type	Number of Carriages	Count
Voyager 220	4	81
Voyager 220+220	8	131
Voyager 220+221	9	200
Voyager 221	5	165
Class 185	3	310
Class 185+185	6	834
Class 43	9	280
Northern 2 Car	2	511
Northern 3 Car	3	1244
Northern 4 Car	4	20
Nova 1	5	940
Nova 3	6	154
Total:		4870

Of about 7,100 FBG records collected between July 2020 and October 2021, about 5,300 are matched to the timetable in this way. Any records containing missing data are removed, as these correspond to instances when the analyzer was unable to automatically connect to the sensors upon booting. This leaves about 4,900 matched records.

In total there are 12 different train types, operated by 4 different train operating companies. A breakdown of the different train types observed in the 2020-2021 data is shown in Table 1.

2.2 *Proposed approach: Extracted features*

The approach proposed relies on the classification of trains based on extracted features from an FBG located near the midspan of the northern track. A truncated version of the timeseries data from this FBG are stored for analysis, as the original file contains buffer data at the start and end of the train event. In the unloaded bridge state, these buffer data mostly contain noise, whereas the portion of the record that contains relevant strain information as the train passes over is of more interest. This portion of the record that contains the train passing over the bridge is noted as the importance region.

The importance region is determined as follows. First, the mean of the last 1000 measurements of the record, called the tail, is calculated. The entire record is then vertically shifted to the mean of the tail, to remove any offset in absolute strain. Then, the standard deviation of the tail is calculated. Any portion of the record that exceeds 10 times the standard deviation of the tail is classified as the importance region. Records without an importance region are discarded, as these correspond to measurements that have been incorrectly triggered, resulting in a record with only noise. There are about 300 such records.

To focus on the shape of each signal and to reduce the variation in signal due to differences in train weight and speed, each record from this FBG is then normalized by its maximum recorded value. Finally, the signal is passed through a Savitzky-Golay filter, followed by a low-pass filter, to smooth out the measurements and remove any high frequency noise.

As seen in Table 1, some train classes have unique carriage numbers, resulting in signals with a unique number of peaks. However, simple peak counting is not sufficient for classification, as there remain other train types with equal carriage counts. Thus, features aim to characterize the peaks of the signal not just by counts, but also by peak heights and widths.

As shown in Table 2, Some of these features are calculated using a ring-up-ring-down scheme. That is, a threshold is set for the signal, and in this case, it is 0.05. The area where the signal surpasses this threshold before descending past it again denotes a ring up zone. Each ring up zone contains a peak, and the width of each ring up zone is also recorded.

Table 2. Features extracted from strain timeseries.

Feature	Description
len_ratio	Ratio of truncated length to untruncated length
peak_count	Sum of the numbers of peaks and valleys in timeseries
peak_spacing	peak_count divided by length of truncated timeseries
peak_freq	Highest-density frequency in PS of timeseries after filtering
avg_freq	Average frequency in PS after filtering
pass_count	Number of times signal exceeds threshold
height_avg	Average value of peaks within each ring up zone
height_std	Standard deviation of peaks within each ring up zone
widths_avg	Average length of the ring up zones
widths_std	Standard deviation of the widths of the ring up zones
width_max	Maximum width of the ring up zones

Lastly, other features are derived from the power spectrum (PS) of the signal, with the power spectrum being calculated via the SciPy Python library. Further details regarding these features are summarized in Table 2.

2.3 *Proposed classification model*

Gradient boosted decision trees are used for classification, namely XGBoost. Gradient boosting iteratively combines weak learners, into a single strong learner. In this case, the features are provided to the model, and through gradient boosting an ensemble of decision trees are combined to produce a stronger predictive model.

The objective function to be optimized at step t is defined as

$$obj(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t} w(f_i) \tag{1}$$

where $l\left(y_i, \hat{y}_i^{(t)}\right)$ is the multi-class softmax loss function and $w(f_i)$ is the complexity of the i-th tree. The complexity of a tree is defined by

$$w(f) = \gamma T + \frac{1}{2}\lambda \sum_{i=1}^{t} s_i \tag{2}$$

where s is the vector of weights assigned to each tree leaf and T is the number of leaves. The regularization parameters γ and λ are set to 0.2 and 0.01 respectively.

