



UNIVERSITY OF
CAMBRIDGE

**Near-real-time Country-wide
Estimation of Susceptibility and
Settlement Exposure from
Norwegian Mass Movements via
Inter-graph Representation Learning**

Joshua Dimasaka

A 4995-word report submitted in partial fulfilment
of the requirements for the degree of Master of Research

Abstract

Near-real-time Country-wide Estimation of Susceptibility and Settlement Exposure from Norwegian Mass Movements via Inter-graph Representation Learning

Joshua Dimasaka

Norway with its sensitive climate-physiographic characteristics faces a serious threat from mass movements such as landslides and avalanches, causing significant annual economic losses and risks to human lives and communities. The current national early warning system provides four to five categories of danger reports at a county or village level. However, its predictions are very sensitive to pre-defined categorization, leading to a poor perception of risk and costly local mitigation measures due to its over-estimated and unlocalized information. In this study, we developed a rapid inter-graph approach based on two graph-based machine-learning techniques that model the graphical representation of the hydrological and geological characteristics of 68,934 incidents of mass movements since 1957 and the connectivity information of 4,778 formal settlements and 257,000-km road networks, in both supervised and unsupervised ways, to produce a daily 1km-by-1km susceptibility map with a quantified assessment of intra- and inter-settlement exposure. Our findings achieved an aggregated performance of 86.25%, providing a distribution of improved susceptibility estimates and identifying settlements with high exposure levels using the 2020 Gjerdrum quick clay incident as a case study. With the increasing trend of rainfall due to the changing climate, our proposed inter-graph approach has opened an opportunity to estimate critical information for developing local adaptation and mitigation measures for national policymakers and regional county governors.

Contents

1	Introduction	6
2	Literature Review	7
3	Methodology	8
3.1	Data Preparation	8
3.1.1	Grid Discretization	8
3.1.2	Labels and Features	9
3.1.3	Ensemble and Dataset Split	10
3.1.4	Formal Settlements and Road Networks	10
3.2	Inter-graph Representation Learning	10
3.2.1	Supervised Ensemble Graph Neural Network	11
3.2.2	Unsupervised Spectral Graph Clustering	13
4	Results and Discussion	15
4.1	Mapping Susceptibility	15
4.1.1	Trained Ensemble Model	15
4.1.2	Efficient Multi-Step Hyperparameter Search	15
4.1.3	2020 Gjerdrum Mass Movement Incident	16
4.1.4	Comparison with Existing Warning System	18
4.2	Assessing Settlements	18
4.2.1	County-level Two-Step Graph Structuring	18
4.2.2	2020 Gjerdrum Mass Movement Incident	19
4.2.2.1	Minimum Triggering Exposure Probability of Mass-Movement-Susceptible Roads for Inter-Settlement Isolation	20
4.2.2.2	Intra-Settlement Exposure Probability of Being a Mass-Movement-Susceptible Area	20
5	Conclusion and Recommendation	22
6	Data and Code Availability	23
7	Declaration	24
	References	25

A Supplementary Information	30
A.1 Trained Ensemble Model	30
A.2 Efficient Multi-Step Hyperparameter Search	32
A.3 Exposure Probabilities on December 30, 2020	33

Chapter 1

Introduction

In the early hours of the 30th of December 2020, over 1,600 displaced Norwegian residents started their day with a national news headline on the widespread quick clay landslide at Ask village, Gjerdrum, resulting in 10 fatalities [1, 2]. This incident is one of the over 84,500 reported mass movements such as avalanches, rockslides, and slippage since 1900 [3], which is expected to occur more frequently due to the increasing trend of rainfall for the last 120 years and its sensitive climate-physiographic characteristics [4, 5]. As the villagers started their new year with loss and continued recovery efforts, many parliament members stressed the need to increase access to risk information [6].

Current state-of-the-art efforts of the Norwegian early warning system use a matrix-based approach to combine four classes of catchment susceptibility and daily rainfall intensity to determine warning levels (low, moderate, high, very high) [7, 8], which is very sensitive and dependent on the pre-defined categorization. The outputs are deterministic with only a single category assigned to the county or smaller villages, which does not show a disaggregated distribution of the warning levels along road networks or within the vicinity of settlements [9, 10], potentially leading to a poor perception of risk and costly local mitigation measures.

Recently, sophisticated models and high-resolution mapping have been introduced at the nexus of graph-based machine learning and remote sensing as they combine the nonlinearity and graphical information of various geospatial datasets (e.g., spatial, causal, or any relational connections). Therefore, in this study, we introduced a rapid inter-graph approach based on two graph-based machine-learning techniques that model the graphical representation of the hydrological and geological characteristics of 68,934 incidents of mass movements since 1957 and the connectivity information of 4,778 formal settlements and 257,000-km road networks, in both supervised and unsupervised ways, to produce a daily 1km-by-1km susceptibility map with a quantified assessment of intra- and inter-settlement exposure.

Chapter 2

Literature Review

Prior studies on susceptibility mapping, particularly in Norway, have underscored various modelling challenges such as the lack of spatial correlation [11] due to regional climatic differences [12, 13] and highly conservative estimates [14, 15] with the use of matrix-based approach [7]. Early efforts in applying statistical and machine-learning techniques [11, 16, 17, 18] have achieved satisfactory results but each with significantly different performance because of their respective inherent uncertainties. Among these several techniques, the ensemble approach, which could handle different model uncertainties, has gained the highest precision in combination with decision trees [19], neural networks [20], logistic regression [21], and logistic model trees [22]. However, these still have limited explainability as they disregarded spatial correlation and region-specific attributes.

Recent studies on integrating graphical information with machine learning (collectively called geometric deep learning [23] such as graph neural networks [24, 25]) have allowed several applications in disaster management [26], air pollutant estimation [27], transportation disruption detection [28], and early warning system [29, 30], as they capture the spatial connectivity and the attribute similarity of every data point. Despite these potential improvements, no studies have also been done to systematically integrate the insights from susceptibility maps with exposure patterns such as population, settlements, and roads, which also explains why the current warning system [9, 10] is limited to county-level or pre-defined large polygons that excessively aggregate the information on the map. Thus, in this study, we combined the *ensemble approach* and *graph neural networks* to map the susceptibility of Norway, and relate its graphical information with the exposure patterns that are expressed in a graphical (*but spectral*) way as well, to assess the degree of exposure of every populated settlement connected by roads.

Modelling the graphical interaction of the susceptibility map and exposure patterns improves the local community risk information and provides an opportunity to capture any region-specific similarities and differences. Additionally, the ensemble approach also provides a perspective on how model uncertainties could influence the predictions that traditional catastrophe models can benefit from. This inter-graph representation learning offers an alternative methodology to evaluate the widely-used disaster risk equation by explicitly modelling the graphical relationships not only between exposure and susceptibility, but also between hazard, exposure, and vulnerability, for a holistic assessment of regional adaptation and mitigation to the perils induced by the changing climate.

Chapter 3

Methodology

In this section, we present a detailed summary of the major procedures from data preparation to the implementation of our proposed inter-graph representation learning, which consists of two main steps: [3.2.1 Supervised Ensemble Graph Neural Network](#) for susceptibility mapping and [3.2.2 Unsupervised Spectral Graph Clustering](#) for settlement exposure assessment.

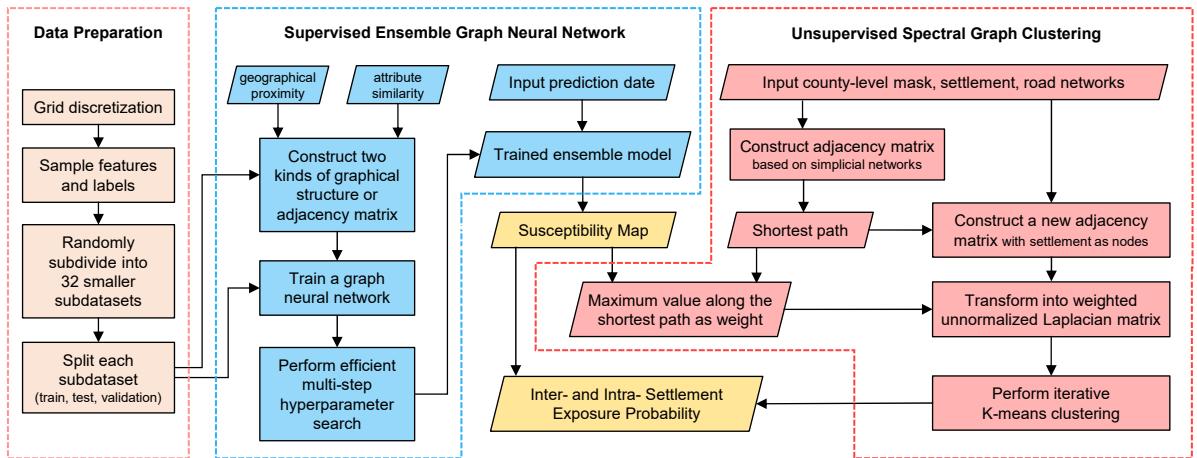


Figure 3.1: A schematic workflow diagram.

