Messfehler Zusammenfassung

Wie bereits zuvor erwähnt, haben die Projektpartner festgestellt, dass die Checkpoints in ihrer Messung Messfehler haben. Messfehler sind Abweichungen der erhobenen Daten vom wahren Wert.

*(Quelle: https://www.wortbedeutung.info/Messfehler/)*

In unserem Fall bedeutet das, dass die Checkpointmessungen nicht exakt den vorbeigehenden Personen mit oder ohne LVS-Gerät entsprechen, somit ist eine mögliche Unter-/Überschätzung des Anteils an LVS-Geräten vorhanden. Dabei können Messfehler verschiedener Art vorkommen, die jeweils eigene Ursachen haben, und so die Stärke und Richtung der Fehlschätzung unterschiedlich beeinflussen. Im folgenden Teil werden erst von den Projektpartnern gesammelte Daten über die Genauigkeit der Checkpointmessungen an zwei Tagen untersucht und danach (unter anderem) daraus verschiedene Szenarien für mögliche Messfehler erstellt. Anhand des jeweiligen Szenarios werden Messungen aus dem originalen Datensatz ergänzt oder entfernt und es wird verglichen, inwiefern sich die Ergebnisse unserer zwei Modelle für die gewählten Szenarien unterscheiden.

Messfehler Beschreibung

Um die Schwere und Art der Messfehler zu untersuchen, wurde eine Gruppe von Studenten mehrmals zum Aufstiegsort geschickt. Unter Anderem wurden dabei verschiedene Tests angestellt, z.B. wurde die Erfassung der Personen abhängig von unterschiedlichen Entfernungen und Winkeln zum Checkpoint untersucht oder man ließ einen Hund mehrmals am Checkpoint vorbeilaufen und markierte, wann dieser vom Checkpoint registriert wurde und wann nicht. Für die folgende Analyse werden Daten benutzt, die von einer manuellen Zählung einer solchen Gruppe von Studenten am 27.02. und 28.02.2020 stammen. Genauer gesagt fanden die Zählungen zu drei unterschiedlichen Zeiträumen statt: am 27.02. von 10:32 bis 12:58, am 28.02. von 10:38 bis 12:26 und von 14:40 bis 16:50. Dabei wurde eingetragen, wie viele mögliche Kontakte zu jeder Minute am Checkpoint vorbeigegangen sind. Zudem werden die Kontakte auf zwei verschiedene Arten unterschieden. Einerseits gibt es eine Unterscheidung zwischen „Skitourengeher“, die eigentlich zu interessierende Art von Besuchern mit eventuellen LVS-Geräten, und „andere Kontakte“, also andere Besucher wie Spaziergänger oder z.B. ein Hund. Andererseits wird zusätzlich zwischen „erfassten“ und „nicht erfassten“ Kontakten unterschieden, wobei ein Kontakt als erfasst gilt, wenn er augenscheinlich vom Checkpoint erkannt und registriert wurde. Wichtig zu erwähnen ist, dass bei der manuellen Zählung keine Unterscheidung zwischen Besuchern mit und ohne LVS-Gerät gemacht wurde. Zusätzlich zu den manuellen Zählungen liegen die tatsächlich von den Checkpoints erhobenen Messungen zur Saison 19/20 vor, wobei einer der Checkpoints früh in der Saison ausgefallen ist und zu den zwei Tagen der manuellen Zählungen somit nur ein Checkpoint aktiv war.

Messfehler deskriptive Analyse

Die Anzahl der in der manuellen Zählung eingetragenen Kontakte beträgt insgesamt 208. Davon sind 122 Skitourengeher und 86 andere Kontakte. Als erfasst vom Checkpoint wurden 161 Kontakte eingetragen, als nicht erfasst somit 47.