Taking the second order Taylor expansion of the loss function, the objective function for additive training can be re-expressed as

$$obj^{(t)}(\theta) = -\frac{1}{2}\sum\nolimits_{j=1}^{n} \frac{G_j^2}{H_j + \lambda} + \gamma T \tag{3}$$

where G_i and H_i are defined as

$$G_j = \sum\nolimits_{i \in I_j} \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{4}$$

$$H_j = \sum\nolimits_{i \in I_j} \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) \tag{5}$$

I_j is the set of indices of the input that are assigned to the j-th leaf (Chen & Guestrin 2016).

2.4 *Experimental method*

To replicate the performance of the model during online implementation, 4370 records between July 2020 and September 2nd, 2021 were used for training, and 500 records between September 2nd and October 19th 2021 were used for testing.

3 RESULTS

3.1 *Model performance*

Using the framework described in Section 2, a test accuracy of 96.8% is achieved. Class-wise precision, recall, and F_1 scores are displayed in Table 3. The F_1 score is defined as the harmonic mean of precision and recall and can be calculated as

$$F_1 = \frac{2}{recall^{-1} + precision^{-1}} \tag{6}$$

As seen in Table 3 and Figure 3, most train classes are classified with few or no mistakes, including rarer classes like the Northern 4 car. Additional performance may be achieved by fine-tuning filter parameters, model parameters, or extraction of further characteristic features.

Table 3. Test precision, recall, and F_1 score by class.

Train Type	Precision	Recall	F_1 Score
Voyager 220	0.667	1.00	0.800
Voyager 220+220	1.00	1.00	1.00
Voyager 220+221	0.962	1.00	0.980
Voyager 221	1.00	1.00	1.00
Class 185	0.903	0.800	0.848
Class 185+185	1.00	0.966	0.982
Class 43	1.00	0.964	0.982
Northern 2 Car	1.00	1.00	1.00
Northern 3 Car	0.925	0.976	0.950
Northern 4 Car	1.00	0.875	0.933
Nova 1	1.00	0.990	0.995
Nova 3	1.00	1.00	1.00

Figure 2 shows the gain, or average improvement in accuracy, of each feature after training. As expected, XGBoost uses features related to the number of peaks in the signal for much of the classification, as this relates to the number of carriages of each train.

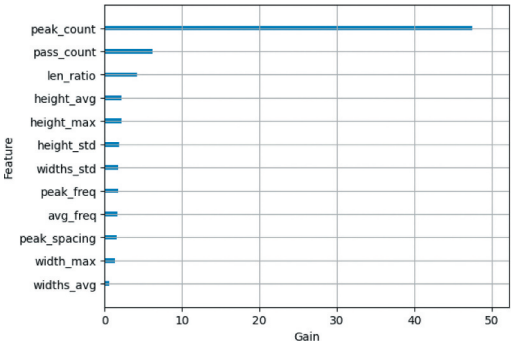

Figure 2. Feature importances.

This is supported by Figure 3, which displays information regarding the misclassification of trains. There are a significant proportion of Class 185 trains misclassified as Northern 3 car trains. This is likely due to actual and misclassified train types having the same number of carriages, as seen in Table 2. While peak counting and other features may be sufficient to distinguish between other train types, more complex classification models or features may be needed to fully separate signals of these similar train types.

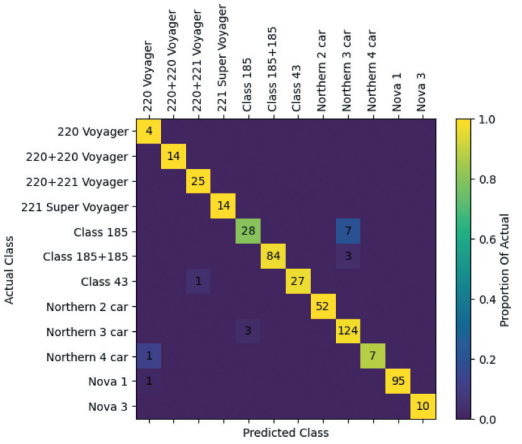

Figure 3. Confusion plot of XGBoost classifier on test set. Counts displayed in squares.

3.2 *Temperature-strain correlation*

With a robust classification framework, further analysis conditioned on train type can be conducted. A further 930 train events, which were collected from January 2023 through August 2023, have been labelled using the framework presented in Section 2. As seen in Figure 4, the train type compositions of both datasets are similar, and there is very little increase in the frequencies of relatively poorly classified train types (i.e., Class 185 and Northern 4 car trains).