3.1 Data Preparation

3.1.1 Grid Discretization

To facilitate a consistent representation of features for the susceptibility mapping, we discretized all geospatial data using a 1km-by-1km grid, resulting in a map with a 1195x1550-array. For the settlement exposure assessment to realistically capture the detail of road networks, we changed the discretization from 1km to 50m (i.e., 20 times finer), resulting in a higher data storage requirement and longer processing time. As shown in Figure 3.2, the use of county-level masking (e.g., Nordland) has resolved this computational

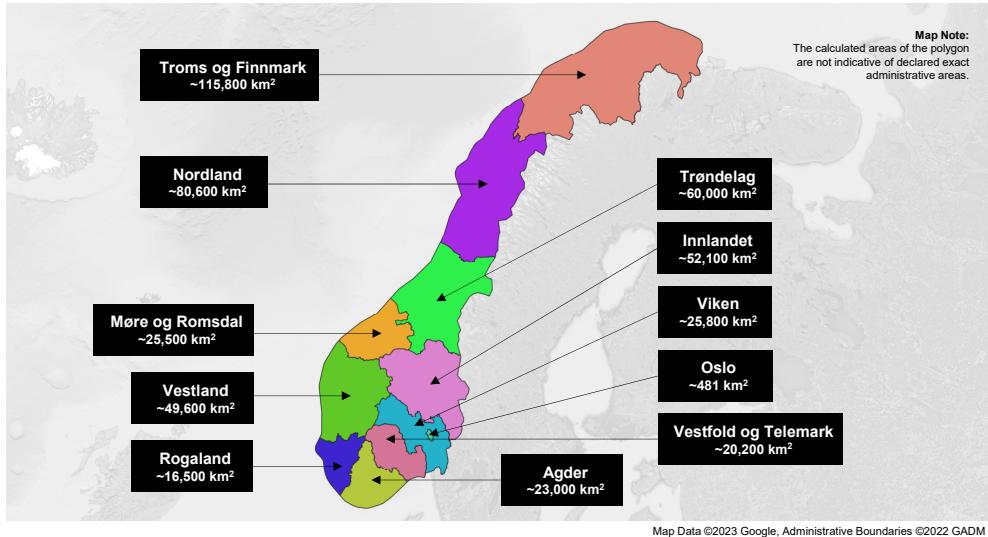


Figure 3.2: Norway counties and estimated areal extent.

issue by reducing the processing time to a few minutes (e.g., the county of Agder needed an 11-minute processing time) from several hours of a data-expensive analysis using the entire map of Norway with 11 counties. We also combined the counties of Oslo and Viken because they are geographically overlapping, reducing to a total of 10 groups of counties.

3.1.2 Labels and Features

We identified 68,934 incidents of mass movements (rockslide, landslide, avalanche, slippage, mudslide, flood landslide, and unspecified) reported since 1957 [31]. Their geographical and temporal information guided the extraction of relevant hydrological and geological characteristics. While these ground truth observations provide direct information on the sampling of mass movement incidents for our model training, it is important to carefully note that this inventory of incidents still has limitations such as incomplete and biased observations due to their remoteness and inadequate manpower to perform incident surveying work.

To relate these mass movements to generate country-wide susceptibility maps, we used two kinds of publicly available features: static and dynamic. Static features include steepness [32] and ancillary maps of the European Landslide Susceptibility Map (ELsus v2) [33, 34], which consists of shallow subsurface lithology class [35], land cover class [36], and slope angle class [37]. Date-specific dynamic features include total rainfall and mean temperature [38], and snow depth, snow water equivalent, and fresh snow water equivalent [39]. The size of these dynamic datasets has become a computational challenge because a year-specific single file exceeds 2GB, containing 365 daily values. To efficiently extract this information for each mass movement incident, we stored every single file into a hard drive with adequate space, requiring a total of 162GB, so that the data preparation becomes faster compared to using internet-enabled remote options like HTTP or OPeNDAP.

3.1.3 Ensemble and Dataset Split

For computational efficiency, we introduced an ensemble approach to subdivide the 508,182 grid points that collectively cover the entire map of Norway. As a result, the ensemble consists of 32 machine-learning models with randomly sampled points with a balanced representation between incident and non-incident. Even though we observed 68,934 incidents, these occupy only about 16,000 unique grid points, which is equivalent to 32 subdatasets, if each dataset has 500 points. Thus, each subdataset and its corresponding model used 500 incident and 500 non-incident labels that are randomly sampled with their respective features at a given date and location. To avoid model overfitting (which may lead to poor test performance), we split each subdataset using 70%, 15%, and 15%, for training, validating, and testing, respectively.

3.1.4 Formal Settlements and Road Networks

To understand how the susceptibility values may affect the exposure patterns (e.g., houses, people, and transportation routes), we used the latest available datasets of urban formal settlements (i.e., a minimum population of 200 and a maximum inter-house distance of 50m) [40] and road networks [41], respectively last updated in 2022 and 2020.

3.2 Inter-graph Representation Learning

We introduced an inter-graph approach that explicitly designs a one-way interaction between two graph-based machine-learning techniques wherein the node outputs of the first method ([3.2.1 Supervised Ensemble Graph Neural Network](#)) are used as an indicator of the strength of the edge connection between the nodes in the graphical structure of the second method ([3.2.2 Unsupervised Spectral Graph Clustering](#)). Figure [3.3](#) shows a schematic illustration of the interaction of different graphical representations or adjacency matrices for each step.

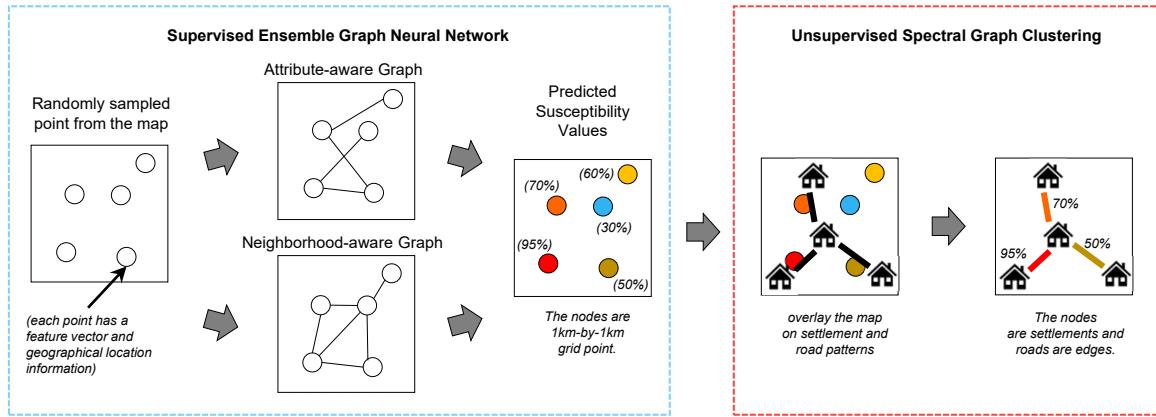


Figure 3.3: Interaction of graphical representations.

After generating a susceptibility map, we examined the exposure of formal settlements and their connections by road networks. Our results consist of two key insights: (1)

4.2.2.1 Minimum Triggering Exposure Probability of Mass-Movement-Susceptible Road for Inter-Settlement Isolation, which takes into account the network connectivity information between a particular settlement and all other settlements, and is also useful in understanding the level of exposure for succeeding evaluation of the risk of isolation; and (2) **4.2.2.2 Intra-Settlement Exposure Probability of Being a Mass-Movement-Susceptible Area**, which analyses the inherent susceptibility characteristics within the boundaries of any formal settlement.

3.2.1 Supervised Ensemble Graph Neural Network

As shown in Figure 3.4, there are only four key concepts - (1) supervised, (2) ensemble, (3) graph, and (4) neural network - to intuitively understand the implementation.

First, our records of mass movement incidents '*supervised*' the training of our machine-learning model to optimize its parameters such as weights and biases, so that, when the model tries to predict the probability, these incidents have high probability assigned.

Second, the 32 machine-learning models (as detailed in 3.1.3) formed as an '*ensemble*' to efficiently handle the very large dataset of over 500,000 points that cover the country. As each member or model of our ensemble predicts the susceptibility on the map, we democratically aggregate their estimates and present their differences or variation. In this way, any percentiles can also be used to specify a range of reasonable estimates of susceptibility because of epistemic (model) uncertainty.