Da die angegebenen Uhrzeiten in der manuellen Zählung und der tatsächlichen Registrierung im Checkpoint voneinander abweichen, werden beim genaueren Vergleich von Erfassungen und Nichterfassungen nur die Daten der manuellen Zählung hergenommen. Für den gesamten Zeitraum der manuellen Zählung kann man jedoch die tatsächlich vom Checkpoint eingetragene Anzahl an erfassten Kontakten hernehmen, die mit 170 etwas höher ist als die 161 von den Studenten angegebenen Nichterfassungen. Für die Berechnung der Unterschätzung insgesamt wird deshalb die Anzahl der vom Checkpoint gemessenen Kontakte, also 170, mit den insgesamt durch die manuelle Zählung festgehaltenen Kontakten, also 208, verglichen. Damit kommt man auf eine allgemeine Unterschätzung des Checkpoints von ca. 22%. Hierbei wird „Unterschätzung“ definiert als der Anteil an vorhandenen Messungen, die man dem Checkpoint hinzufügen müsste, um auf den wahren Wert der vorbeigehenden Kontakte zu bekommen. In unserem Fall müsste man dem Checkpoint ca. 22% von 170 erfassten Kontakten hinzufügen, um auf den wahren Wert von 208 Kontakten zu kommen.

Für die Berechnung der Unterschätzung wurden keine Messungen aus den manuellen Zählungen entfernt. Insbesondere auch keine Fälle, zu denen angegeben war, dass das Gerät durch fälschliche Messungen wie Mehrfacherfassungen einer Person den wahren Wert überschätzt. In diesem Sinne ist die berechnete Unterschätzung keine „reine“ Unterschätzung sondern die Unterschätzung für alle gemessenen Beobachtungen.

Um zu testen, wie genau die Erfassung des Checkpoints bei größeren Gruppen an vorbeigehenden Kontakten funktioniert, wurde eine Untersuchung des Anteils an erfassten Kontakten in Abhängigkeit von Gruppengröße durchgeführt. Als Gruppe zählen hierbei alle Kontakte, die zur selben Uhrzeit in der manuellen Zählung vermerkt wurden. Sind also beispielsweise um 15:23 vier Kontakte in der manuellen Zählung vermerkt, so gilt dies als Gruppe von vier Personen. In den Daten kommen jeweils mehrere Gruppen in einer Größe von 1 bis 5 und eine Gruppe von 8 Kontakten vor. Wie man in der Abbildung sieht, schwankt der Anteil von Nichterfassungen an allen Messungen bei den fünf kleineren Gruppengrößen um einen Wert von 18%, der einer Unterschätzung von 22% entspricht. (0.18/0.82=0.22) In der Achtergruppe ist der Anteil von Nichterfassungen allerdings über 50%, d.h. es wurden in dieser Gruppe mehr Kontakte vom Checkpoint nicht erkannt als erkannt.

Messfehlerszenarien Einleitung

Ursprünglich war mit den Projektpartnern vereinbart, die Daten zu den Messfehlern herzunehmen, um das bereits aufgestellte Modell auf Grundlage der Checkpointdaten der Saison 18/19 anzupassen und das neue, verbesserte Modell an den Daten zur Saison 19/20 zu testen. Da das Frühjahr 2020, zu dem die manuellen Zählungen geplant waren, von den Wetterbedingungen sehr milde war und somit vergleichsweise sehr wenige Besucher am Untersuchungsort unterwegs waren, ist der Umfang der Daten zu den Messfehlern allerdings deutlich geringer als vorgesehen. Da die meisten Kovariablen im Modell täglich erhoben werden und für die Messfehler nur Daten zu zwei Tagen mit nur insgesamt 208 Messungen vorliegen, ist der Informationsgehalt zu gering für eine sinnvolle Anpassung der Modelle.

Stattdessen wurden in Abstimmung mit den Projektpartnern aus den Erkenntnissen über die Messfehler vier verschiedene Szenarien entworfen. Für jedes Szenario wird der originale Datensatz aus der Saison 18/19 angepasst und die bestehenden Modelle auf den jeweils neu erzeugten Datensatz angewandt. Die Ergebnisse für die unterschiedlichen Szenarien werden mit den originalen Ergebnissen verglichen, um Aussagen darüber treffen zu können, inwiefern sich unterschiedliche Messfehler auf die Ergebnisse des Modells auswirken. Da aus den Tests hervorgegangen ist, dass der Checkpoint bei der Messung von LVS-Geräten sehr genau ist und Messfehler eher bei der Infrarotmessung entstehen, werden in jedem Szenario nur Messungen ohne LVS-Gerät hinzugefügt (bzw. entfernt).

