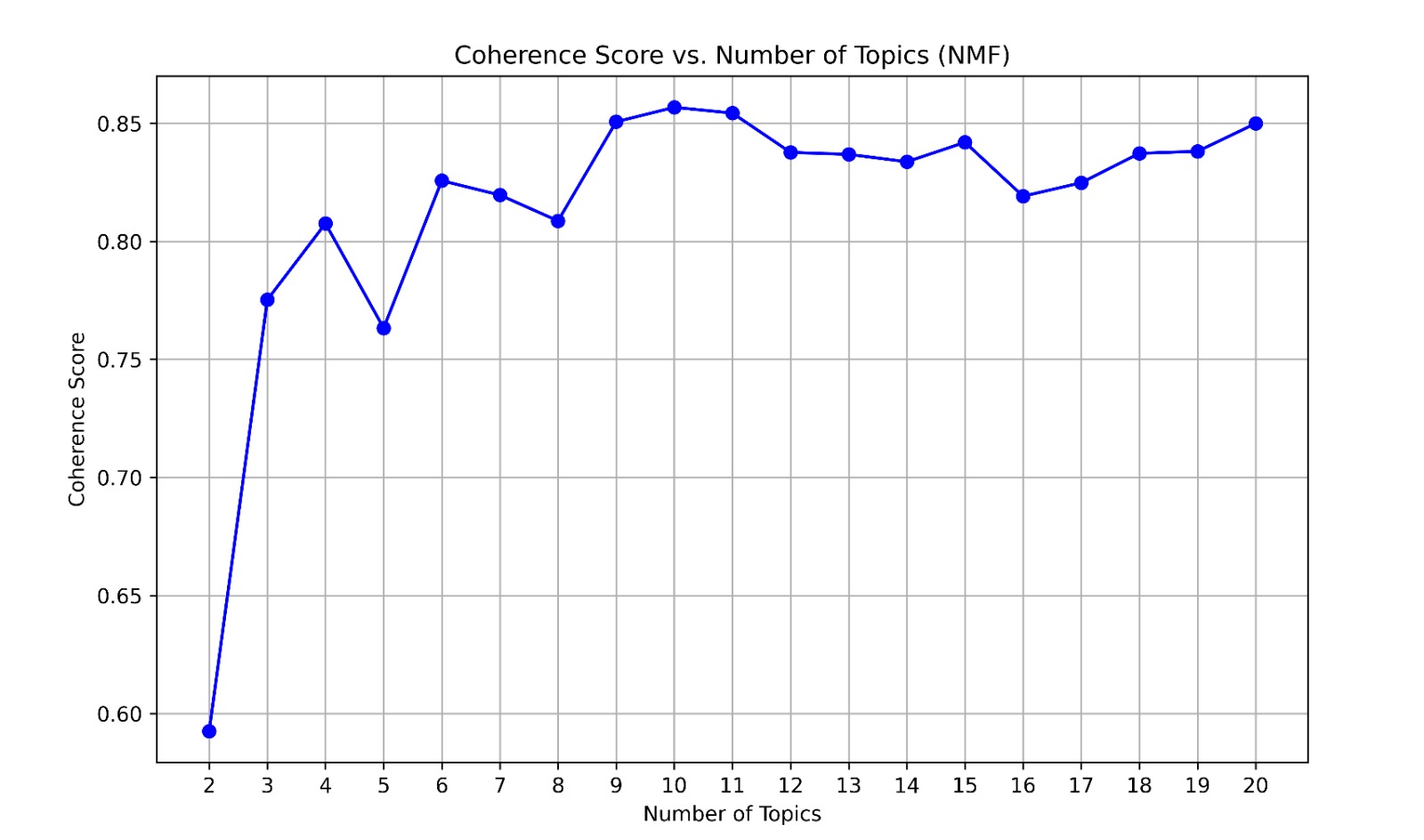
*Supplement of*

**An automated approach for developing geohazard inventories using news: Integrating NLP, machine learning, and mapping.**

Avcioglu et al.

Correspondence to: Aydogan Avcioglu ([a.avcioglu@brgm.fr](mailto:a.avcioglu@brgm.fr))



**Figure S1: The evolution of coherence score by increasing number of topics within the non-negative matrix factorization (NMF) algorithm.**

**Table S1: A sample set of an unrefined dataframe of scraped newspapers through the tr-news-scraper tool online for multiple hazards; flood, landslide, wildfire, and sinkhole formation.**

**Table S2: The frequency of the words both in English and Turkish within the WordClouds for different geohazards.**

**Table S3: The examples for news identified as an incident of geohazards and non-incident news with the 4 determined groups: common words, misinterpretation, review news, and warning alerts.**