Third, for the k th model of the ensemble, we created an undirected '*graph*' $G_k = (V_k, A_k, X_k)$, where V is the set of nodes represented by randomly sampled 1km-by-1km

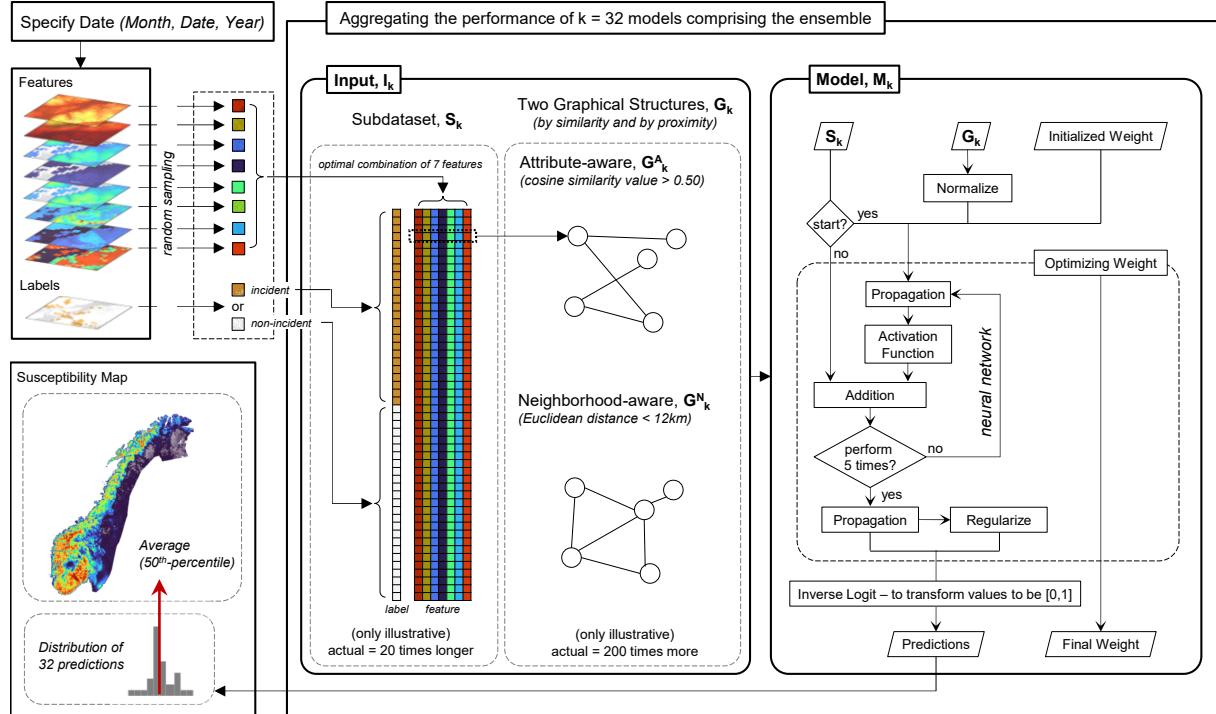


Figure 3.4: Schematic flowchart of the supervised ensemble graph neural network.

grid points, A is the adjacency matrix that contains the connectivity information between nodes, and X is the feature matrix of subdataset S_k represented by the optimal combination of 7 features, as a result of hyperparameter investigation.

Our combined A_k depends on the interaction of attribute-aware A_k^A and neighborhood-aware A_k^N . We tested their varying contribution by assigning weights [0, 1] in our hyperparameter investigation as well to determine their optimal interaction.

Fourth, the '*neural network*' $f(X_k, A_k)$ implements a layer-wise propagation (i.e., in every layer of our neural network, we propagate the update to our G_k by following the connectivity from A_k [24]). In symbols,

$$H_k^{l+1} = \sigma(\widehat{D_k}^{-1/2} \widehat{A_k} \widehat{D_k}^{-1/2} H_k^l W_k^l) + H_k^l \quad (3.1)$$

Here, $\widehat{A_k} = A_k + I_N$ is the adjacency matrix of our graph with added self-connections (I_N is the identity matrix). $\widehat{D_k}$ is the degree matrix of $\widehat{A_k}$. W_k^l is the weight matrix for the multiplication or propagation with learnable weights. H_k^l is the matrix of activations in the l^{th} layer; $H_k^{l=0} = X_k$. σ is the activation function $\text{ReLU}(\cdot) = \max(0, \cdot)$. The final operation does not apply an addition step, but uses the last result of the ReLU layer for propagation for the last layer L , which will be used as an input to our inverse logit equation (3.2) to return the output P_k with values [0, 1].

$$P_k = \frac{1}{1 + e^{-H_k^L}} \quad (3.2)$$

To avoid the poor effects of over-propagation, we applied regularization to our loss function by controlling the changes in the next propagated value [42].

$$\begin{aligned} Loss = & -\frac{1}{N} \sum_{n=1}^N (y^n \log(p^n) + (1 - y^n) \log(1 - p^n)) \\ & - \frac{1}{N} \sum_{n=1}^N (p^n \log(q^n) + (1 - p^n) \log(1 - q^n)) \end{aligned} \quad (3.3)$$

Here, the first term is the regular cross-entropy loss function that compares probabilistic outputs p^n with values [0, 1] to the binary ground truth data y^n with values of 0 if non-incident and 1 if incident. The second term regularizes the model by using:

$$q^n = \frac{1}{1 + e^{\widehat{D_k}^{-1/2} \widehat{A_k} \widehat{D_k}^{-1/2} H_k^{L+1}}} \quad (3.4)$$

We trained the model until the validation loss reaches a minimum. To analyze the predictive performance, we generated the receiver operating characteristic (ROC) curve to observe how true positive rates (TPR) change with false positive rates (FPR). A desirable area under the curve (AUC) should be very close to 1.0, wherein 0.5 indicates random guessing only while (0.7, 0.8] is acceptable, (0.8, 0.9] is excellent, and (0.9, 1.0] is outstanding [43]. The balance between TPR and FPR is determined using the misclassification cost equation below, which also gives the optimal operating threshold in [0, 1].

$$S = \frac{Cost(P/N) - Cost(N/N)}{Cost(N/P) - Cost(P/P)} \times \frac{N}{P} \quad (3.5)$$

where S is the slope of the line that intersects the ROC curve at the point of optimal operating threshold, $Cost(N/P)$ is the cost of misclassifying a positive class as a negative class, and $Cost(P/N)$ is the cost of misclassifying a negative class as a positive class. P and N are the number of estimates in the positive and negative class, respectively.

3.2.2 Unsupervised Spectral Graph Clustering

As shown in Figure 3.5, we used a county-level two-step graph structuring approach to model the connections between formal settlements and roads. The first graphical structure treats every pixel or grid point of formal settlement and road network on the map as nodes of a simplicial network, which means that a connection exists if two pixels or points are adjacent to each other by one step in all directions. This simplicial network generally describes the topology of roads and settlements, which has enabled us to extract other information such as the shortest path between any two pixels.

Using the length and profile of extracted shortest paths that could indicate a probable edge between any two settlements as nodes, we developed the second graphical structure wherein (1) a random point within the vicinity of a particular settlement represents the node and (2) the maximum susceptibility value along the shortest path between any two settlement serves as the relative weight of their connection.

We performed an iterative spectral graph clustering wherein we observed which settlement gets isolated when we increase the susceptibility cutoff value, from 0 to 1, which

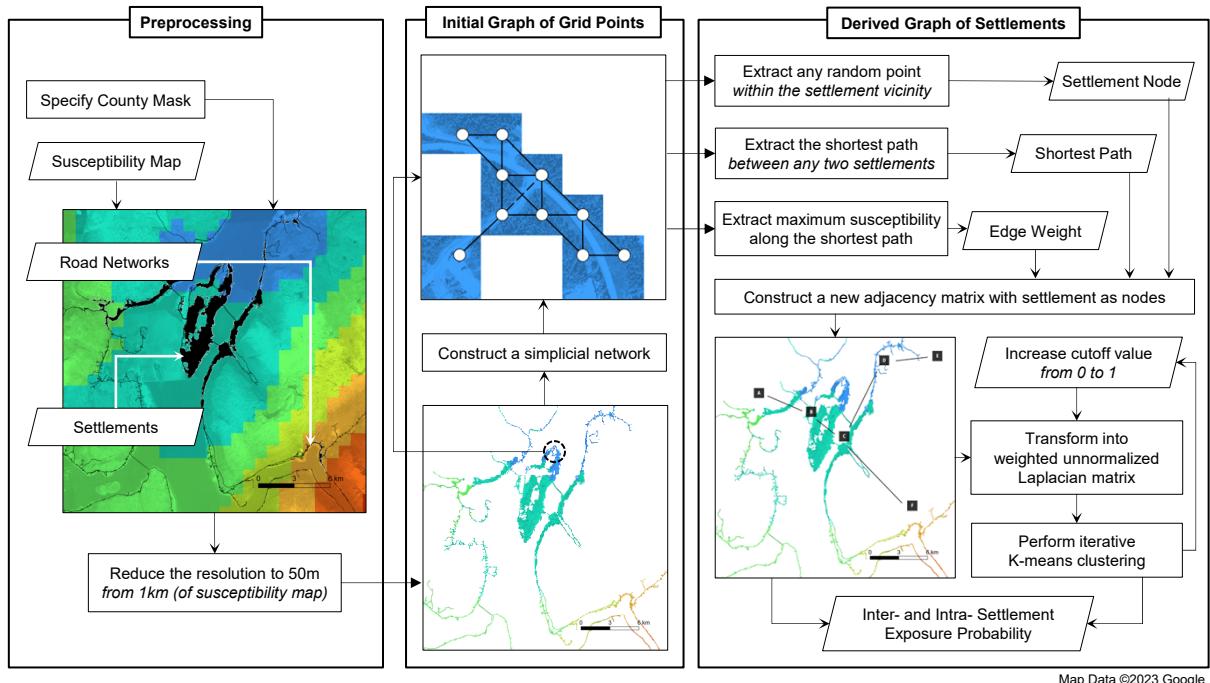


Figure 3.5: Schematic flowchart of the unsupervised spectral graph clustering.

removes the edge connection. For every iteration, we implemented K-means clustering using the spectrum (i.e., eigenvectors) of the unnormalized graph Laplacian transformation L of our derived graphical representation of settlements with adjacency matrix W [44]. In symbols,

$$L = D - W \quad (3.6)$$

Here, D is the degree matrix of W . We used the eigenvectors $\vec{v}_{\lambda=0}$ that correspond to zero eigenvalues (λ) because it indicates the number of connected sub-components of the graph.

$$L\vec{v} = \lambda\vec{v} \quad (3.7)$$

As a result, for all the iterated values tested, we then extracted the lowest cutoff value that would cause the isolation (in a graphical sense) of the settlement node. This value serves as the minimum triggering exposure probability of mass-movement-susceptible road for inter-settlement isolation because we assess the susceptibility value for each road segment or shortest path between any two settlements, which means that an edge or connection removal would potentially cause an inter-settlement isolation. We also presented the aggregated susceptibility value within the vicinity of each settlement to provide a holistic assessment of exposure. A detailed discussion of the assumptions and applications is presented in 4.2.2.1 and 4.2.2.2.