Szenario 1 geht von einer generellen Unterschätzung von 22% aus, nach dem Wert aus der Analyse der Messfehlerdaten. Hier werden dem Datensatz zu zufälligen Zeitpunkten insgesamt 22% neue Beobachtungen von Kontakten ohne LVS-Gerät hinzugefügt.

Szenario 2 geht von einer ansteigenden Unterschätzung bei zunehmender Gruppengröße aus. Hier wird für jede Beobachtung mit einer bestimmten Wahrscheinlichkeit eine neue Beobachtung hinzugefügt, so dass bei größeren Gruppen im Durchschnitt auch mehr neue Beobachtungen hinzukommen.

Szenario 3 geht von einer nächtlichen Überschätzung aus. Die Tests haben gezeigt, dass teilweise auch ein vorbeilaufender Hund vom Checkpoint erfasst wird. Da nachts wenige Menschen unterwegs und besonders in den Dämmerungszeiten vermehrt Tiere unterwegs sind, wird nachts ein fester Anteil an Beobachtungen ohne LVS-Gerät aus dem Datensatz entfernt, der für irrtümlich erfasste Tiere steht.

Szenario 4 geht von einer wachsenden Unterschätzung bei fallenden Temperaturen aus. Hier wird davon ausgegangen, dass der Checkpoint bei größerer Kälte fehleranfälliger wird und es wird für sinkende Temperaturen ein bestimmter Anteil an neuen Beobachtungen hinzugefügt.

Die Ergebnisse für die Szenarien im Vergleich zum originalen Datensatz werden erneut erst für das Tagesmodell, danach für das Zeitmodell verglichen.