Due to the relative abundance of Class 185+185 trains throughout the survey period, this class of train was selected for long-term temperature analysis. The following analysis and

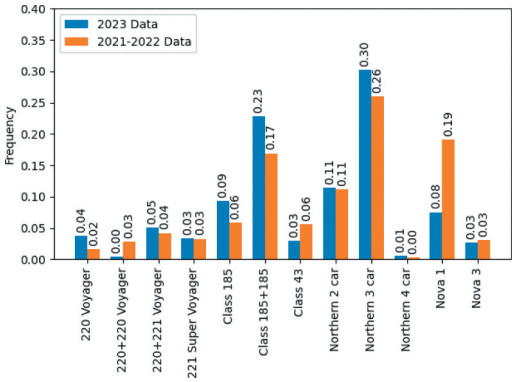

Figure 4. Frequency of predicted train classes.

results consider XGBoost labelled data from both the 2020-2021 and 2023 monitoring periods.

FBGs N_MID, C_MID, LCRACK_MID, and LCRACK_TIP were selected for observation. FBGs N_MID and C_MID are located near the middle of the span under the northern track and bridge centerline respectively. This northern portion of the bridge is relatively undamaged and exhibits the most linear behavior. LCRACK_MID and LCRACK_TIP span across a longitudinal crack in the southeast corner of the arch soffit, with LCRACK_MID positioned at the middle and LCRACK_TIP at the tip of the crack. If the crack progresses, it is expected that LCRACK_TIP, will show increased movement as the crack length grows. Opening displacements at LCRACK_MID may well also increase.

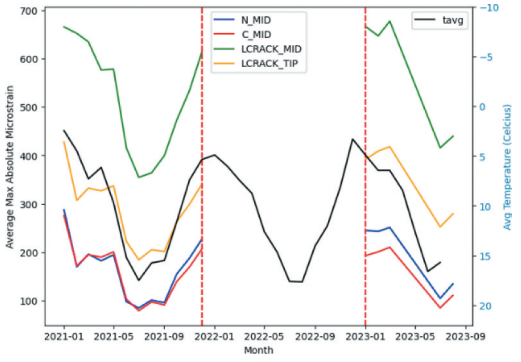

Figure 5. Monthly average maximum strain readings and monthly average temperature for Class 185 +185 trains travelling eastward.

It is known that maximum dynamic strain is well-correlated with temperature (Alexakis et al. 2021). Figure 5, which plots the average maximum dynamic strain reading across each month, for the four FBGs, along with monthly average temperature of a nearby weather station in Bradford, UK, clearly confirms this correlation. This value is calculated by averaging the monthly low and high temperatures.

To observe long-term differences more clearly, without the overlying effect of temperature, months with similar average temperatures were analyzed to compare the maximum dynamic strains of each Class 185+185 train event. Three spring months and three summer months with average temperatures around 6.4°C and 15.5°C respectively were selected for this purpose.

Figure 6 presents violin plots of the maximum dynamic strain measurements for Class 185 +185 trains. The thick dashed lines denote the distributions' median values, and the thin dashed

lines denote distributions' first and third quartile values. The median maximum dynamic strains for February 2023 and March 2023 are similar while the measurements for April 2021 are notably lower. This could be an indication of damage progression from 2021 to 2023, or possibly a sign of remaining temperature variation which the use of monthly average temperatures cannot remove. More complex normalization of the measurements against temperature will be explored in future work. Violin plots for summer months of similar monthly average temperatures yield similar results in the evolution of average maximum dynamic strain over time. Again, there is no clear ordering of the average maximum dynamic strains with respect to temperature. Generally, the average strains in 2023 are larger than those in 2021, but the extent to which the 0.2°C difference in average monthly temperature and operational damage contribute to this difference is unclear, and this will be explored further in future work, potentially by observing changes in correlation between temperature and maximum dynamic strain over time.