Chapter 4

Results and Discussion

4.1 Mapping Susceptibility

4.1.1 Trained Ensemble Model

Using our proposed efficient multi-step hyperparameter search, the trained ensemble of 32 machine-learning models optimally achieved an aggregate AUC of 86.25%, as shown on the ROC curve in Figure 4.1. As the curve comprises different pairs of points of false and true positive rates, minimizing the misclassification costs yields 0.48 as the corresponding optimal operating point, which becomes the new relative midpoint for the resulting susceptibility values on the map. A summary of the selected feature selection for each machine-learning model is presented in Appendix A.1.

4.1.2 Efficient Multi-Step Hyperparameter Search

We determined the parameters for all 32 models that compose the ensemble using six major hyperparameter search steps: (1) feature selection; (2) the interaction between neighborhood- and attribute-aware adjacency matrices; (3) the threshold for distance and cosine similarity; (4) model architecture about the depth and width of hidden layers of the neural network; (5) choice of propagation-related regularization; and (6) feature selection using a reduced number of possibilities and fixed configurations based on preceding steps. Because of the weight-related sensitivity of hyperparameter search, we implemented the first five steps ten times with different random initialization of weights of neural networks. This procedure has enabled us to efficiently perform the neural network training, instead of an expensive and exhaustive hyperparameter search with over 1.5 billion runs.

First, of 176 possibilities with 7-10 features, the optimal selection contains seven features that consistently include ELSUS susceptibility, steepness, slope angle class, rainfall, and temperature. To inform the succeeding steps, we defined a new subspace with these five consistent features and any two features that were randomly chosen from non-consistent.

Second, of 12 possibilities, the optimal interaction can be any of the seven options: 0 (or merely attribute-aware adjacency matrix), 0.10, 0.20, 0.30, 0.40, 0.50, and the AND operation (i.e., the connection between any two points should exist in both attribute- and neighborhood-aware adjacency matrices). This implies that the neighborhood-aware

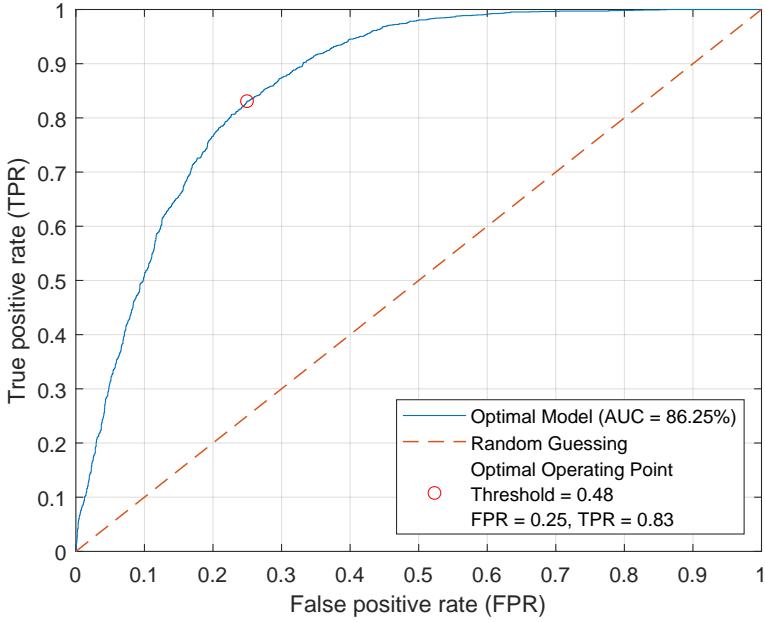


Figure 4.1: ROC curve of the trained ensemble model.

adjacency matrix has less significant information than that of the attribute-aware adjacency matrix, which may be due to the large space of locations for random sampling of data points that could be very far from each other.

Third, of 200 possibilities, the optimal threshold for the neighborhood-aware adjacency matrix is 12 km, which means that the edge or connection between any two data points exists if the geographical distance is at most 12 km. For the attribute-aware adjacency matrix, any of the cosine similarity values (e.g., 0.1, 0.2, ..., 0.9, 1.0) can be used, given the 12-km distance threshold. As these first three steps defined a new subspace of possibilities of parameters, a combined hyperparameter search showed that the optimal interaction has changed to the AND operation.

Fourth, of 28 possibilities, the optimal architecture selects the simplest and most frequent option (i.e., lowest) with a depth of six hidden layers and a width of 240 because the performance on the test dataset under varying depth and width was generally indifferent.

Fifth, of four possibilities, the optimal regularization is the cross-entropy loss equation, which yields slightly higher test performance compared to those of the Kullback–Leibler divergence, mean-square error, and without regularization.

Sixth, for each of the 32 models, we used the optimal configurations from previous steps to determine the respective combination of seven features using 20 different random initialization of weights of neural networks. The first three steps of this hyperparameter search are presented in Appendix A.2.

4.1.3 2020 Gjerdrum Mass Movement Incident

To demonstrate the supervised ensemble graph neural network, we derived the susceptibility map for December 30, 2020, when the Gjerdrum mass movement incident occurred (see

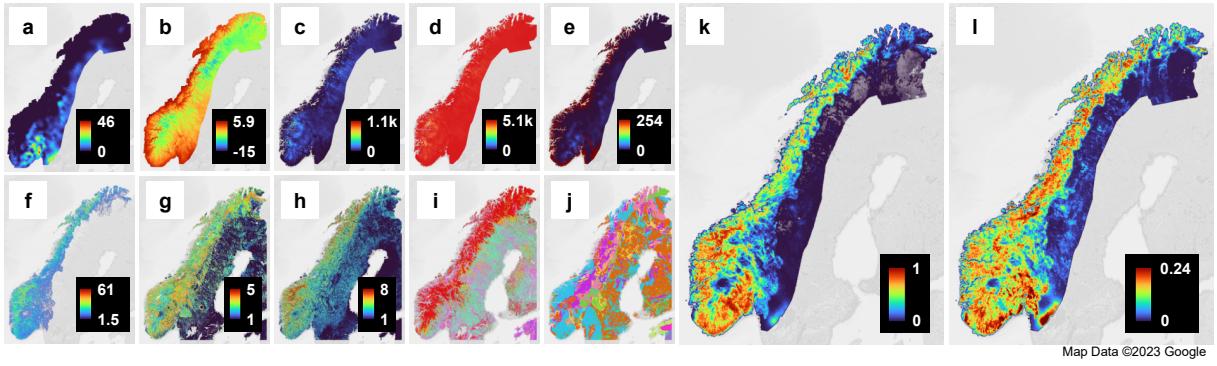


Figure 4.2: Date-specific maps of (a) total rainfall (mm/day), (b) mean temperature (Celsius degrees), (c) snow depth (cm/day), (d) snow water equivalent (mm/day), and (e) fresh snow water equivalent (mm/day) for December 30, 2020. Static maps of (f) steepness (degrees), (g) ELSUS susceptibility (categorical integer), (h) slope angle class (categorical integer), (i) land cover class (categorical integer), and (j) lithology class (categorical integer). Resulting (k) average (50th-percentile) susceptibility map with (l) calculated standard deviation using the estimates of 32 machine-learning models.

Figure 4.2). At 3:45-4:00 A.M. (GMT+1), a catastrophic quick clay landslide caused 10 fatalities, displaced over 1,600 individuals, and damaged 31 houses because of the combined triggering effect of erosion and moderate rainfall event preceding the failure of the ground [2]. Figure 4.2a also confirms the rainfall patterns reaching a daily total of 46mm in the southern region of Norway, where the incident occurred.

Moreover, our results in Figure 4.2k estimated that the incident vicinity has an average susceptibility of 58% with a standard deviation of 25%. Assuming normally distributed predictions from 32 models, this corresponds to an upper limit value of 83%, which is the 84th-percentile estimate (i.e., mean plus one standard deviation) that is commonly used in probabilistic decision-making and risk management [45]. As a result, this estimated band of 58%-83% for the incident vicinity could have significantly influenced the risk perception and tolerance of many affected residents [46, 47].

However, in informing our policymakers and decision-makers, it is important to note these numerical susceptibility values merely provide the location-specific probability of mass movement based on hydrological and geological characteristics such as rainfall, steepness, land cover, and other features considered in the training of our machine learning model. This should not be used interchangeably with hazard (which incorporates temporal information such as the frequency to derive the probability of a magnitude-specified mass movement), exposure (which considers the number of residents, houses, and road segments that are within the vicinity), physical vulnerability (which accounts for the system capacity such as the structural material of roads or buildings), and risk (e.g., probability of fatalities or damaged roads). To avoid misinterpretation of the resulting susceptibility maps, this also explains why our approach in deriving susceptibility maps does not include any features such as the occurrence rate of mass movements, population density, and the year of construction of road segments, which could respectively relate to the hazard, exposure, vulnerability, or risk due to mass movements.