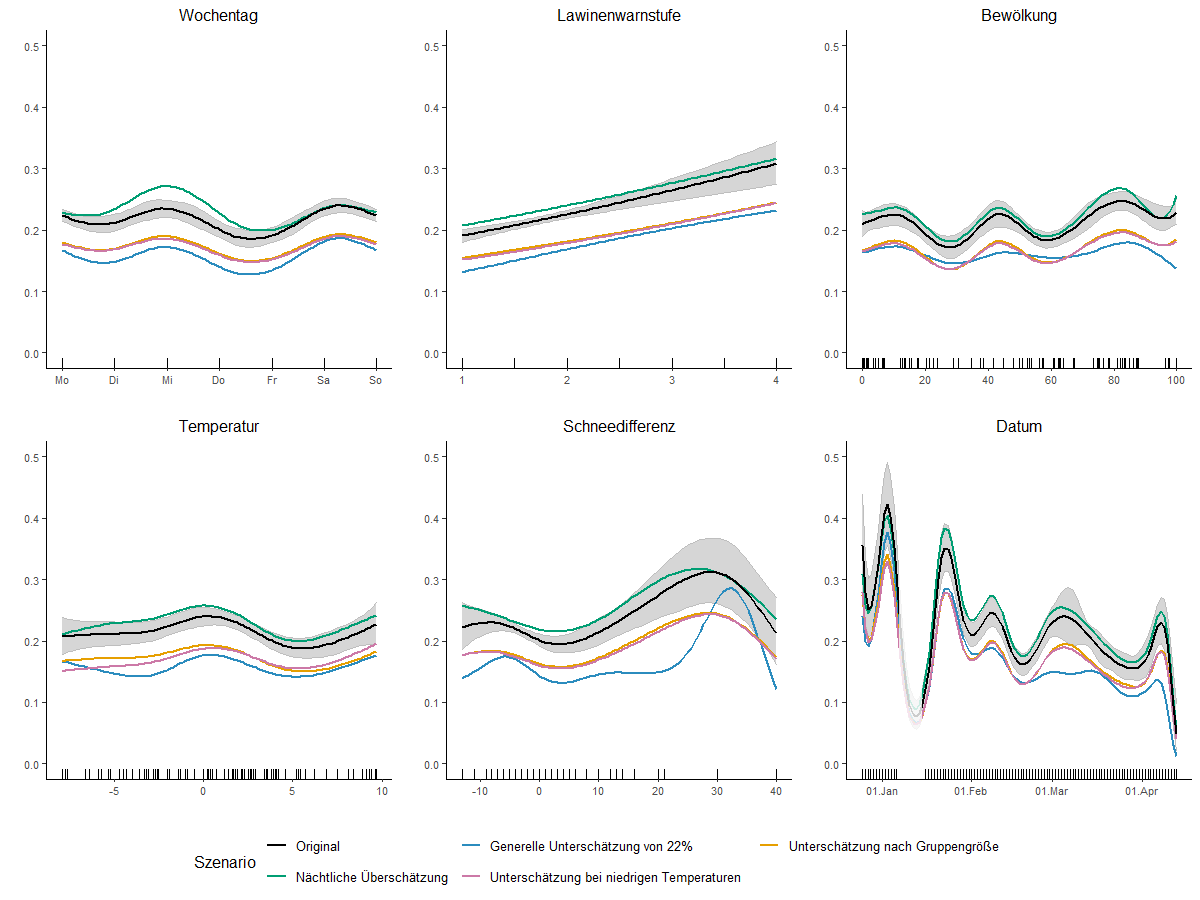
Szenarien im Tagesmodell

TABELLE 1

*Tabelle 1 zeigt im Tagesmodell die Werte für die Koeffizienten von Intercept und Ferientag, die p-Werte und Signifikanzen für alle Kovariablen für den originalen Datensatz und jedes Messfehlerszenario im Vergleich.*

Da in den Szenarien nur Beobachtungen ohne LVS-Gerät hinzugefügt bzw. entfernt werden, sinken bzw. steigen die Koeffizienten für die zwei parametrischen Variablen entsprechend. Szenario 1 mit genereller Unterschätzung hat die höchste Menge an hinzugefügten Beobachtungen, deswegen ist der Intercept, also der Anteil an Messungen mit LVS-Geräten bei konstant gehaltenen Kovariablen und keinem Ferientag, mit einem Wert von 15,8% am geringsten. Szenario 3 mit nächtlicher Überschätzung ist das einzige Szenario, bei dem Beobachtungen entfernt werden, deswegen steigt hier der Koeffizient für den Intercept auf einen Wert von 23,2%. Analog gelten die Aussagen für den Anteil an Ferientagen bei konstant gehaltenen Kovariablen. Bei der generellen Unterschätzung ist der Wert für Ferientage mit 16,2% höher als der Wert für den Intercept.

Bei einem Signifikanzniveau von 5% haben nur der Intercept und das Datum bei allen Szenarien wie im Original einen statistisch signifikanten Zusammenhang mit dem Anteil an LVS-Geräten. Die Kovariable für Lawinengefahr hat im Original und bei jedem Szenario außer der nächtlichen Überschätzung einen statistisch signifikanten Zusammenhang. Alle anderen Kovariablen, als für Ferientag, Bewölkung, Wochentag, Temperatur, Bewölkung und Schneedifferenz, haben wie im Original keinen statistischen Zusammenhang mit dem Anteil auf dem gewählten Signifikanziveau.

ABBILDUNG 1

*Abbildung ??? zeigt im Tagesmodell den Zusammenhang zwischen Anteil an LVS-Geräten und allen nonparametrischen Kovariablen im Original und den verschiedenen Messfehlerszenarien.*

Es lässt sich anhand von Abbildung xyz feststellen, dass die Zusammenhänge der einzelnen Kovariablen mit dem Anteil an LVS-Geräten für alle Szenarien zumeist einen ähnlichen Verlauf nehmen wie im Original. So hat die glatte Funnktion z.B. für den Wochentag bei jedem Szenario wie im Original höhere Werte für Montag, Mittwoch und Samstag als für Dienstag, Donnerstag oder Freitag. Auch für die anderen Kovariablen unterscheidet sich der Verlauf des Zusammenhangs insbesondere für Szenarien 2 bis 4 nur geringfügig vom Original.