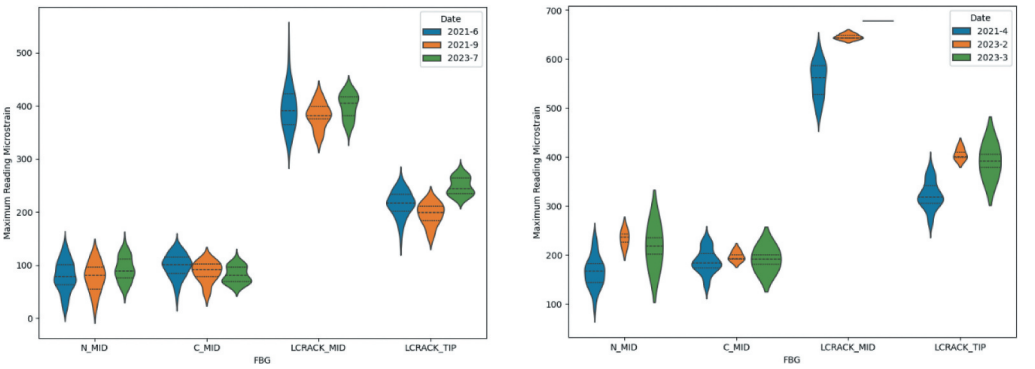

Figure 6. Maximum dynamic strain measurements for Class 185+185 trains, grouped by similar monthly average temperatures in (left) spring and (right) summer.

4 CONCLUSIONS

In this study, a comprehensive framework for post-processing and interpreting FBG monitoring data from the CFM-5 masonry arch bridge is presented. The primary focus is on automated train classification using gradient-boosted decision trees, specifically XGBoost, in order to study the bridge response normalized for different train types. The classification model exhibits a high test accuracy of 96.8%, demonstrating its efficacy across diverse train types, including rarer train classes. Further refinement of feature extraction could potentially enhance its performance.

Additionally, the study explores the correlation between temperature and strain, a critical factor in understanding long-term bridge performance. The observed correlation between maximum dynamic strain readings and temperature variations underscores the importance of considering environmental factors when assessing structural health.

The results highlight the significance of automated monitoring solutions, such as FBG sensors coupled with advanced classification models, for evaluating the health of aging infrastructure such as masonry arch bridges. Beyond streamlining the monitoring process, the proposed framework offers valuable insights into the structural response under varying operational and environmental conditions.

ACKNOWLEDGEMENTS

The work described here is part of a monitoring collaboration with engineering consultant AECOM, along with financial support from the asset owners Network Rail. The authors are grateful for the help and support of both organizations. The 1st author of this research was

financially supported by the U.S. National Science Foundation. Any findings, opinions, conclusions, and recommendations of this paper are those of the authors and do not necessarily reflect the views of the research sponsor.

REFERENCES

Acikgoz S., DeJong M. J., & Soga K. 2018. Sensing dynamic displacements in masonry rail bridges using 2D digital image correlation. *Struct Control Hlth* 25:e2187.

Acikgoz S., DeJong M. J., Kechavarzi C., Soga K. 2018. Dynamic response of a damaged masonry rail viaduct: measurement and interpretation. *Eng Struct* 168:544–558.

Alexakis, H., Lau, F. DH., & DeJong, M. J., 2021. Fibre optic sensing of ageing railway infrastructure enhanced with statistical shape analysis. *J Civil Struct Health Monit* 11:49–67.

Augusthus-Nelson, L., Swift, G., Smith, C., Gilbert, M., & Melbourne, C. 2016. Behaviour of backfilled masonry arch bridges subjected to cyclic loading. *Proceedings of ARCH'16, the 8th International Conference on Arch Bridges*: 1039–1048.

Brencich, A., & Morbiducci, R. 2007. Masonry Arches: Historical Rules and Modern Mechanics. *International Journal of Architectural Heritage* 1(2), 165–189.

Chaiyasarn K., Khan W., Luqman Ali, Sharma M., Brackenbury D., & DeJong M. 2018. Crack detection in masonry structures using convolutional neural networks and support vector machines. *Proceedings of the 35th ISARC*: 118–125.

Chen, T. & Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: 785–794.

Cocking, S. H., Ye C., & DeJong M. J. 2019. Damage Assessment of a Railway Bridge Using Fibre Optic Sensing and Lidar Data. *International Conference on Smart Infrastructure and Construction 2019*: 701–710.

Ye C., Acikgoz S., Pendrigh S., Riley E., & DeJong M. J. 2018. Mapping deformations and inferring movements of masonry arch bridges using point cloud data. *Eng Struct* 173:530–545.