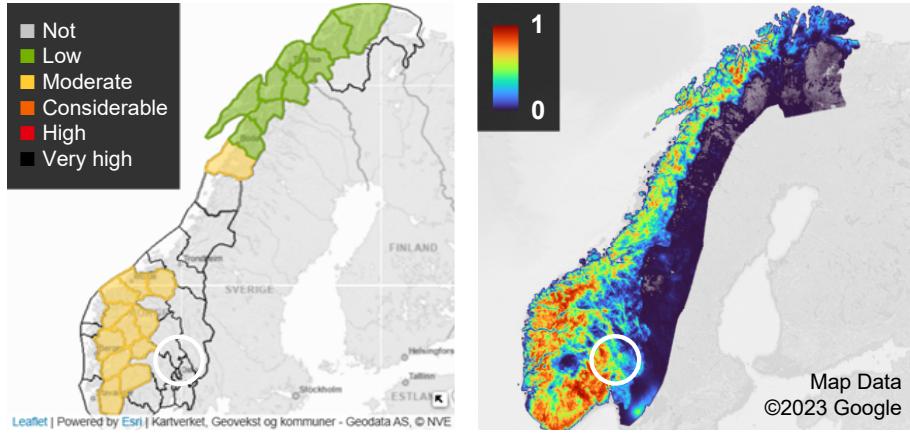


Figure 4.3: Comparison between the prediction of the Norwegian Avalanche Warning Service (*left*) and our estimates (*right*) for December 30, 2020.

4.1.4 Comparison with Existing Warning System

As shown in Figure 4.3, the current Norwegian early warning system reported a very low probability for landslides [9] and no significant danger for avalanches [10] for the municipality of Gjerdrum. In contrast, as previously presented, our estimates showed a higher prediction, which verified the occurrence of the said incident. This particular case validates the performance of our model, which both generalizes various types of mass movements because of their shared causal drivers and offers high-resolution maps compared to the county-level presentation of the current system.

4.2 Assessing Settlements

Inter- and intra-settlement analyses are equally important in understanding the total exposure characteristic of a settlement. For instance, even though a settlement (e.g., danger zones but well-connected to roads) may have a low minimum probability to trigger inter-settlement isolation, it may still have a high intra-settlement exposure probability because it may be situated next to an eroding soil mass with imminent danger within its vicinity. A simple case study is presented in Figure 4.4, wherein the relatively remote village of Leinesfjord has high inter-settlement exposure (69%) but low intra-settlement exposure (21%) because it is not situated in danger zones whereas the opposite case occurred for the village of Brattvåg, which is connected to many neighboring villages. This assessment of exposure probability for roads to trigger inter-settlement isolation uniquely presents a new perspective by capturing the effect of meaningful interaction of settlements at the county level.

4.2.1 County-level Two-Step Graph Structuring

Although our proposed *county-level two-step graph structuring* has allowed a large-scale assessment of settlement exposure, it is important to discuss its three major assumptions:

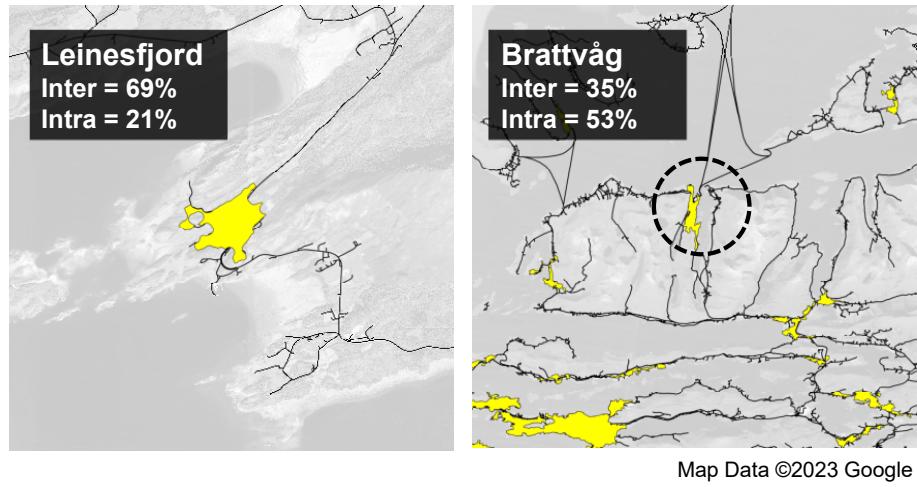


Figure 4.4: Comparison between the villages of Leinesfjord (*left*) and Brattvåg (*right*).

(1) the neglected inter-county connectivity, which may have provided additional information on how a settlement may access mitigation assistance from neighboring settlements from different counties instead of assuming strict and non-cooperative administrative county boundaries; (2) the use of multiple shortest paths, which may have corrected our underestimated probable connections between any two settlements; and (3) the statistical distribution of susceptibility values along the shortest path, which may have provided a more precise approach to represent the relative weight of the connection between any two settlements. Nonetheless, our proposed methodology is still a reasonable initial effort in conducting large-scale settlement exposure assessment that could aid the decision-making process of Norwegian county governors in charge of coordinating the mitigation efforts of various municipalities within their respective county-level jurisdictions.

4.2.2 2020 Gjerdrum Mass Movement Incident

In Figure 4.5, we revisit the quick clay incident. Figure 4.5f and Figure 4.5g showed slight differences between the two probabilities for the five illustrated settlements. This may be attributed to their close proximity to each other, which reduces the spatial variation of susceptibility values for each settlement. Interestingly, however, this implies that these two probability measures are more informative for formal settlements that are located far from each other (i.e., relatively remote areas from the cities) and less dense road networks where the susceptibility values take more variation along a longer path connecting two settlements. A summary of these two probabilities with population information of all settlements in the combined counties of Oslo and Viken, where the significant rainfall pattern was mostly observed, is presented in Appendix A.3.

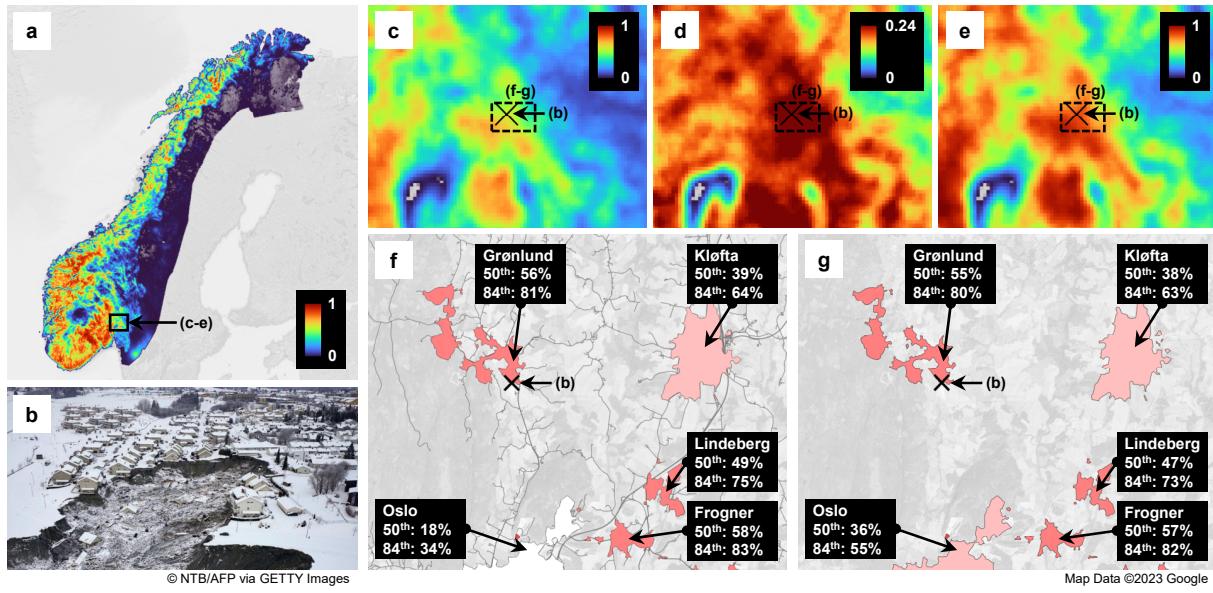


Figure 4.5: (a) Country-wide map of average (50th-percentile) susceptibility December 30, 2020; (b) a photo showing the extent of the quick clay incident; a set of magnified maps: (c) average, (d) standard deviation, and (e) upper limit or 84th-percentile; Five neighboring settlements around the incident, each showing (f) the minimum triggering exposure probability of mass-movement-susceptible roads for inter-settlement isolation and (g) the intra-settlement exposure probability of being a mass-movement-susceptible area.

4.2.2.1 Minimum Triggering Exposure Probability of Mass-Movement-Susceptible Roads for Inter-Settlement Isolation

Our findings showed that, on average, the formal settlement of Grønlund only needs a minimum triggering value of 56%. This means that, if all the roads connecting Grønlund to other settlements are very vulnerable with inadequate capacity to handle the effects of mass movements, Grønlund could result in isolation from the group of settlements. This also indicates that, if any of these roads have a susceptibility value of more than 56% (e.g., 70%), the same outcome can also be expected.

Similarly, among the five formal settlements, Oslo has the smallest estimate, which implies that it has a very low exposure to becoming isolated because of its wide and central areal extent that connects to many neighboring settlements, effectively increasing its total degree of connectivity. However, these estimates assumed a uniform vulnerability of all connecting roads whereas, some road segments may have varying characteristics such as a low vulnerability due to an applied engineering intervention, which would also affect the cutoff mechanism of our clustering.