Einzig in Szenario 1, der generellen Unterschätzung, sind an mehreren Stellen größere Unterschiede zu bemerken. So schwankt der Anteil im Zusammenhang mit der Bewölkung weniger stark als im Original und den anderen Szenarien. Für niedrige Temperaturen bis -3 Grad sinkt der Anteil im Gegensatz zu den anderen Szenarien, verläuft danach aber ähnlich. Bei einer Schneedifferenz von 5 bis 20 cm steigt der Anteil geringer als in den anderen Szenarien bis nur ca. 15%, für Werte zwischen 20 und 32 cm dagegen steigt der Anteil in Szenario 1 deutlich stärker und sinkt danach auch deutlich stärker ab als in den anderen Szenarien, mit einem Höchstwert von ca. 28%. Wie schon erwähnt ist die Unsicherheit in diesem Bereich allerdings erhöht, da nur wenige Tage mit hohen Werten für die Schneedifferenz vorkommen. Beim Datum unterscheidet sich Szenario 1 im Bereich vom 01. März bis zum15. März, da der Anteil dort fast konstant bei ca. 15% liegt und nicht wie in den anderen Szenarien ansteigt und wieder abfällt.

Szenarien im Zeitmodell

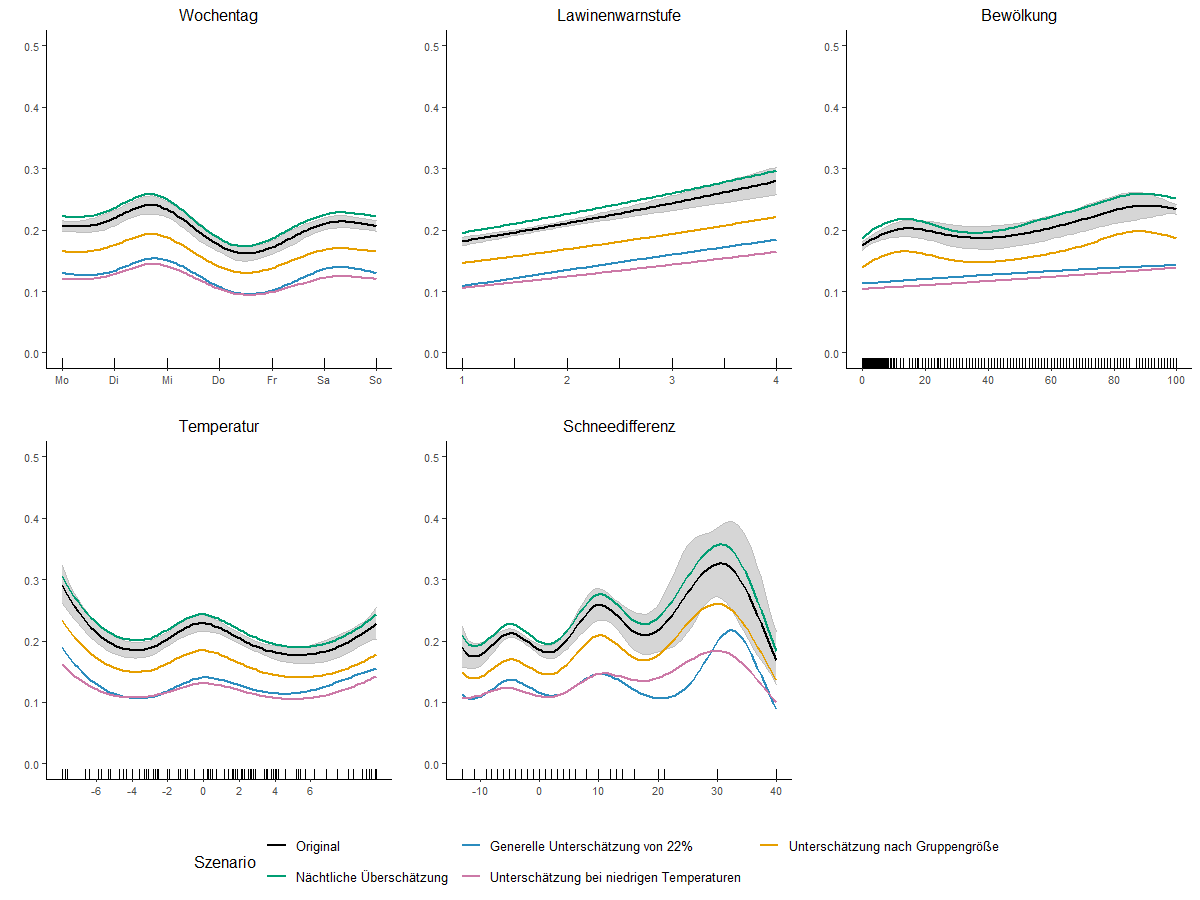
TABELLE 2

*Tabelle 2 zeigt im Zeitmodell die Werte für die Koeffizienten von Intercept und Ferientag, die p-Werte und Signifikanzen für alle Kovariablen für den originalen Datensatz und jedes Messfehlerszenario im Vergleich.*