4.2.2.2 Intra-Settlement Exposure Probability of Being a Mass-Movement-Susceptible Area

In Figure 4.5g, our findings showed that the formal settlements of Grønlund and Frogner achieved significantly high values of intra-settlement exposure probabilities above 50% for

both the average and upper limits. The other three settlements - Kløfta, Lindeberg, and Oslo - resulted in relatively lower values, which may be due to their relatively flat surface (i.e., low steepness) that decreases their susceptibility to being a mass-movement area. In this particular example, Grønlund was also found to be three times more susceptible than Oslo, on average.

Furthermore, we also observed that the upper limits of susceptibility increased due to the high standard deviation of our ensemble predictions. This underscores the benefits of using ensemble learning in which a distributional characteristic of predictions can be used to model the uncertainty that is related to the differing predictions of our 32 trained machine-learning models. Furthermore, our ensemble learning has enabled us to provide a more informative range of predicted estimates (e.g., average and upper limit), instead of a single value that could be misleading for risk perception because of various sources of model uncertainty.

Chapter 5

Conclusion and Recommendation

Our proposed inter-graph representation learning has opened an opportunity to develop solutions for climate change adaptation and mitigation by understanding the graphical relationships of different geospatial datasets of susceptibility and exposure patterns. It also creates a robust, reusable, and flexible multi-modal methodology that accounts for the similarities and differences based on proximity and attributes, which can be extended for any other relationships such as causal and time-series.

Our study has specifically developed a proof of concept that is computationally efficient for the future advancement of the automated Norwegian early warning system from mass movements, which could effectively identify the high-risk settlements as pertinent and urgent information for policymakers and regional county governors. Nevertheless, we recommend investigating the usefulness of (1) real-time satellite-derived vegetation data, (2) hypergraph modelling as an alternative for settlement-road connectivity, and (3) other geometric deep learning methodologies.

Chapter 6

Data and Code Availability

The complete set of data and code are available in our publicly available Zenodo repository (<https://github.com/jtdimasaka/InterGraphNorwayMM>). The code and its documentation are available in our GitHub repository (<https://github.com/jtdimasaka/InterGraphNorwayMM>).

Chapter 7

Declaration

This report is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text and/or bibliography.

References

- [1] Ivanna Penna and Inger-Lise Solberg. Landscape changes and bedrock reconstruction in gjerdrum area. methodological approach and main results. Technical Report 394800, Norges Geologiske Undersøkelse, Fence Room, Creek, 2021.
- [2] Jean-Sebastien L'Heureux, Bjørn Kristian Bache, Solve Hov, and Suzanne Lacasse. Mitigation measures following landslides in sensitive clays, examples from Norway. In *Geohazards VIII*, December 2022.
- [3] Christian Jaedicke, Karstein Lied, and Kalle Kronholm. Integrated database for rapid mass movements in norway. *Natural Hazards and Earth System Sciences*, 9(2):469–479, 2009.
- [4] Katja Laute and Achim A. Beylich. Potential effects of climate change on future snow avalanche activity in western norway deduced from meteorological data. *Geografiska Annaler: Series A, Physical Geography*, 100(2):163–184, 2018.
- [5] I Hanssen-Bauer, H Drange, EJ Førland, LA Roald, KY Børshheim, H Hisdal, D Lawrence, A Nesje, S Sandven, A Sorteberg, et al. Climate in norway 2100. *Background information to NOU Climate adaptation (In Norwegian: Klima i Norge 2100. Bakgrunnsmateriale til NOU Klimatilpassing)*, Oslo: Norsk klimasenter, 2009.
- [6] Stortinget. Written question from Une Bastholm (MDG) to the Minister of Petroleum and Energy. https://www.stortinget.no/no/Saker-og-publikasjoner/Spo_rsmal/Skriftlige-sporsmal-og-svar/Skriftlig-sporsmal/?qid=82767, 2021. Accessed: 2023-06-20.
- [7] Graziella Devoli, Rainer Bell, and José Cepeda. Susceptibility map at catchment level, to be used in landslide forecasting, norway. 2019.
- [8] Emma Barfod, Karsten Müller, Tuomo Saloranta, Jess Andersen, Nils Kristian Orthe, Anders Wartianen, Tore Humstad, Steinar Myrabø, and Rune Engeset. The expert tool xgeo and its applications in the norwegian avalanche forecasting service. In *International Snow Science Workshop Grenoble*, pages 7–11, 2013.
- [9] Norwegian Water Resources and Energy Directorate. Flood and landslide warning service. <https://www.varsom.no/en/flood-and-landslide-warning-service/>, n.d. Accessed: 2023-06-20.

- [10] Norwegian Water Resources and Energy Directorate. Avalanche warnings. <https://www.varsom.no/en/avalanches/avalanche-warnings/>, n.d. Accessed: 2023-06-20.
- [11] Arzu Erener and H Sebnem B Düzgün. Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of more and romsdal (norway). *Landslides*, 7:55–68, 2010.
- [12] C Jaedicke, A Solheim, LH Blikra, K Stalsberg, A Sorteberg, A Aaheim, K Kronholm, D Vikhamar-Schuler, K Isaksen, K Sletten, et al. Spatial and temporal variations of norwegian geohazards in a changing climate, the geoextreme project. *Natural hazards and earth system sciences*, 8(4):893–904, 2008.
- [13] L Fischer, L Rubensdotter, K Sletten, K Stalsberg, C Melchiorre, P Horton, and M Jaboyedoff. Debris flow modeling for susceptibility mapping at regional to national scale in norway. In *Proceedings of the 11th international and 2nd North American symposium on landslides*, pages 3–8, 2012.
- [14] MH Derron and K Sletten. Method for the susceptibility mapping of snow avalanches in norway. Technical report, Technical report, 2016.
- [15] Anders Solheim, Kjetil Sverdrup-Thygeson, and Bjørn Kalsnes. Hazard and risk assessment for early phase road planning in norway. *Natural Hazards*, pages 1–21, 2023.
- [16] Haoyu Luo, Zhongqiang Liu, Yutao Pan, and Irene Rocchi. Gis-based rainfall-induced landslide susceptibility mapping: a comparative analysis of machine learning algorithms and a numerical method in kvam, norway. Technical report, Copernicus Meetings, 2023.
- [17] Israr Ullah, Bilal Aslam, Syed Hassan Iqbal Ahmad Shah, Aqil Tariq, Shujing Qin, Muhammad Majeed, and Hans-Balder Havenith. An integrated approach of machine learning, remote sensing, and gis data for the landslide susceptibility mapping. *Land*, 11(8):1265, 2022.
- [18] Binh Thai Pham, Ataollah Shirzadi, Dieu Tien Bui, Indra Prakash, and MB Dholakia. A hybrid machine learning ensemble approach based on a radial basis function neural network and rotation forest for landslide susceptibility modeling: A case study in the himalayan area, india. *International Journal of Sediment Research*, 33(2):157–170, 2018.
- [19] Abdelaziz Merghadi, Ali P Yunus, Jie Dou, Jim Whiteley, Binh ThaiPham, Dieu Tien Bui, Ram Avtar, and Boumezbeur Abderrahmane. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207:103225, 2020.
- [20] Binh Thai Pham, Dieu Tien Bui, Indra Prakash, and MB Dholakia. Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide

susceptibility assessment at himalayan area (india) using gis. *Catena*, 149:52–63, 2017.

- [21] Wei Chen, Himan Shahabi, Shuai Zhang, Khabat Khosravi, Ataollah Shirzadi, Kamran Chapi, Binh Thai Pham, Tingyu Zhang, Lingyu Zhang, Huichan Chai, et al. Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. *Applied Sciences*, 8(12):2540, 2018.
- [22] Xuan Luan Truong, Muneki Mitamura, Yasuyuki Kono, Venkatesh Raghavan, Go Yonezawa, Xuan Quang Truong, Thi Hang Do, Dieu Tien Bui, and Saro Lee. Enhancing prediction performance of landslide susceptibility model using hybrid machine learning approach of bagging ensemble and logistic model tree. *Applied Sciences*, 8(7):1046, 2018.
- [23] Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.
- [24] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [25] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [26] Chao Fan, Cheng Zhang, Alex Yahja, and Ali Mostafavi. Disaster city digital twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management*, 56:102049, 2021.
- [27] Shuo Wang, Yanran Li, Jiang Zhang, Qingye Meng, Lingwei Meng, and Fei Gao. Pm2.5-gnn: A domain knowledge enhanced graph neural network for pm2.5 forecasting. In *Proceedings of the 28th international conference on advances in geographic information systems*, pages 163–166, 2020.
- [28] Bruno Scalzo Dees, Yao Lei Xu, Anthony G Constantinides, and Danilo P Mandic. Graph theory for metro traffic modelling. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–5. IEEE, 2021.
- [29] Hamed Farahmand, Yuanchang Xu, and Ali Mostafavi. A spatial-temporal graph deep learning model for urban flood nowcasting leveraging heterogeneous community features. *Scientific Reports*, 13(1):6768, 2023.
- [30] Ian W McBrearty and Gregory C Beroza. Earthquake location and magnitude estimation with graph neural networks. In *2022 IEEE International Conference on Image Processing (ICIP)*, pages 3858–3862. IEEE, 2022.
- [31] Norwegian Water Resources and Energy Directorate. Skredhendelser - WMS. <https://nve.geodataonline.no/arcgis/services/SkredHendelser/MapServer/WMServer?request=GetCapabilities&service=WMS>, 2018. Accessed: 2023-05-01.