Ähnlich wie beim Tagesmodell verhalten sich die Koeffizienten für den Intercept und Ferientag. In den Szenarien mit Unterschätzung sinken beide Anteile im Vergleich zum Original bei einem Tiefstwert von 12,7% für den Intercept bei der generellen Unterschätzung und einem erhöhten Wert von 21,8% beim Szenario der nächtlichen Überschätzung, analog für den Anteil an Ferientagen. Bei der generellen Unterschätzung ist der Anteil an Ferientagen mit 12,6% geringer ist als an Tagen, die keine Ferientage sind.

Wie im Original sind in allen Szenarien alle Variablen bis auf Ferientag statistisch signifikant.

ABBILDUNG 2



*Abbildung ??? zeigt im Zeitmodell den Zusammenhang zwischen Anteil an LVS-Geräten und allen nonparametrischen Kovariablen im Original und den verschiedenen Messfehlerszenarien.*

Wie im Tagesmodell verlaufen die Zusammenhänge der einzelnen Kovariablen mit dem Anteil an LVS-Geräten bei allen Szenarien zumeist ähnlich zum Original. Größere Unterschiede bestehen bei der Bewölkung, wo für das Szenario mit genereller Unterschätzung und einer Unterschätzung bei niedrigen Temperaturen ein konstanter Anstieg bei wachsender durchschnittlicher Bewölkung zu beobachten ist, und bei der Schneedifferenz, wo im Bereich von 10 bis 20 cm der Anteil im Szenario mit Unterschätzung bei niedrigen Temperaturen verhältnismäßig weniger Schwankungen unterliegt. Auffällig ist, dass der Anteil bei Werten für die Schneedifferenz von 20 cm und mehr für das Original und Szenarien 2 bis 4 nun im Vergleich zum Tagesmodell stärker ansteigt und wieder sinkt und somit dem Verlauf von Szenario 1 ähnlicher ist.