- [32] Norwegian Water Resources and Energy Directorate. Avalanche steepness map - WMS. <https://nve.geodataonline.no/arcgis/services/Bratthet/MapServer/WMServer?request=GetCapabilities&service=WMS>, 2010. Accessed: 2023-05-01.
- [33] Martina Wilde, Andreas Günther, Paola Reichenbach, Jean-Philippe Malet, and Javier Hervás. Pan-european landslide susceptibility mapping: Elsus version 2. *Journal of maps*, 14(2):97–104, 2018.
- [34] Andreas Günther, Miet Van Den Eeckhaut, Jean-Philippe Malet, Paola Reichenbach, and Javier Hervás. Climate-physiographically differentiated pan-european landslide susceptibility assessment using spatial multi-criteria evaluation and transnational landslide information. *Geomorphology*, 224:69–85, 2014.
- [35] Klaus Duscher, Andreas Günther, Andrea Richts, Patrick Clos, Uta Philipp, and Wilhelm Struckmeier. The gis layers of the “international hydrogeological map of europe 1: 1,500,000” in a vector format. *Hydrogeology journal*, 23(8):1867, 2015.
- [36] European Space Agency. Globcover 2009, 2010.
- [37] HI Reuter. A europe-wide digital elevation model based on srtm and russian topographic contours. *Data set and documentation for the contract*, 2007.
- [38] Cristian Lussana, Ole Einar Tveito, Andreas Dobler, and Ketil Tunheim. senorge_2018, daily precipitation, and temperature datasets over norway. *Earth System Science Data*, 11(4):1531–1551, 2019.
- [39] T. M. Saloranta. Simulating snow maps for norway: description and statistical evaluation of the senorge snow model. *The Cryosphere*, 6(6):1323–1337, 2012.
- [40] Statistics Norway. Urban settlements 2022. <https://kartkatalog.geonorge.no/metadata/tettsteder-2022/28906bc8-a644-43d5-bbdb-33a217f458ef>, 2022. Accessed: 2023-05-01.
- [41] Norwegian Mapping Authority. Inspire transport network road. <https://kartkatalog.geonorge.no/metadata/inspire-transport-network-road/a76704df-fdd4-4e09-92ac-a7ad1a77b328>, 2018. Accessed: 2023-05-01.
- [42] Han Yang, Kaili Ma, and James Cheng. Rethinking graph regularization for graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 4573–4581, 2021.
- [43] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.
- [44] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17:395–416, 2007.

- [45] Saskia de Vilder, Chris Massey, Biljana Lukovic, Tony Taig, and Regine Morgenstern. What drives landslide risk? disaggregating risk analyses, an example from the franz josef glacier and fox glacier valleys, new zealand. *Natural Hazards and Earth System Sciences*, 22(7):2289–2316, 2022.
- [46] T Taig. Guidelines for doc on dealing with natural hazard risk. *Cheshire (GB), TTAC Ltd.*
- [47] T. Taig, C. Massey, and T. Webb. Canterbury earthquakes port hills slope stability: Principles and criteria for the assessment of risk from slope instability in the Port Hills, Christchurch, 2012. GNS Science Consultancy Report 2011/319.

Appendix A

Supplementary Information

A.1 Trained Ensemble Model

Table A.1: Feature selection for 32 machine-learning models of the ensemble. Legend: A - ELSUS susceptibility, B - steepness, C - total rainfall, D - mean temperature, E - snow depth, F - snow water equivalent, G - fresh snow water equivalent, H - slope angle, I - lithology class, and J - land cover class.

Model Number	Feature Selection Set Index	A	B	C	D	E	F	G	H	I	J
1	10	✓	✓	✓	✓				✓	✓	✓
2	1	✓	✓	✓	✓	✓	✓		✓		
3	8	✓	✓	✓	✓			✓	✓	✓	
4	1	✓	✓	✓	✓	✓	✓		✓		
5	4	✓	✓	✓	✓	✓			✓		✓
6	4	✓	✓	✓	✓	✓			✓		✓
7	2	✓	✓	✓	✓	✓		✓	✓		
8	1	✓	✓	✓	✓	✓	✓		✓		
9	7	✓	✓	✓	✓			✓		✓	
10	11	✓	✓		✓		✓		✓	✓	✓
11	3	✓	✓	✓	✓	✓			✓	✓	
12	1	✓	✓	✓	✓	✓	✓		✓		
13	9	✓	✓	✓	✓			✓	✓		✓
14	1	✓	✓	✓	✓	✓	✓		✓		
15	9	✓	✓	✓	✓			✓	✓		✓
16	10	✓	✓	✓	✓				✓	✓	✓
17	9	✓	✓	✓	✓			✓	✓		✓
18	10	✓	✓	✓	✓				✓	✓	✓
19	11	✓	✓		✓		✓		✓	✓	✓

Table A.1 continued from previous page

Model Number	Feature Selection Set Index	A	B	C	D	E	F	G	H	I	J
20	4	✓	✓	✓	✓	✓			✓		✓
21	10	✓	✓	✓	✓				✓	✓	✓
22	9	✓	✓	✓	✓			✓	✓		✓
23	7	✓	✓	✓	✓		✓		✓		✓
24	11	✓	✓		✓		✓		✓	✓	✓
25	10	✓	✓	✓	✓				✓	✓	✓
26	10	✓	✓	✓	✓				✓	✓	✓
27	5	✓	✓	✓	✓		✓	✓	✓		
28	2	✓	✓	✓	✓	✓		✓	✓		
29	7	✓	✓	✓	✓		✓		✓		✓
30	9	✓	✓	✓	✓			✓	✓		✓
31	4	✓	✓	✓	✓	✓			✓		✓
32	4	✓	✓	✓	✓	✓			✓		✓

A.2 Efficient Multi-Step Hyperparameter Search

Table A.2: Step 1 - Hyperparameter search for feature selection at different random seeds of weight initialization.

Random Seed	100	1	51	68	94	27	84	47	32	12
ELSUS V2 susceptibility	✓		✓	✓	✓	✓	✓	✓	✓	✓
steepness	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
total rainfall	✓	✓	✓		✓	✓	✓	✓	✓	✓
mean temperature	✓	✓	✓	✓	✓	✓	✓	✓	✓	
snow depth	✓				✓	✓	✓		✓	✓
snow water equivalent		✓	✓	✓		✓		✓	✓	✓
fresh snow water equivalent		✓			✓	✓			✓	✓
slope angle class	✓	✓	✓	✓	✓	✓	✓	✓		
lithology class		✓	✓	✓				✓	✓	
land cover class	✓	✓	✓	✓	✓	✓	✓			✓
Number of Features	7	7	8	7	8	9	7	7	7	7
Training AUC	89.48	85.49	90.49	90.43	86.88	90.75	89.16	89.71	90.54	87.23
Validation AUC	81.40	84.34	87.60	87.44	89.32	88.48	87.32	86.33	84.89	86.03
Test AUC	86.93	81.94	88.38	90.83	79.87	85.76	83.91	83.98	84.89	88.43
Epoch	49	23	43	67	20	47	37	37	41	44

Table A.3: Step 2 - Hyperparameter search for interaction of two adjacency matrices at different random seeds of weight initialization.

Random Seed	7	18	29	35	48	57	69	78	85	91
0 or Aattribute	84.65%	85.96%	85.45%	85.16%	83.56%	86.12%	84.27%	85.60%	84.86%	85.13%
0.1	84.65%	85.96%	85.45%	85.16%	83.56%	86.12%	84.27%	85.60%	84.86%	85.13%
0.2	84.65%	85.96%	85.45%	85.16%	83.56%	86.12%	84.27%	85.60%	84.86%	85.13%
0.3	84.65%	85.96%	85.45%	85.16%	83.56%	86.12%	84.27%	85.60%	84.86%	85.13%
0.4	84.65%	85.96%	85.45%	85.16%	83.56%	86.12%	84.27%	85.60%	84.86%	85.13%
0.5	83.28%	84.49%	85.38%	84.47%	82.62%	84.16%	82.82%	84.17%	82.94%	82.22%
0.6	73.41%	73.24%	72.00%	73.25%	72.92%	73.36%	73.47%	72.48%	72.77%	72.96%
0.7	71.58%	71.49%	68.93%	70.43%	70.21%	68.81%	71.84%	68.64%	70.13%	60.43%
0.8	68.96%	69.39%	55.81%	69.12%	67.28%	62.12%	63.83%	68.05%	67.71%	49.57%
0.9	69.52%	70.55%	53.80%	71.94%	68.60%	65.54%	64.16%	70.60%	70.17%	50.78%
1 or Aneighborhood	54.26%	64.39%	61.22%	59.62%	54.86%	51.19%	58.29%	58.36%	61.16%	51.04%
AND	84.55%	85.72%	85.47%	85.30%	83.68%	86.00%	84.38%	85.83%	84.39%	84.96%

Table A.4: Step 3 - Hyperparameter search for the distance and cosine similarity thresholds at different random seeds of weight initialization.

Random Seed	9	17	25	35	47	54	69	75	84	95
optimal distance	12000	12000	12000	12000	12000	12000	20000	18000	12000	12000
optimal cosine similarity	any									

A.3 Exposure Probabilities on December 30, 2020

Table A.5: 190 settlements or villages in Oslo-Viken.

Village	Intra	Inter	Population
Askgrenda	82.73%	82.30%	522
Askim	21.03%	20.73%	14651
Aulifeltet	25.02%	25.33%	2,875
Aursmoen	15.40%	15.92%	3493
Berger	57.26%	58.43%	1110
Bjertnestunet	9.50%	9.65%	415
Bjørkelangen	25.25%	25.74%	4084
Blaker	26.04%	26.83%	480
Bomansvik	25.58%	25.53%	233
Borgen	27.51%	28.50%	1477
Brårud	15.50%	15.37%	461
Darbu	63.10%	62.95%	630
Degernes	8.92%	15.00%	324
Drammen	63.91%	64.03%	111036
Drøbak	39.59%	40.04%	13675
Eidsvoll	29.79%	30.94%	6530
Eltonåsen	49.19%	49.25%	1,696
Fagerliåsen/Poverudbyen	81.93%	82.80%	390
Fagerstrand	26.62%	30.10%	2,885
Fetsund-Østersund	44.90%	46.57%	8,966
Finnestad	18.31%	18.50%	220
Fjellfoten	9.70%	9.70%	1,328
Fjellsrud	36.96%	37.74%	936
Fjellstad	47.27%	49.80%	1059
Fjellstrand	25.45%	24.73%	1321
Flateby	34.66%	35.19%	3803
Fredrikstad/Sarpsborg	13.53%	13.45%	118992
Frogner	57.36%	57.73%	1,555
Fuglevik	6.73%	6.80%	495
Geilo	36.56%	36.98%	2020
Geilolie	31.56%	35.90%	552
Glosli	9.13%	9.13%	773
Gol	56.46%	57.54%	2954
Gomnes	19.79%	19.98%	373
Grua	63.03%	62.65%	1532
Grønlund	51.24%	51.20%	5,208
Haga	18.25%	18.35%	701
Hakadal	34.32%	33.70%	846

Table A.5 continued from previous page

Village	Intra	Inter	Population
Halden	17.66%	18.01%	25,916
Hallingby	58.70%	59.70%	889
Halmstad	6.05%	6.13%	3,408
Hamnås	22.52%	21.70%	234
Harestua	37.61%	38.00%	2,426
Hasle	19.66%	20.20%	694
Hauge	15.89%	15.85%	407
Heiås	19.65%	19.27%	896
Helglandsmoen	14.80%	14.88%	574
Hellerud	64.24%	64.37%	280
Hemnes	8.85%	8.83%	766
Hemsedal	65.45%	65.62%	944
Hogsetfeltet	22.75%	23.10%	912
Hol	58.02%	58.00%	297
Holm	8.08%	8.20%	253
Holmsbu	57.63%	57.20%	289
Hvittingfoss	83.22%	83.44%	1,143
Hyggen	58.19%	58.98%	762
Hølen	28.95%	28.95%	1060
Hønefoss	42.53%	43.28%	16547
Høysand	13.81%	14.03%	737
Ise	14.44%	14.67%	718
Isebakke	12.57%	15.00%	861
Jansberg	24.55%	24.10%	263
Jelsnes	34.92%	35.70%	326
Jessheim	20.57%	20.45%	22953
Jevnaker	52.58%	53.06%	5077
Kampå	17.48%	0.1734	959
Karlshus	9.95%	9.94%	2327
Kirkebygda	46.16%	45.85%	596
Kirkebygden	19.97%	20.05%	1484
Kirkegrenda	7.70%	7.80%	329
Klokkarstua	65.71%	65.87%	720
Kløfta	35.04%	35.71%	8,215
Konglungen	12.90%	15.30%	270
Kongsberg	70.83%	71.23%	22,568
Konnerud	82.07%	80.67%	10,403
Kroksund	48.72%	51.40%	342
Krakkstad	13.47%	13.55%	1,121
Krøderen	62.26%	62.45%	570
Lampeland	70.15%	70.50%	588

Table A.5 continued from previous page

Village	Intra	Inter	Population
Langset	33.99%	34.00%	378
Larkollen	6.86%	7.27%	1,507
Leirsund	61.27%	61.17%	1,465
Lervik	12.13%	12.89%	3,343
Lindeberg	47.94%	48.67%	1,232
Lunde	13.71%	13.70%	475
Lundermoen	56.10%	52.95%	1,455
Lunner	52.41%	52.66%	971
Løken	22.55%	22.38%	1,645
Løkenfeltet	52.37%	52.25%	631
Lørenfallet	36.50%	36.58%	1,110
Løstad	41.57%	41.80%	2,675
Maura	30.25%	29.92%	4,534
Minnesund	34.93%	35.47%	544
Missingmyr	16.69%	17.65%	929
Momoen	32.76%	32.98%	508
Moss	20.87%	20.08%	48,477
Movatn	67.04%	66.80%	312
Mysen	27.75%	27.97%	6,882
Møvik	10.38%	10.13%	244
Nakkerud	24.65%	33.60%	355
Nesbyen	66.28%	66.70%	2,129
Nesbygda	44.98%	47.80%	1,020
Neskollen	10.42%	10.50%	2,825
Nesoddtangen	13.78%	14.94%	13,189
Nordal	79.85%	80.00%	329
Norderhaug	9.74%	10.03%	603
Norderhov	28.79%	28.45%	314
Nordkisa	13.05%	13.33%	1,287
Noresund	67.30%	67.08%	349
Nærnes	38.05%	40.47%	1,593
Oddevall/Sjåstad	89.42%	89.00%	421
Opakermoen	6.38%	6.48%	527
Ormåsen	50.19%	48.63%	1,899
Oslo	48.78%	51.00%	1,064,235
Prestegårdshagen	34.89%	35.17%	379
Prestfoss	70.73%	71.90%	488
Rakkestad	16.06%	16.43%	4,855
Ringvoll	16.13%	16.03%	530
Roa	58.24%	58.45%	957
Rostadneset	15.29%	15.50%	280

Table A.5 continued from previous page

Village	Intra	Inter	Population
Rotnes	59.82%	61.10%	7,452
Rustad	35.42%	35.40%	322
Råholt	22.91%	23.06%	14,830
Rød	12.82%	12.94%	687
Rødberg	87.46%	89.24%	495
Røyken	70.98%	71.10%	4,005
Sand	31.31%	32.13%	2,443
Sandvoll	44.14%	46.10%	269
Sessvollmoen	20.15%	20.22%	1,283
Setskog	15.57%	22.70%	203
Siggerud	48.13%	48.50%	1,493
Ski	29.57%	29.46%	20,835
Skiptvet	14.59%	15.10%	1,829
Skivika	9.73%	9.90%	1,075
Skjeberg	11.28%	11.37%	1,522
Skjærhalden	7.14%	8.77%	936
Skjønhaug	21.38%	22.80%	2,094
Skogrand	7.54%	8.50%	329
Skotbu	20.79%	20.60%	462
Skotselv	72.54%	72.58%	729
Slevik	21.70%	22.36%	1,166
Slitu	20.35%	20.63%	778
Smestad	48.13%	52.40%	2,721
Sokna	51.99%	54.80%	616
Solbakken-Sofienberg	12.70%	13.60%	218
Spetalen	8.54%	10.40%	2,318
Sponvika	14.97%	24.80%	538
Spydeberg	20.72%	20.24%	6,419
Steinsåsen	26.44%	26.00%	2,045
Stikkaåsen	16.76%	16.50%	209
Sundvollen	55.65%	56.23%	1,171
Svelvik	55.99%	56.62%	4,089
Svene	73.67%	73.68%	373
Svinndal	21.14%	21.87%	582
Svøo	52.40%	52.58%	606
Sylling	80.93%	81.00%	747
Sysle	55.63%	55.20%	246
Såstadbråten	7.45%	7.63%	428
Sætre	47.76%	48.87%	6,202
Sørumsand	42.94%	42.70%	5,596
Teigebyen	31.12%	31.60%	3,280

Table A.5 continued from previous page

Village	Intra	Inter	Population
Tofte	36.13%	38.13%	3,064
Tomb	14.99%	15.00%	284
Tomter	21.26%	21.55%	2,112
Tomteråsen	12.83%	12.80%	727
Torget	34.83%	35.56%	765
Torpo	53.50%	53.83%	375
Torrvik	24.48%	24.75%	408
Tranby	76.15%	75.90%	5,296
Trømborg	21.57%	22.70%	242
Tyristrand	31.12%	34.90%	934
Utgård	20.08%	20.40%	324
Vang	59.65%	59.41%	2,312
Veggli	87.78%	87.90%	362
Vestby	24.73%	25.41%	8,199
Vestfossen	57.28%	56.60%	3,246
Vikersund	53.11%	53.86%	3,171
Vormsund	9.19%	9.10%	428
Våk	16.42%	16.40%	944
Ytre Enebakk	37.06%	37.24%	4,528
Åkrene	46.02%	47.00%	375
Ål	64.06%	64.36%	2,346
Åmot/Geithus	67.25%	67.11%	6,839
Åneby	49.41%	49.35%	1,501
Årnes	8.61%	8.85%	4,918
Ås	17.29%	17.25%	10,762
Åsgrenda	44.60%	42.15%	1,456
Ørje	14.79%	15.25%	1,849
Øyenkilen	13.14%	22.40%	467