Szenario 1 mit unterschiedlichen Anteilen

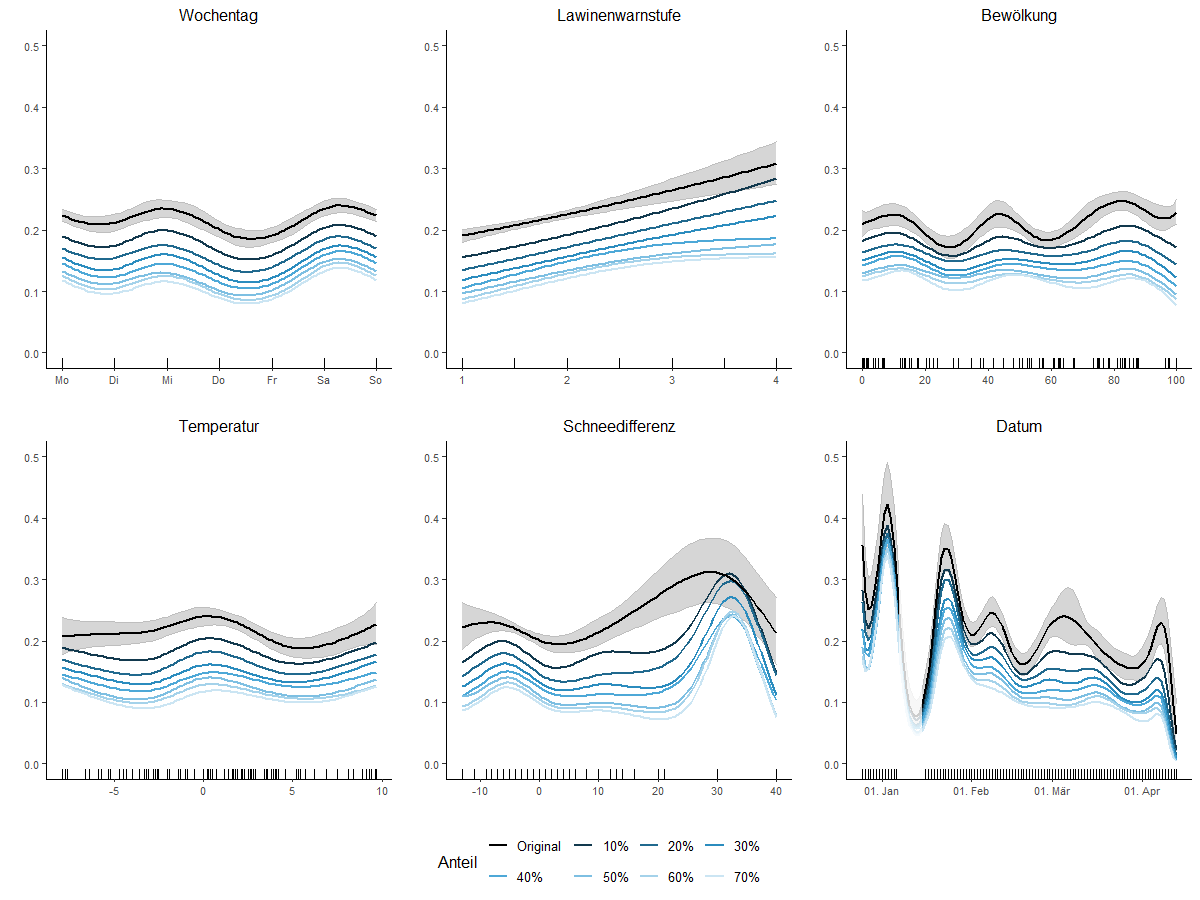
Für das Szenario mit genereller Unterschätzung wird im obigen Vergleich eine Unterschätzung von 22% angenommen. Im Folgenden wird verglichen, wie sich die Ergebnisse der beiden Modelle für unterschiedliche Werte für die allgemeine Unterschätzung verhalten.

Szenario 1 im Tagesmodell

TABELLE 3

Da für steigende Unterschätzung ein größerer Anteil an Beobachtungen ohne LVS-Geräte hinzugefügt wird, sinken die Werte für Intercept und Ferientag. Einen höheren Anteil an Ferientagen gibt es bei einer Unterschätzung von 20%, 30% und 50%. Dies wird allerdings hauptsächlich davon beeinflusst, an welchen Tagen neue Beobachtungen zufällig generiert werden. Wie im Original sind in jedem Fall der Intercept, Datum und Lawinengefahr statistisch signifikant, ab einer Unterschätzung von 30% ist auch Wochentag signifikant.

ABBILDUNG 3



*Abbildung ??? zeigt im Tagesmodell den Zusammenhang zwischen Anteil an LVS-Geräten und allen nonparametrischen Kovariablen im Original und steigenden Werten für den Anteil an Unterschätzung.*

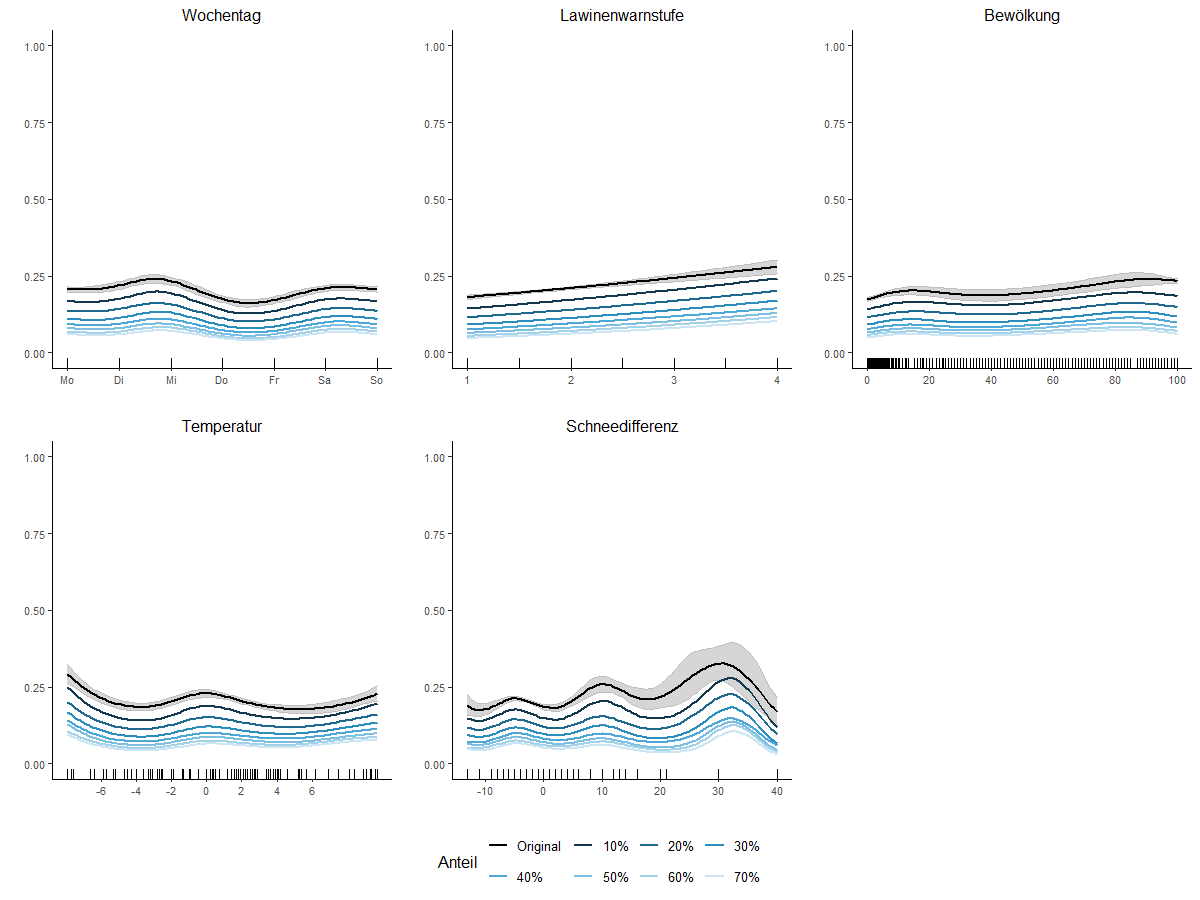
Die Verläufe des Zusammenhangs zwischen Anteil an LVS-Geräten und den einzelnen Kovariablen sind für alle Anteile an Unterschätzung sehr ähnlich. Eine kleine Auffälligkeit ist ein größeres Abflachen des Anteils an LVS-Geräten bei einer Lawinenwarnstufe ab 3 bei Anteilen an Unterschätzung von 40% und höher. Unterschiede zum Original verhalten sich bei allen Werten für Unterschätzung wie vorher für eine Unterschätzung von 22% beschrieben.

Szenario 1 im Zeitmodell

TABELLE 4

Analog wie im Tagesmodell sinken die Koeffizienten für Intercept und Ferientag mit wachsender Unterschätzung. Ab einer Unterschätzung von 30% ist der Anteil an LVS-Geräten an Ferientagen geringer. Wie im Original sind für alle Werte für die Unterschätzung alle Kovariablen bis auf Ferientag signifikant, ab einer Unterschätzung von 60% ist auch Ferientag signifikant.

ABBILDUNG 4



*Abbildung ??? zeigt im Zeitmodell den Zusammenhang zwischen Anteil an LVS-Geräten und allen nonparametrischen Kovariablen im Original und den verschiedenen Messfehlerszenarien.*

Wie im Tagesmodell verlaufen die Zusammenhänge zwischen Anteil an LVS-Geräten und den einzelnen Kovariablen für alle Werte an Unterschätzung sehr ähnlich. Größere Unterschiede gibt es weder untereinander noch zum Original.